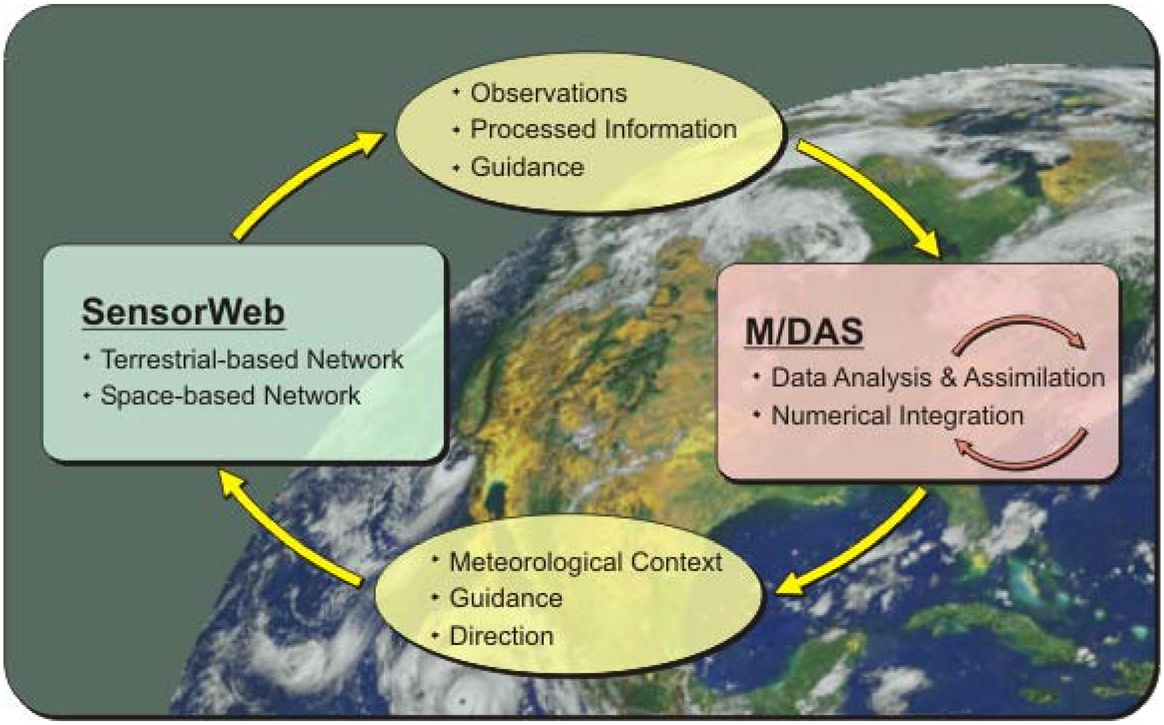
Table of Contents

Advanced Weather Prediction Technologies: NASA's Contribution to the Operational Agencies

Vision Architecture Study



Mark Clausen
Michael Kalb
Gail McConaughy
Ronald Muller
Steven Neeck
Michael Seablom
Mark Steiner

A report prepared for
NASA's Earth Science Technology Office

May 31, 2002

Core Study Team:

Robert Atlas	NASA Goddard Space Flight Center (Science Advisor)
Mark Clausen	TASC/ Northrop Grumman, Inc.
Michael Kalb	Global Science & Technology, Inc.
Gail McConaughy	NASA Goddard Space Flight Center
Ronald Muller	QSS Group, Inc.
Steven Neeck	NASA Goddard Space Flight Center
Michael Seablom	NASA Goddard Space Flight Center
Mark Steiner	NASA Goddard Space Flight Center (Team Leader)

Contributors:

Neil Dennehy	NASA Goddard Space Flight Center
Terrence Doiron	NASA Goddard Space Flight Center
Jaime Esper	NASA Goddard Space Flight Center
James Gatlin	NASA Goddard Space Flight Center
Bruce Gentry	NASA Goddard Space Flight Center
William Heaps	NASA Goddard Space Flight Center
Glenn Higgins	TASC/ Northrop Grumman, Inc.
Peter Hughes	NASA Goddard Space Flight Center
Jesse Leitner	NASA Goddard Space Flight Center
Mark Matsumura	NASA Goddard Space Flight Center
Rick Schnurr	NASA Goddard Space Flight Center

Discussions:

Robert Adler	NASA Goddard Space Flight Center
Joe Ardizzone	Science & Applied Information Corporation
Steve Bloom	Science & Applied Information Corporation
Michael Bosilovich	NASA Goddard Space Flight Center
Steve Cohn	NASA Goddard Space Flight Center
Arlindo da Silva	NASA Goddard Space Flight Center
John Derber	Oceanic & Atmospheric Administration
David Emmit	Simpson Weather Associates, Inc.
Ronald Gelaro	NASA Goddard Space Flight Center
Jeffrey Halverson	University of Maryland Baltimore County
Peter Hildebrand	NASA Goddard Space Flight Center
Larry Korb	NASA Goddard Space Flight Center
S.J. Lin	NASA Goddard Space Flight Center
Peter Norris	University of Maryland Baltimore County
John Pietras	Global Science & Technology, Inc.
Paul Poli	University of Maryland College Park
Lars-Pieter Riishogaard	University of Maryland Baltimore County
Mark Schoberl	NASA Goddard Space Flight Center
Siegfried Schubert	NASA Goddard Space Flight Center
Ed Torres	NASA/ESTO
Eric Travis	Global Science & Technology, Inc.
Steve Ungar	NASA Goddard Space Flight Center
Phillip Webster	NASA Goddard Space Flight Center

ABSTRACT

NASA's Earth Science Technology Office (ESTO) appointed NASA's Goddard Space Flight Center to perform a study that would identify science knowledge and technology improvements needed to enable skilled weather forecasts of 10 - 14 days in the 2025 timeframe.

This report describes the (a) goal of the study, (b) science needs derived for the analysis and how they were developed, and (c) development of a notional architecture concept for a future operational system. The architecture concept is used as a starting point for gap analyses that identify technologies needed to provide the technical capabilities needed to enable the proposed improvements in weather forecasting.

The central premise of the study approach is that a global observing system and a forecast modeling system interactively feeding information to one another in real-time could constitute a new type of operational weather forecast system whose skill improvements would represent a leap forward, and whose performance would be self-optimizing for any level of resources invested. The primary capabilities needed to execute this concept for 14-day weather forecasting are a global mesoscale model with 1 – 25 km horizontal resolution, and a global observing system (SensorWeb) providing high time resolution (1 –3 hourly) and space resolution (1- 25 km) global coverage of key geophysical variables. A seamless ground–space communications network linking these two segments will enable both coordinated global observing among spacecraft and delivery of those observations in near real time for integration into the model forecast system.

Among the larger challenges identified is designing a software system that would possess such an unprecedented level of semi-autonomous intelligence, that it would actually be able to make informed "scientific" judgments, weigh priorities, then direct the coordinated tasking of space-based platforms and instruments based on observational needs identified from a ground-based modeling system. All this would have to occur continuously and in real-time in order to meet operational forecast requirements. Greater reliance on high performance on-board computing seems essential for supporting both system intelligence as well as data analysis functions.

This report also offers preliminary ideas on how the entire system must be designed and operated in order to provide the needed coordination between and among space platforms, instruments and ground. Given some notion of the desired interactions, next steps will be to consider in more detail the system logic, architectures and technologies, as well as advances in system theory, communications and artificial intelligence that could provide the necessary interactivity and results from a highly intelligent, highly integrated operational weather forecast system.

1. Introduction

This study report was sponsored by NASA's Earth Science (ES) Technology Office (ESTO) in as part of its continued development of NASA's ES Enterprise Vision. The ES Vision promotes development of advanced scientific and technical capabilities that will enable proactive Earth system prediction of natural phenomena such as climate, weather, and natural hazards. Preliminary work performed during ES Vision workshops conducted at the University of Maryland in the summer of 2000 identified notional research and applications scenarios which, when developed, will help NASA to understand the range of scientific and technical advancements needed to realize the Earth Science Vision. Findings from these studies help identify and prioritize future investment in ES technologies and thus provide focus to for future ES technology development-activities.

1.1 Purpose of Study

The purpose of this study is to identify the science applications and technology improvements needed to enable skilled weather forecasts of 10 - 14 days in the 2025 timeframe. The science inputs reflect largely the views of Dr. Bob Atlas and members of the NASA Data Assimilation Office. This study was organized and funded under the auspices of ESTO.

The following white paper does not cover all aspects of the science justification in the same detail; some areas receive more attention than others. The purpose of this paper is to provide documentation that others may reference at a future date, regarding the thinking and discussions that took place in this study.

1.2 Questions to be Answered

The following questions were formulated early-on by the study team in order to focus the study approach. The team identified scientific and technological challenges to be addressed in order to advance the current state of the art sufficiently to approach skilled weather forecasting at ten to fourteen-days.

- What observations do we need? What Earth observations need to be gathered, and at what time and space scales and accuracies, in order to support a capability to produce operational 10 - 14 day weather forecasts with the same skill as year 2000 5-7 day weather forecasts?
- What do we need to do acquire these observations? What candidate future observing technologies and observing system architectures should be studied further and/or developed to accomplish this goal?
- How do we bring these observations to bear for the science challenges at hand? What space and ground-based communication and data system capabilities and supporting technologies will be needed to gather, disseminate, and stage these data for science processing and assimilation into numerical weather prediction models?
- How do we best incorporate data and advance the capabilities of the predictive models? What advances in the science and/or engineering of numerical models of the earth-atmosphere system and statistical-dynamical data assimilation are necessary to achieve 10 – 14 day predictive capability?

- <u>What concurrent advances in computing technology and infrastructure will be needed</u> to accomplish the necessary advances in numerical modeling and data assimilation systems that will make use of these observations to produce operational forecasts?

1.3 Deliverables

The deliverables of this study are a high-level architecture concept and the identification of key enabling technologies that will require NASA research and development investments.

2. SCIENCE OVERVIEW

Modern operational weather forecasting is based on using super-computers to solve complex systems of fluid dynamical equations. Only 25 years ago, forecasts in the U.S. were based on not much more than extrapolation techniques and the accumulated subjective experience of local forecasters. In general, a forecast could not be relied on beyond two days. With the convergence in the late 1970's of advances in supercomputing and satellite observing, refinement of numerical methods for integrating by computer, and basic research in meteorology, the period of useful forecast skill has increased to more than five days.

2.1 Weather Prediction Models

A weather prediction model is a system of equations that is based on the Navier-Stokes fluid dynamic equations, but re-formulated for atmosphere flows. The atmospheric equations, although complex in appearance and in their solution, are merely restatements of fundamental physical laws such as the first and second laws of thermodynamics, Newton's laws of motion, conservation of mass and energy, and hydrostatic balance.

Making a numerical "prediction" for the atmosphere involves using a computer to numerically solve a set of five (at minimum) simultaneous non-linear time-dependent partial differential equations at every point over a pre-defined 3D gridded model domain, stepping forward in time. In the simplest models, the weather variables that are directly predicted by the model (the so-called prognostic variables) are temperature, humidity, surface pressure, and north-south and east-west components of the horizontal wind. More sophisticated model systems include additional predictive equations for clouds, precipitation, vertical motion, turbulent kinetic energy, chemical constituents, etc. The more variables being explicitly predicted, the larger and more complex becomes the equation set, and the more computing resources are needed to integrate those equations forward in time.

2.2 Limits to Accuracy of Weather Forecast Models

Based on theoretical considerations first expressed by Lorenz (1963, 1969) the theoretical upper limit of deterministic weather forecasting is about two weeks. To approach this limit would require 1) a model with perfectly prescribed initial conditions, 2) a model that has no numerical errors and that perfectly accounts for all known physical processes that affect weather on practically all scales, and 3) open boundary conditions that are known and prescribed perfectly. Lorenz demonstrated that even if it was possible to approach such perfection, infinitesimally small errors from any source, too small even to be measured, will grow and accumulate at such a rate as to render any forecast useless over a two week period. (Note that the limits of deterministic forecasting may to some degree be overcome through stochastic methods. Use of ensemble or Monte Carlo forecast methods is clearly an avenue that will be important, and is addressed later as an important strategy).

At the outset it must be recognized that an "accurate" deterministic 14-day forecast would be extremely difficult to achieve, and might not be possible. But how far can we push the limits of useful weather prediction? The answer will depend on how well we can construct a system that satisfies four items below. Accuracy of numerical weather prediction models is affected by:

- Quality (accuracy and completeness) of the initial conditions
- Knowledge and treatment boundary conditions
- Accuracy of numerical approximations and solution methods
- Accuracy and relevance of parameterizations of sub-grid scale processes

Below we discuss each of these factors, as they relate to the framing the approach and recommendations in this report.

2.2.1 Initial Conditions

Numerical weather prediction is an initial value problem. Thus, a model forecast state of the atmosphere at any time depends only on the state of the atmosphere at the previous time, and by extension, on the initial state of the atmosphere as it was originally cast into model grid coordinates. This means that a model predicted state of the atmosphere can be no more *inherently* accurate than the initial state; and from the initial time forward, differences between the real atmosphere and a model forecast (i.e. forecast error) must increase. The larger the error in the initial conditions, the more rapid growing and profound are the errors that necessarily accumulate as the model is stepped forward in time. Thus, it is of paramount importance to start the model off with the smallest possible initial state error and, as the model is stepped forward, to keep growth of those errors as small as possible for as long as possible, thereby extending the usefulness of a model forecast.

Other factors excluded, extending forecast skill much beyond 10 or 12 days will require an initial starting point for the model that nearly perfectly describes the true state of the atmosphere. In the context of a global model, this would seem to be a nearly impossible challenge, being contingent on the ability to observe the 3-D atmosphere everywhere perfectly. This requirement alone points to the necessity for global space-based observing.

In so pushing the limits of forecast skill in an operational system, one of the paramount challenges and conceptual drivers in this study will be the need to prescribe, as perfectly as possible, the complete 3D [initial] state vector of the global atmosphere as well as the trajectory of the state vector at any arbitrary model starting time, t=0.

2.2.2 Discretization Of Equations

Numerical Weather Prediction models are not the real atmosphere. The model equations merely approximate the true equations, which like the atmospheric fluid itself are continuous. By discretizing the equations in order to solve them with a computer using finite differencing or other numerical techniques, errors are introduced. The growth of the smallest of these errors alone is more than sufficient to prove Lorenz's point. The numerical accuracy of the model depends on numerical techniques used, of which there are many (spectral methods, Cartesian finite differencing, finite elements, finite volume). However, there are no perfect models.

In general, the shorter the time step and the finer (more detailed) the grid resolution, the more accurate is integrative solution of the equations. Time and space resolution are linked to one another in interesting ways. Finer grid resolution also permits smaller, faster moving phenomena to be generated and propagated by the model. In order for a numerical integration to be stable (in terms of a constrained error growth), those phenomena must also be resolved in the time domain. Thus, finer spatial grid resolution (more grid points) also requires shorter (more) time steps, further increasing computational demands. However, some phenomena (physically enabled by the increased resolution) may not be meteorologically relevant and can destructively interfere with proper evolution of meteorologically important larger scale structures.

Numerical modeling is a complex undertaking, and the few statements made here barely do justice to the fact. Suffice it to say that deciding on the numerical formulation of the weather prediction model, such as grid resolution, time step, and forward integration technique involves making numerous trades. These trades however are primarily are not as much technological as they are scientific.

2.2.3 Boundary Conditions

Lateral Boundaries

The size and complexity of models is constrained by limits on total computing resources (e.g. numbers of processors), computing capacity (memory), and processor speed -- and just as importantly, dollars required to gain access to these capabilities. The most fundamental demands placed on computing resources by the models are numbers of computational grid points and sophistication of the physical processes included in the model. The number of grid points carried by a particular model depends on the total geographic area to be represented (size of the computational domain), and the space and time scale (resolution) of the smallest phenomena desired to be represented by the model.

For a given level of available computing resources, the formulation of a model requires that trade-offs be made between domain size and domain resolution (which together determine the total number of grid points). For example, a current-day model run over a *global* domain might typically have an effective horizontal resolution of about 250 km. With the same computational resources, a model with a 25 km resolution must be confined to smaller geographic regions (typically 5000km x 5000 km). Similarly, models that operate at 1 km resolution may only cover an area 100 km on a side. Yet even finer, cloud scale models may be run with a 100 m grid spacing over areas 10 km on a side. Each type of model is tuned for a specific type of forecast. We believe that computational resources will, based on currently emerging technologies, be sufficient in the 2025 timeframe to operate very high resolution atmospheric models globally (1- 10 km grids). The state of the art operational type forecasting model in 2002 is a global model run at about 100 km horizontal resolution, although experimental models are being run globally at 25 km resolution.

In current practice, limited-area models require nesting the finer scale grid domain within a larger coarser resolution model, allowing essential information from outside the limited area grid to enter through the sides and propagate into and through the limited-area domain. Most operational weather forecast models are run over continental size areas with boundary conditions specified from a prior lower resolution, usually spectral global model. Engineering the numerical interface between the two model meshes is problematic, and can result in, among other boundary effects, spurious waves and internal boundary reflections.

Lateral boundary issues are obviated with a global domain; and use of global domain is a fundamental recommendation of this study. Potential gains made in accuracy, from the point of view of lateral boundary conditions alone, far outweigh the increase in computing resource requirements, especially given the rate of increase in computing power.

We envision in this study use of global model domains operating with horizontal resolutions from 10 km down to 1 km. The required increases in computing power on this basis alone are on the order of 10^6 over the current state of practice, but potentially worth the price. Alternately, we can envision a 25 km global model, but with significant research investments given to the development of variable adaptive grid techniques that will allow for accurate two-way interaction between large-scale coarse and embedded finer-scale grid meshes.

Upper Boundary Conditions

Changes in upper boundaries within the stratosphere and lower mesosphere are not believed to significantly affect meteorologically relevant tropospheric dynamics on two-week time scale.

Surface Boundary Conditions

The Earth's surface is the only significant open boundary for a global atmosphere model system. Proper treatment of the exchange of moisture and energy between the surface and atmosphere will represent the single largest R&D challenge for forecasts from a few days to two weeks.

Surface heating and evaporation of water from the surface ultimately drive atmospheric development on almost every time and space scale. Water reaching the surface by precipitation may be recycled between the surface and atmosphere several times over a two-week period; yet by the rules of our two-week forecast scenario, surface heating and moisture fluxes may only be specified one time -- the initial time. While it may be possible to characterize the state of the surface (both land and ocean) in a model initial state, unless one can account realistically for the exchanges on heat and moisture between the surface and the atmosphere continuously throughout the forecast period, it will not be possible to achieve fourteen day forecast skill.

Model precipitation, surface hydrology and the surface and atmospheric energy budgets are highly non-linearly coupled. Fluxes of heat and water from the surface into the atmosphere depend on antecedent model precipitation, (and also soil type, vegetation, near-surface wind, temperature and moisture, and solar insolation (clouds)). Conversely, precipitation depends on antecedent evapotranspiration and on the spatial distribution of surface moisture-modulated sensible heat flux. Specification of these critical surface boundary fluxes at all times in a model simulation depends on the capability to separately and accurately model the surface hydrology and energy budgets, and to couple them continuously with the atmospheric model.

2.2.4 Parameterization Of Sub-Grid Processes

Like any fluid, the atmospheric is continuous from the global scale to the molecular scale. The generation and dissipation of all energy, material and motion at any scale in the atmosphere depends on the exchange processes that occur among all these scales. These processes have to be included in a forecast model. However, a model cannot deterministically include any process whose operative scale is smaller than the spacing between the model grid points. The aggregate effects of such "sub-grid scale" processes on the "grid scale" must be parameterized (estimated statistically or empirically) in the model in terms of other atmospheric properties or processes that can be resolved on the model grid. Examples of such essential sub-grid scale processes include:

- Radiation (clouds, water vapor, aerosols, surface properties)

- Turbulence and diffusion (planetary boundary layer)
- Planetary boundary layer & surface fluxes (land and ocean)
- Surface hydrology and energy budget (land and ocean)
- Moist convection (cloud scale dynamics)
- Cloud and precipitation processes (microphysical)

The complexity of parameterizations can range from simple formulas to embedded fully 4-D process models, each of which may be as sophisticated and computationally demanding as the basic atmospheric model itself. In addition, as grid spacing of the atmospheric model is decreased, the model becomes more capable of explicitly resolving on the model grid some of what previously was parameterized on the coarser grid -- so the formulation of the parameterization may need to change accordingly. The complexities presented are enormous, yet they must be dealt with. Parameterizations or sub-grid models will need to become more, not less, complex with computing power being the largest impediment. Tremendous investments are also required in understanding the basic physics, the scale dependencies and interactions related to these processes, and their incorporation into atmospheric models.

Based on all the above considerations, our base recommendation is for a global mesoscale model with a) grid resolution as high as one to ten kilometers, b) sophisticated parameterizations formulated and tuned to these scales, and c) fully coupled surface energy budget and surface hydrological models (including land and ocean).

2.3 Measuring Forecast Accuracy and Skill

What constitutes an accurate forecast? Actual experience and reflection quickly leads to the realization that "accuracy", based on any lay usage of the term, depends on the expectations of the user. To be scientifically useful, forecast accuracy requires some objective basis. Measuring forecast accuracy has been debated in meteorology for decades, and has become a specialty in itself. The construction of suitable metrics for assessing forecast accuracy depends on the space and time scale of what is being forecast, and on the parameters being forecast (e.g. temperature, cloudiness, wind speed, precipitation occurrence, precipitation amount, time of occurrence of precipitation). This study emphasizes verification at the time and space scale of synoptic scale weather patterns.

On average, the weight/unit area of a column of atmosphere extending from sea level to space is about 1000 millibars (mb). 500 mb is the pressure at which approximately one half of the mass of the atmosphere lies both above and below. Variations in the height of this point (typically around 18,000 feet in altitude) when mapped geographically reveal the existence of dynamically evolving wave structures in the atmosphere. While there are many parameters that might (and will) be used in model verification, we argue that if only a single metric were allowed it would have to be 500 mb height fields. 500 mb height patterns are especially useful in weather forecasting because they are highly correlated with jet stream locations, vertical motion patterns, cloud and precipitation patterns, and with the locations, development, and future movements of surface pressure systems.

The Anomaly Correlation (AC) of the 500 mb height field is the measure of forecast skill accepted currently among operational numerical forecast centers around the globe, and will be assumed for this study. It is a statistical measure of agreement between a model forecast of 500 mb height fields and observed 500 mb height fields, averaged over many model forecasts (weekly, monthly, seasonally, or annually). In making this comparison, climatological 500 mb heights are subtracted out to prevent improper attribution of model forecast skill to an inter-seasonal or instra-seasonal climate signal, whose prediction requires no real skill. The 500 mb AC is calculated over some defined computational (geographical) model domain and typically graphed as a function of forecast length. At the initial time, the AC by definition equals 1.0, but decreases as the forecast length increases time and as model error (relative to actual) grows.

Based on empirical studies, it is generally accepted that a forecast ceases to have predictive value (skill) when the value of the 500 mb AC falls to 0.6. Among major operational forecast centers presently, this level is reached at about 7 days. Our goal is to extend this limit to 14 days, essentially doubling the current range of skillful prediction. It is important to recognize that because numerical weather prediction is an initial value problem, extending forecast skill to 14 days by definition must accompany a simultaneous increase forecast skill for <u>all</u> weather events at shorter time and space scales. Thus, if in 2025 a two-week forecast could be made with the same skill as today's five day forecast, one might expect that today's two-day forecast skill could be extended to four days, today's three-hour forecast skill extended to six hours, etc.

Figure 1 shows three AC curves, representing the state-of-the-art in 1989, 2000 and 2025. Over the last decade, forecast skill as measured by 500 mb AC has increased only by about one day, even with improvements in computing technology, advanced data assimilation strategies and new data sources. That progress has been hard won. Comparing the 2000 with 2025 gives some idea as to the enormity of the challenge of reaching even minimal forecast skill by 2025. However, the comparison also helps illustrate that any significant progress made in increasing longer range forecast skill delivers, by default, similarly impressive increases in shorter range forecast skill. Whether or not we are able to reach the 0.6 AC threshold at 14 days, the effort made will likely result in greatly improved 3 - 10 day forecasts.

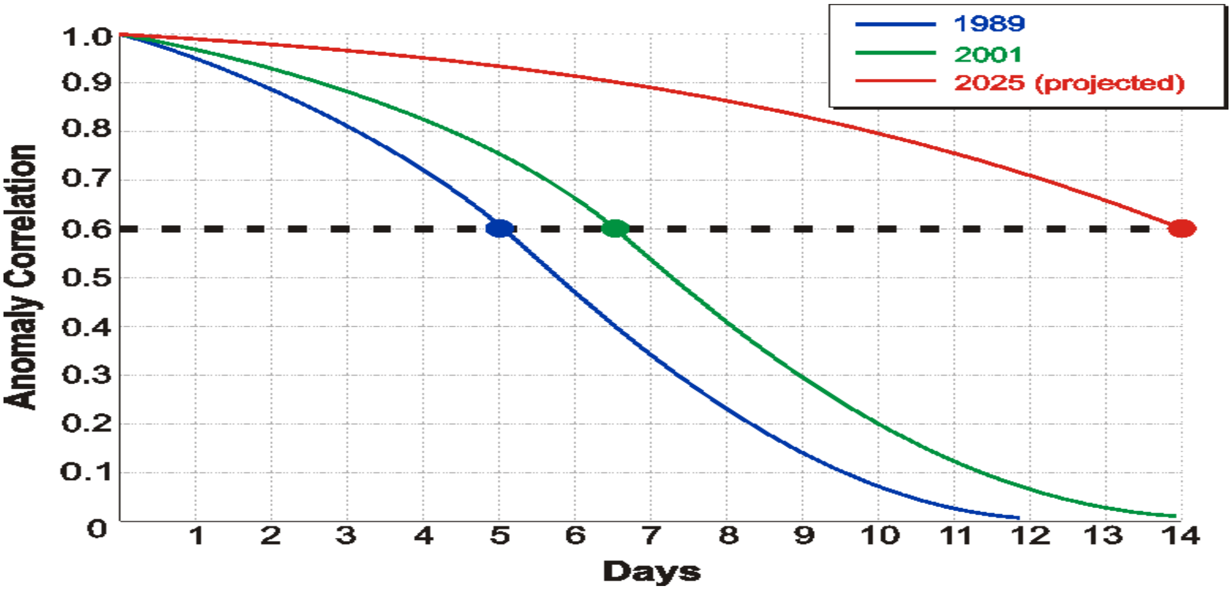


Figure 1. Illustration of decreasing anomaly correlation with length of forecast as a metric of forecast skill.

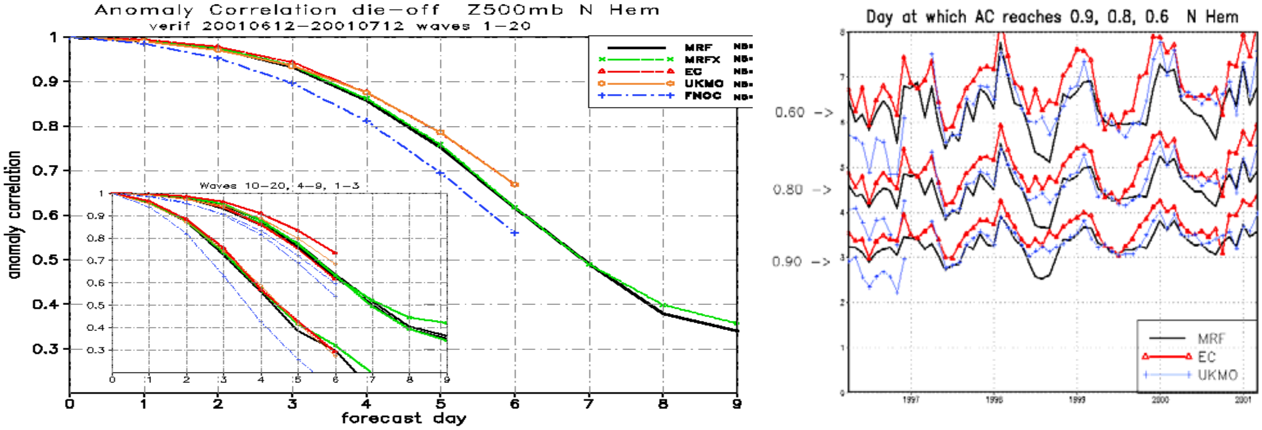


Figure 2. The panel at left shows actual graphs plots of actual AC calculated for forecasts from five different models each showing AC=0.6 at about 6 days. The figure at right illustrates that there are seasonal variations in AC, with better forecasts produced during winter seasons when temperature gradients and dynamics are strong and more easily captured by model equations.

2.4 Data Assimilation

To the extent that a given model permits only physically possible and meteorologically relevant evolution of the atmospheric fluid, the model can be employed to impose a degree of internal dynamical consistency among independent observations obtained at different times and locations, and upon which are superimposed various inherent measurement errors. Conversely, to the extent that the errors in those same observations are small [and are properly understood and characterized] relative to the information they contain about the actual state of the atmosphere, those observations can be used to correct for model forecast errors that accumulate with time, by pulling the model back into alignment with reality.

The objective of data assimilation is to find, using all available observations, the mathematically optimal depiction of the atmosphere state that will exactly satisfy (balance) the model equations at the initial time. This initial state is a compromise between the observations and the model. While this compromise initial state does not necessarily perfectly describe the one true state of the atmosphere, it does none-the-less result in the best possible forecast (in a statistical sense) of the state of the atmosphere at future times. A good assimilation system must take into account:

- Dynamical dependencies among observational variables
- Complementarities, redundancies, and disagreement among variables
- Accuracy and reliability of the observations
- Spatial and temporal structure of errors inherent in various data types
- Observations taken at times different from initialization time
- Inhomogeneities in data densities and types

It is important to recognize that the assimilation process is computationally intensive, and places much larger demand (even 1 to 2 orders of magnitude potentially) on computing resources than does the forward integration of the model prediction codes. State-of-the-art assimilation methods include: kalman filtering based methods which are theoretically optimal; finding and solving adjoint forms of the model equations; and methods based on the calculus of variations. All these techniques involve minimizing some cost function related to differences between model and observations.

Even truncated variants of the kalman filter based assimilation methods being developed still require inversions of $10^6 \times 10^6$ element matrices. Moreover, the size of the matrices depends on the number of observations used in the assimilation, with computing requirements scaling as the square of the number of observations. Computations also depend non-linearly on number of grid points in the model, so that increasing model resolution also greatly impacts the computational cost. At present, a *fully implemented* kalman filter is not computationally feasible for operational weather forecasting, nor is it likely to be as long as the number of observations to be assimilated from satellites and other sources grows. In operational practice only about 10 – 20% of satellite observations are being assimilated largely because of the demands on assimilation related computing resources.

2.5 Emphasis of Study

For 10 – 14 day weather forecasting to be operationally reliable and useful, advances are required on many scentific and technological fronts. Providing the best possible initial conditions for a model -- the top priority, depends most heavily on the ability to provide comprehensive observing of the Earth's atmosphere and surface. This study emphasizes the importance of initial conditions, because it is here that NASA's core capabilities in global observing might be best exploited.

To make significant progress toward our 14 day forecast goals, equivalent large investments are required in the numerical design and development of models and modeling systems, large scale development of Observing System Simulation Experiment test-bed and research capabilities, basic research on fundamental physical processes and their incorporation in models either explicitly or through parameterization. To a large extent, dealing with issues of improved engineering of model boundary conditions and numerical methods, and development of parameterizations and other model improvements fall to the realm of basic science activities and science infrastructure -- areas where progress might be best facilitated by advances in computing sciences and technologies.

In this study, we aim neither to minimize the importance of these scientific endeavors nor to minimize the investments that will be needed. But, due to time and resource constraints, this study was primarily focused on observing system strategies and the interfaces between future observing systems and modeling systems.

3. OBSERVATIONAL NEEDS

3.1 Baseline Requirements

Our baseline requirement is for an observing system that is capable of providing an initial state for the free atmosphere that depicts horizontal and vertical structure “equivalent” to that which could be provided by today’s radiosondes operating every 25 km, every three hours globally. This does not say that we actually need radiosondes, nor even satellite-retrieved temperature, moisture and wind profiles every 25 km. Even AIRS and GIFTS hyperspectral interferometric retrieval techniques cannot match radiosonde performance.

The important point is that no single current or future measurement system alone can prescribe the atmospheric state with the accuracy, reliability, frequency and coverage that we need. In any case, it is the combination of forward-integrated model state (at grid resolution), with all wind, temperature, moisture, cloud and other data from all sources at different times and resolutions and accuracies, optimally combined through the mathematical assimilation process that must yield this equivalency to radiosondes every 25 km every three hours globally. The diversity of data types each having different, but known spatial error characteristics, overlapping in time and space and complementing one another is important to achieving this sought after high-resolution global depiction of the atmosphere and surface. A certain amount of redundancy among observations is essential. Too much is wasteful. The ability to manage redundancy within an integrated observing system for optimal results is a key intelligence that the overall system must possess.

Figure 3 lists, in decreasing general order of priority, the observations thought to have reasonable bearing on the 14-day weather forecasting problem. At the top are 3-dimensional structure of atmospheric temperature, moisture and wind, since to first order the evolution of the free troposphere depends on proper specification of these variables. Variables shown as green in Figure 3 are thought not to present significant technology challenges, but will be obtainable though evolution of current measurement technologies and systems. Variables shown as red are difficult to carry out remotely and will require significant technological developments. In general these are the observations that will involve active remote sensors using such space-deployed Lasers (Lidar) and Radars (See Section 3.3).

	GLOBAL MEASUREMENT	Temporal Resolution	Horizontal Resolution	Vertical Resolution
*****	3-D Atmospheric Wind Speed & Direction	3 hours	25 km	250 meters
****	3-D Atmospheric Temperature (T)	3 hours	25 km	250 meters
****	3-D Atmospheric Humidity (Td, RH)	3 hours	25 km	250 meters
****	Barometric Pressure (Psfc)	3 hours	25 km	NA
***	3-D Precipitation (accumulation, rate, phase)	1 hour	1 km	250 meters
***	3-D Cloud (water content, phase & other properties)	1 hour	1 km	250 meters
***	Land-surface / Soil Moisture (LSM)	3 hours	25 km	NA
**	Land-surface / Soil Temperature (LST skin)	3 hours	25 km	NA
**	Land-Sea Snow-Ice (extent, depth & properties)	3 hours	25 km	NA
**	Sea Surface Skin Temperature (SST skin)	3 hours	25 km	NA
**	Planetary Boundary Layer (PBL) Height	1 hour	25 km	25 meters
*	Aerosols (size dist., conc., & other properties)	6 hours	25 km	NA
*	Albedo (%)	3 hours	25 km	NA
*	Vegetation (e.g. NDVI)	1 week	1 km	NA
*	Surface Roughness (R_0)	2 weeks	1 km	NA

Figure 3. Measurement needs. The number of asterisk (****) reflects the relative criticality (priority) of the measurement for the 14-day forecast problem. Measurements depicted in **green** will likely be provided through evolutionary development of existing and currently conceived & planned technologies without significant special new investments. Measurements depicted in **red** are likely to require overcoming significant technology gaps, and therefore significant additional investment focused primarily in the area of active remote sensing.

The higher the resolution of the model, the greater is the requirement for higher density observations for both model initialization and for model verification. In the case of initialization, arguably, initial data need not be provided at a density as high as the 1–10 grid resolution proposed. The model must be allowed to generate its own internally consistent structures down to grid scale -- over-prescribing detail would probably not helpful. For this reason we believe it is probably sufficient that initial state atmosphere structure information be provided at a 25 km resolution globally.

However, there are reasons to observe surface boundary conditions (soil moisture, vegetation, snow cover, sea surface temperature, etc.) down to a 1–10 km model resolution even if they are ultimately averaged to a coarser grid resolution. The most obvious justification for doing so is the recognition that many of these surface parameters are *already* being measured at these high resolutions. Even while recommending that important surface parameters be observed at very high spatial resolution, variables such as sea surface temperature and vegetation need only be made at a frequency commensurate with change in those variables (days or weeks). High temporal (1-3 hours) and spatial (1-10 km) resolution measurements will need to be carried globally out using space-based remote sensing techniques.

The main reason for measuring clouds and precipitation hourly at high resolution is for their value in providing real-time feed-back on (and potentially correction to) model performance and in detecting rapid change. In addition, a model initialized moist diabatically with precipitation and clouds can have a direct impact on model forecast accuracy beyond 12-24 hours.

A good summary of current observing capabilities can be found in World Meteorological Technical Document WMO/TD No. 1052: “Statement of Guidance Regarding how well Satellite and in situ Sensor Capabilities meet WMO user Requirements in Several Application Areas”. This document provides annually updated community inputs regarding observing system capabilities and requirements to support forecasting in six categories: *Global Numerical Weather Prediction, Regional Numerical Weather Prediction*, Synoptic Meteorology, Nowcasting and very short range forecasting, Seasonal to Inter-annual forecasting, and requirements for aeronautical forecasting. Our requirements in fig. 3 are generally consistent with the maximum data requirements (accuracy, frequency, resolution and coverage) in the WMO document for global and regional numerical prediction.

3.2 IN SITU Observations

Presently, most observations, both space-based and in situ are made without any specific regard to priority needs for those observations based on the meteorology. Radiosondes are launched at the same times and locations. Both in situ observing and satellite observing modes are more or less "set at the factory". As a result, current observing systems do not use resources efficiently -- taking observations where they are not especially needed, or worse, unable to provide observations where and when they might be most useful. Altogether, the coverage and timeliness of these data are insufficient to satisfy the input needs of the *proposed* weather forecast system.

Today about 20% of all data input to synoptic scale operational numerical weather prediction is terrestrial-based or airborne. These so-called "conventional" observations include: land-based surface observations, balloon–borne radiosondes, aircraft-based, ship and ocean buoy reports. The land-based surface network observations include near-surface wind speed and direction, temperature, humidity, barometric pressure, clouds, precipitation and net radiation. These data are reported hourly or three-hourly. There is a current trend toward deployment of networks of remote automated surface observing and reporting stations. Associated applications of new technologies might also focus on automated local data processing, quality control, and real-time conveyance of processed and calibrated information to central collection facilities via satellite.

Balloon–borne radiosondes provide vertical profiles of wind, temperature, relative humidity, and pressure. Radiosonde data is available, by convention from about 700 locations around the globe on a synchronized twelve-hour schedule. Radiosondes are not evenly distributed geographically, but mainly launched from populated land areas in the Northern Hemisphere. By itself, radiosonde coverage is inadequate for the needs of global weather prediction modeling, with coverage being very sparse over the 70% of the globe that is ocean, plus Southern Hemisphere landmasses. Although radiosondes are far from error-free, they are none-the-less considered the standard by which other techniques, notably satellite retrieval methods, are assessed. GPS-based tracking (in contrast with radio-based tracking) is a recent innovation lending greater accuracy in computation of winds to better than 1 m/s relative to 3-4 m/s for older radio tracked sondes. The response time of radiosonde instruments combined with the rise-rate of the balloon is capable of providing a maximum effective vertical resolution of 5 to 20 meters depending on altitude. In practice, observations are always reported at standard ("mandatory") pressure levels, and reported at higher resolution only where significant vertical changes are detected (at so-called "significant" pressure levels).

Operational forecast models and data assimilation methods have been developed around use of radiosonde-based depictions of the three-dimensional atmospheric structure. Opinions vary on the future necessity for radiosondes for model initialization (e.g., Atlas and Korb, 1981). One view is that re-engineering modeling and assimilation systems around satellite-only input will be eventually be more effective than current systems, with radiosondes used only for forecast validation and observing system calibration. The counter view is that some radiosonde data will always be needed to provide control points that anchor the initial analysis itself. Whether for actual assimilation or for improving model validation globally, a technology recommendation in this study is for expansion of the radiosonde network through further continued development un-manned automated radiosonde system that is deployable to remote locations, is capable of being activated (launched) by satellite command, and is able to process data on-board and to communicate processed data via satellite. Fully automated radiosonde systems have already begun to see operational use (Vaisaila AutoSonde).

By providing actual (not just hypsometrically calculated) pressure-heights, GPS-tracked radiosondes will enable an additional quality control check on the hypsometric calculations – e.g., discrepancies could reflect errors in measurement of humidity, suggesting an additional check and constraint on final specification temperature and moisture profiles.

Among the priority measurements in fig. 3, perhaps the single most logistically difficult variable to obtain globally is surface pressure, since it currently can only be measured in situ. Given surface pressure and terrain altitude at a given location, the vertical distribution of pressure is calculated hypsometrically from the surface upward from temperature and moisture profile observations. In the absence of in situ surface pressure observations in remote regions, specification of surface pressure is largely dependent a forecast model first guess, and errors in the surface pressure analysis affect the accuracy pressure-height patterns at all levels (e.g. 500 mb heights). Since pressure, temperature, humidity and wind are dynamically linked (the basis of the model equations), it is often assumed that the surface pressure field is derived implicitly with sufficient accuracy from the other variables, particularly when all the variables are combined through an assimilation procedure that is based on model dynamics, and thus to a large degree on actual atmospheric dynamics.

Although much attention is properly given to assimilation of temperature, moisture and winds, if one could *know directly* the distribution of mass (a most fundamental descriptor of atmospheric state) it must be an important constraint on the specification of the combined variable set. Using DIAL and / or GPS occultation methods (referenced below), studies suggest the potential for remotely sensing surface pressure globally and thus depicting the global 3D mass field far better than can be done now. Additionally, using DIAL Lidar techniques proven capable of providing the full pressure profile, it may also be possible to provide an alternate global reference level (such as absolute pressure at 15 km above a mean sea level or geoid reference) for global hypsometric calculations downward toward the surface (as well as upward).

Commercial aircraft data (ACARS) can provide near-radiosonde accuracy wind, temperature (Benjamin & Schwartz, 1999) (and in the near future, humidity) ascent and descent slope profile information near airports, but level data is generally restricted to commercial routes at cruise altitudes. Their real impact in operational assimilation and forecasting (Shue, et al, 2002) is not yet demonstrated, and quality control is an issue. Occasionally dropsondes from aircraft provide supplementary profile data equivalent to that of radiosondes. In Australia, Unmanned Aerial Vehicle ‘Aerosondes” have been used for obtaining temperature, wind and moisture profiles below 25,000 feet in connection with Hurricane studies. Constant level “DriftSonde” balloons that release dropsondes are also being developed. Surface ship and buoys report surface temperature, humidity, wind, pressure, and sea surface temperature and state. All these technologies are well established, and improvements to them will be evolutionary.

Because conventional measurements are direct local measurements of the atmospheric fluid or surface (being immersed in or in direct contact the measured medium), they have the advantage of accuracy. The disadvantages of most conventional, terrestrial-based observations is that they are representative only of conditions at a point, and their spatial density and coverage is inadequate for global observing.

Non-Conventional Terrestrially-based
Ground-based non-conventional ground-based (remote sensing) observations include data from precipitation radars, and Doppler wind profilers, both of which are expensive and confined to land areas. Only a handful of vertically pointing radar wind profilers exists. Their numbers could be increased at great expense, but still would not serve as a major source of wind profile data on anything close to a global basis. Ground-based precipitation radars will continue to be important but will be eventually (by 2025) be integrated with or superseded by space-based precipitation radar, which is necessary to achieve global coverage.

3.3 Space-Based Observations

The majority of spaced-based meteorological observations are derived from U.S. owned Geostationary (GOES) and Polar Orbiting Satellites operated by NOAA, and from DoD polar orbiting Defense Meteorological Satellites (DMSP). These operational observations are supplemented with data from NASA research satellites (such as TERRA, TRMM, EO-1, AQUA). There are also many European operational and research satellite systems existing and planned that will contribute to total global observing requirements. Most operational meteorological and surface remote sensing make use of passive techniques in the visible, infrared and microwave portions of the spectrum.

Space-based observations comprise greater than 90% of all observational input to operational forecast models. Even so, this amount constitutes only about 15% of the total satellite information acquired. The difference between included and excluded data is the result of quality control, and sometimes arbitrary culling of data because of limits on the amount that can be assimilated with current computing resources.

Satellite-generated data that currently find their way into operational numerical weather prediction systems include sea surface and land surface temperature, winds calculated from motions of water vapor and cloud structures in sequential geostationary satellite images, and vertical profiles of temperature and humidity, or radiances. NOAA Geostationary platforms (GOES) provide passive remote sensing in visible and IR parts of the spectrum. Products include 0.5 km (visible) - 2 km (IR) imaging, and IR temperature and moisture sounding covering large contiguous areas with high temporal resolution. Observables include atmospheric 3-D temperature, 3-D humidity, clouds (VIS/IR), winds (cloud and water vapor tracking), precipitation (IR methods), Land and Sea Surface Temperature (IR). Microwave observing is not presently feasible from geostationary orbit because useful ground resolution would require extremely large (~100 meter) antennae, which at present we neither know how to deploy nor control. Large deployable antennae structures, and electronically pointable synthetic aperture radar is an area for continued technological investment.

Current generation NOAA LEO and DoD LEO satellites provide passive Visible, IR, and Microwave imaging, and IR and Microwave sounding with high horizontal resolution along orbital track. Observables include 3-D atmospheric temperature & humidity, clouds and water/ice content, ocean surface winds, precipitation, snow and ice, vegetation, LST & SST, surface moisture, and aerosols. A single LEO satellite will see a given location only twice daily day. The spatial resolution and spectral resolution of the IR (and Microwave at LEO) sounding channels (and retrieval accuracy) from LEO is superior to imagery and soundings from geostationary platforms, but it cannot match the time resolution and spatial coverage afforded from Geostationary orbit. Providing hourly global coverage will require a constellation of LEO platforms. Deployment and coordinated management of such a constellation will present engineering challenges perhaps greater than the technological challenge of making the measurements.

Future directions for visible and infrared imaging and sounding will continue with development and improvements in hyperspectral sensors. The groundwork is already being laid with observing systems such as the Hyperion on EO-1, MODIS for visible and near-IR, and large-format hyperspectral CCDs for IR sounding and imaging such as proposed for GIFTS (Geostationary Interferometric Fourier Transform Spectrometer/ EO-3). Private industry has already begun to exploit hyperspectral techniques using proprietary sensors (such as AISA). Thus, improvements along all these lines are apt to be evolutionary rather than revolutionary.

Passive microwave *sounding* from LEO has a heritage dating from the late 1970's with MSU on NOAA polar orbiting satellites. Thus, the foundations are already laid for the much improved next generation microwave sounders AMSU, ATMS, and CMIS which may provide 2 km vertical resolution temperature profiles while being able to see through many categories of cloud.

Soil Moisture (Passive)

One of the great challenges is the measurement of soil moisture using microwave (Jackson, 1993; O'Neill, et al., 1996; Schmugge, et al. 1992; Soil Moisture Mission Working Group, 2000). However, the challenges are as much scientific as technological. An advanced European mission concept for measuring soil moisture (SMOS, Soil Moisture Observing Satellite) is now in formulation. GSFC is considering candidate concepts, one of which involves a LEO formation of 3 small satellites that would employ a sophisticated interferometric technique to provide global coverage for "soil" moisture with 40 km horizontal resolution and three-day repeat time. This will involve significant technological challenges in formation flying and on-board processing. Achieving 10 km resolution is believed to be possible in principle with this technique in 2025, but obtaining 1km resolution would require revolutionary breakthroughs. In the future soil moisture monitoring will be accomplished eventually by combining active and passive techniques.

Precipitation Radar

Active microwave sensors that have been deployed from LEO have mostly been in connection with proof-of-concept missions. The greatest success to date has been the TRMM satellite, which houses a single frequency downward-looking precipitation radar capable of 4 km horizontal resolution with a 250-meter range gate. TRMM however was not designed to provide real-time data or to support operational needs. Leveraging the success of TRMM in the 2010 timeframe, a planned successor to TRMM is an interagency quasi-operational Global Precipitation Mission (GPM) involving a constellation of LEO platforms with multi-frequency multi-polarized microwave precipitation radar and multi-channel polarized passive microwave imagers. Given the coordination that will be required among spacecraft operated by different agencies, in different orbits, and with different sensor formats, and between ground and space observing, GPM is fertile ground for testing and developing SensorWeb concepts. GPM represents the first best opportunity to demonstrate many essential features of a primitive Earth observing SensorWeb.

Ocean Surface Radar Scatterometry

Data from other space-borne radars, notably ocean wave microwave scatterometers (SeaSat, NSCAT, and SeaWinds) have been assimilated experimentally into forecast systems with indications that scatterometer-derived ocean surface winds can have a significant positive forecast impact through improved boundary layer winds and surface fluxes (Atlas, et al., 2001) and by contributing to better representation of derived surface pressure fields on assimilation (Davis, et al., 2002).

Cloud Radar

A mm wavelength cloud radar (CLOUDSAT) now in development (Stephens, et al., 1998; Austin & Stephens, 2001; Miller & Stephens, 2001) will provide vertical distributions of cloud water and ice. Planned to track CLOUDSAT in the same orbit plane will be the Picasso-Cena spacecraft (now called CALIPSO), which holds an orbiting Laser system for characterization of aerosols and thin cirrus clouds. These proof of concept missions employ technologies that will continue to evolve. The challenge will be how to deploy sufficient numbers to space to meet continuous global operational observing requirements.

Lidar

As recognized two decades ago (Atlas and Korb, 1981), the potential for applications of space-based Lidar is immense; and it remains so. Lidar can provide (if developed) an array of *globally distributed* high vertical resolution measurements of important variables, including 3-D tropospheric winds, temperature, humidity, pressure, aerosols, cloud water / ice content, cloud heights and depth, and planetary boundary layer height. Differential Absorption based (DIAL) techniques alone can provide high vertical resolution pressure profiles (Korb & Weng, 1983), Doppler wind profiles (Flesia & Korb, 1999), and temperature and moisture profiles (Korb & Weng, 1982).

Direct measurement of surface pressure globally is of special interest, although the apparent potential of Lidar to provide such measurements is not widely recognized among the meteorological community. The study team was in fact puzzled by the lack of attention given to this capability, which we believe needs to be carefully re-examined. Already, measurement of the pressure profile and surface pressure has been demonstrated from both ground and airborne platforms (Korb, et al, 1989) and even employed successfully to detection of gravity wave pressure perturbations (O'C. Starr, et al., 1992). As with any laser profiling system, clouds are somewhat of an impediment, but the referenced technique will also provide very precise cloud top heights and pressures.

Basic Raman Lidar techniques for measuring aerosols, humidity and various chemical constituents are already well established based on deployments from the ground and from aircraft (Whiteman, et al., 1992; Turner et al., 2000).

Lidar Winds

Supported by OSSEs, a significant scientific movement has emerged over the last twenty years, asserting that investments in systems to make accurate 3-D space-based global Doppler Lidar winds would offer greatest incremental improvements to weather forecasting relative to comparable investments to enhance measurements for other variables such as temperature or humidity (Baker et al., 1995; Ingman 2001). Although the measurement of winds from a space–based Doppler LIDAR has been proposed in various incarnations since the early 1980's, the technology challenges are immense. Scenarios for Doppler Lidar winds from both LEO and GEO have been advanced (Emmitt, 2001). We did consider both LEO and GEO based Doppler wind scenarios and concluded that the technological issues from GEO were not likely to be surmounted in the 2025 time-frame; and that in any case, superior accuracy and global 3-hourly coverage could be obtained from a LEO constellation of 10-13 satellites, compared to seven required platforms in GEO that would not still not provide polar coverage.

We note that the GSFC- developed “double-edge” Doppler detection technique (Flesia and Korb, 1999) has been adopted by ESA for its laser wind profile program; and being based on molecular backscatter, it is not limited like earlier proposed techniques by presence of sufficient background aerosol.

Radio Occultation

Based on a radio occultation techniques developed originally for exploration of atmospheres of other planets, careful processing of GPS (1.2 – 1.6 GHz) signals offers promise of being able to provide accurate global temperature and moisture profile information. It has been demonstrated that refraction of GPS broadcast signals through the limb of the atmosphere is related to temperature and moisture, and provides a physical basis for retrieving vertical profiles of temperature and humidity with 1– km vertical resolution and excellent accuracy (Kursinski, et al., 1997). Palmer, et al. (2000) have also demonstrated the use of a occultation method for obtaining surface pressure. A single orbiting receiver rising and setting on the existing 26 U.S. GPS will be able to produce 1000 soundings per day uniformly in space and time around the globe. At least one program (COSMIC) is already proposed that would fly eight space-borne receivers to produce 4000 retrievals per day. Other systems have been proposed in which a fleet LEO satellites *both* receive and transmit to one another through the atmospheric limb would offer continuous global 3-D coverage distribution temperature & humidity. Ground-based networks of GPS receivers are also being implemented among US universities (SUOMINET) for making atmospheric temperature and moisture profile measurements along lines of site to many GPS satellites.

Where the OERSTED and SUNSAT space-flight demonstration programs failed for technical reasons, CHAMP (a joint German / US Program) and SAC-C (a joint Argentina-U.S. Program) have produced tens of thousand of retrievals that are now being analyzed. Radio signal refraction is sensitive to both temperature and humidity. Deconvolving temperature structure from moisture structure is one of the challenges made simpler when a good first guess temperature profile is available, from a model forecast for example. Once temperature is determined the remaining signal is remarkably sensitive to humidity. In the future, the addition of a 22 GHz broadcast could allow the determination of humidity more directly.

A number of studies have also demonstrated techniques and benefits of assimilating GPS temperature and humidity into numerical weather prediction models (Kuo, et al, 2000). Recent studies have begun to examine benefits of assimilating GPS retrieved profiles versus assimilation of the GPS refractivity information directly (Palmer, et al, 2000; Poli, et al, 2002; Poli & Joiner, 2002). It is certain that GPS occultation methods will become a routine and important global source of vertical atmospheric structure information; but, in-as-much as the principles are understood and prototypes have been tested, this technology will evolve and be implemented in due course limited only by funding.

Another, novel use of GPS proposed by JPL would make use of GPS signal reflected from the ocean back to an orbiting satellite to infer sea surface wind speed.

Other Orbits

Other potential future capabilities that were noted, but not considered in depth in this study, include exploitation of vantage points from polar geostationary orbit, and from Lagrange point orbits. It was suggested during this study that from L-2 (dark side of the Earth) one could obtain nearly continuous solar occultation measurements of atmospheric constituents; however these were not considered immediately relevant to the fourteen day forecast issue.

4. Current Modeling and Assimilation Practice

This section briefly describes steps involved in operational weather forecasting as practiced by major forecast centers around the world. We will suggest changes to current modeling and assimilation practices from a system perspective, and how the overall forecast system processes might be redesigned and improved in the future to take advantage of new capabilities in computing, communications, artificial intelligence, and SensorWeb concepts.

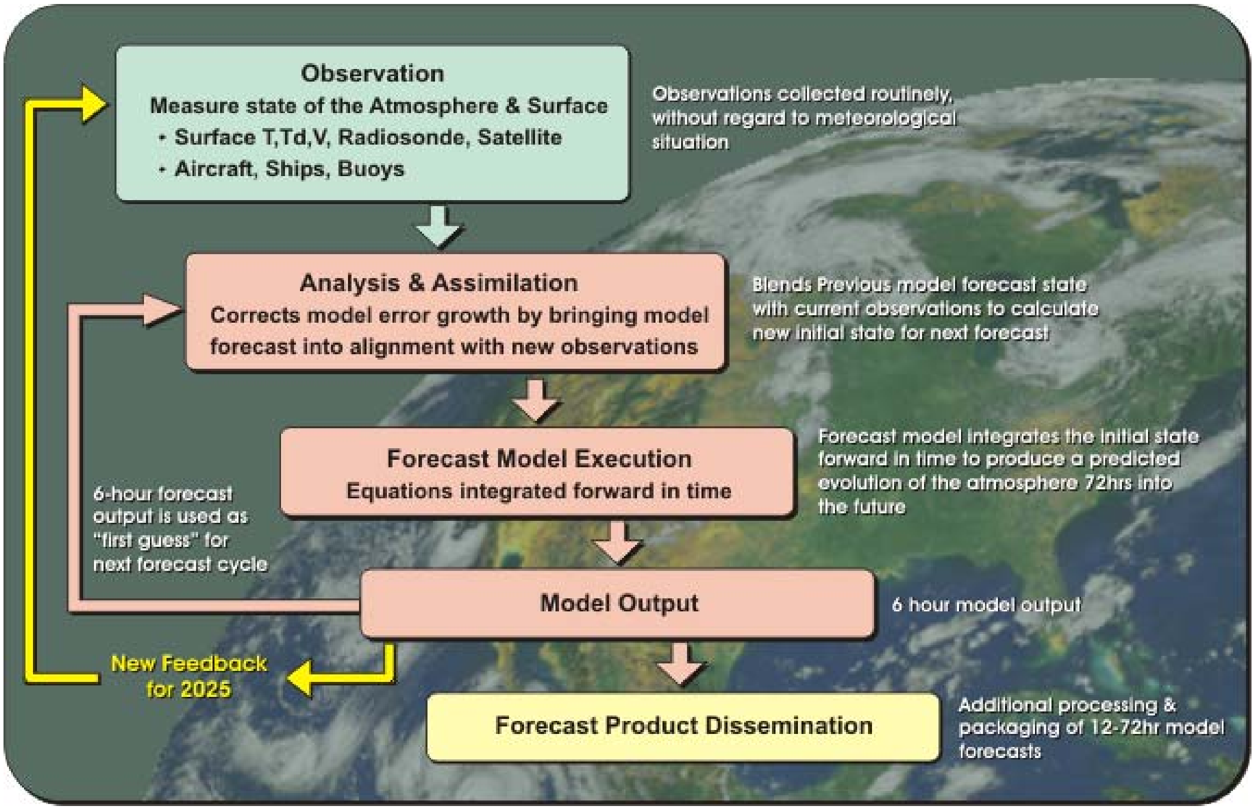


Figure 4. Summary of current and possible future practice in operational weather forecasting.

4.1 Observation

As shown in Figure 4, weather observations are routinely collected from both terrestrial- and space-based sensors. In situ data, such as Radiosondes and surface shelter observations are collected by international agreement globally at the same times (0000 GMT and 1200 GMT) and locations (typically located near airports and population centers). Thus, location and timing of the observations is not determined out of consideration for any specific meteorological need. Space based observations are produced more or less continuously in their orbits, although by a limited number and variety of spacecraft. All these data are communicated to central ground facilities for quality control and processing prior to incorporation into a numerical weather prediction model.

4.2 Analysis and Assimilation

The process of incorporating all these data into a forecast model is performed in a combined analysis and assimilation step. At this stage, the spatial (and sometimes temporal) structure represented inherently in observational data at different locations is projected into a horizontal and vertical grid that defines the three dimensional computational domain of the forecast model. The result is a calculated initial state that both satisfies the model equations and accurately reflects the actual state of the atmosphere. This process is done operationally every twelve hours (sometimes six) at 0000 and 1200 GMT, coinciding with the simultaneous global launches and reporting of radiosondes.

There are a number of mathematical techniques for accomplishing the analysis & assimilation, but in all cases the model initial state at every model grid point (x,y,z) at a targeted time t is essentially computed as a weighted average of nearby observations (collected within some within some time window, typically 3-6 hours) and a first-guess value from an earlier model forecast valid at that same time t and location (x,y,z). Historical observational error and forecast error statistics are introduced into the calculation to help determine the relative weighting between the first guess and the observations such that the resulting adjustment minimizes the overall error in the integration. The main difference between techniques is the way the weights are computed. The end result is that wherever sufficient numbers of reliable observations exist, the initial state will strongly resemble the actual observations. Where data are sparse or unreliable, the initial state will more strongly reflect the first guess given by a previous model forecast.

4.3 Forecast Model Execution

The initial state is the starting point from which the model calculates future atmospheric states. There are many techniques, but all use numerical methods to integrate discretized versions of the model equations forward in time. From a given initial state, current operational global models output forecasts states every 6-12 hours of forecast time. The first 12-hour forecast from a model run provides the first guess that will be used to help determine the initial starting point for the next analysis / forecast cycle. In this way there is a continuous feedback loop between the model and the assimilation process as indicated by the pink arrows in figure 4.

4.4 Model Forecast State

“Snapshots” of raw model results are output every 12 hours of forecast time (out to 120 hours), then subjected to some additional processing prior to being made available for commercial use and public dissemination. The first twelve-hour forecast output is sent back to the assimilation system where it is stored to provide the first guess for the next analysis / assimilation cycle when the next batch radiosondes and other data have been received. This cycle is repeated every twelve hours, each time resulting in a five-day forecast updated with most recently queued observations (a recommendation of this study will be to increase assimilation frequency to one hour).

4.5 Forecast Product Dissemination

In practice, the raw grid point output from a numerical weather prediction model is not usually ready (accurate) enough for public consumption, because of accumulation and propagation of observational, assimilation and numerical integration errors along the entire process. It is fortunate that a significant component of total forecast error for a given location is systematic and is traceable to model biases and local microclimatic factors not addressable in the model. By compiling statistics over many years that compare model predictions with actual weather, the direction and degree to which a model forecast for a given location is in error is statistically predictable. Even with statistical adjustments applied, the model results may require further interpretation and refinement by an experienced local meteorologist. Value-added graphics products based on model output, along with real-time satellite and radar, are much of what a consumer sees on his local TV weather.

5.0 DEVELOPING A NEW ARCHITECTURE

5.1 General Description

The central premise of this study is that it will be possible in the future to fundamentally improve on current operational processes by building in an additional feedback between the forecast model and the observing system, such that the observing system operates flexibly and is responsive to special data acquisition needs identified by the forecast model. Given opportunities to realize key technological advances over the next quarter century, this new feedback (shown as a bold yellow arrow in fig. 4) could significantly advance weather forecasting. The simplest implementation of such a feedback from model to observing system might merely involve increasing the frequency of data collections upstream of locations where the model predicts future development. A more complex implementation might involve targeting specific observations based on Kalman Filter or Singular Vector methods. To complete the feed–back loop, real-time reporting of observations to the model could help to quickly identify discrepancies and enable the model to be appropriately adjusted /corrected.

This interactivity is illustrated simply in figure 5. It is a self-optimizing forecast system made possible by continuous near-real time two-way feedback of information between a Modeling & Data Assimilation System (MDAS) and Observing System. The MDAS development will be based on significant but evolutionary science-based improvements to current-day models and assimilation systems. The Observing System will draw on SensorWeb concepts, which will require a commitment to sensor and space systems technology developments, and require significant advances in the development and implementation of a highly complex, automated intelligent command and control system.

The approach is intended to be scalable -- in which benefits of coordinated observing and modeling may be realized for any level of assets, providing they can be associated in the manner described below. The implementation of such a system does, however, depend on significant advances in computing power both on ground and in space, on communication technologies, and on advances in artificial intelligence and integrated systems. At a working level, the components of the system are: A) an Intelligent Global Observing System (SensorWeb) that provides comprehensive fundamental observations in real-time; that is flexible enough to provide special observations on demand; and that integrates all available observing assets (terrestrial and space-based) under single Command and Control, B) a Global Mesoscale Atmospheric Prediction Model with sophisticated parameterizations, including dynamic coupling to the Earth surface, and having horizontal resolution from 25 km to 1 km, and C) a Data Assimilation System (DAS) that couples in situ and space-based observations to the global mesoscale model, updated as often as hourly with newly available observations.

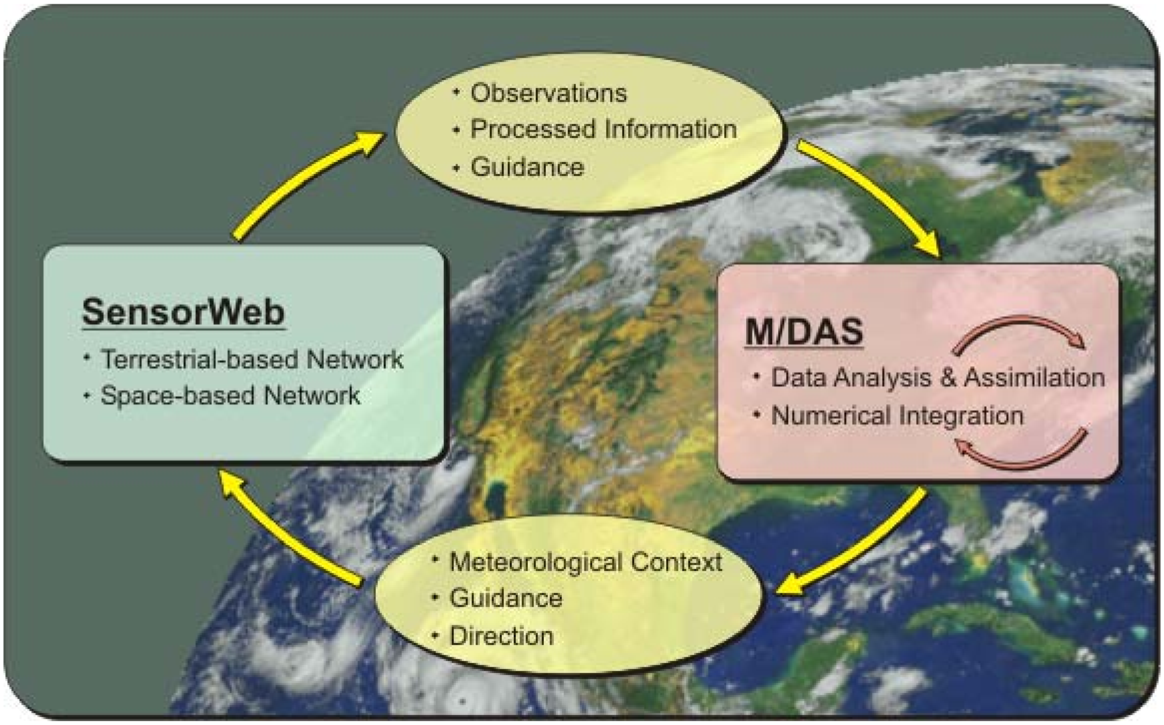


Figure 5. Two-way Interactive SensorWeb and Model / Data Assimilation System

5.1.1 SensorWeb

Within the SensorWeb box (at left in figure 5) is a seamlessly networked observing system comprised of all in situ and space-based observing platforms and sensors, each of which is a potential node. SensorWeb is an emerging concept that allows for intelligent virtual organization of multiple numbers and types of sensors (Space, Terrestrial, Fixed, Mobile) into a coordinated “macro-instrument”. The power of a SensorWeb is that information collected by any one sensor can used by other sensors in the web, as necessary to accomplish some coordinated observing mission. Adaptive behavior can be initiated throughout any or all assets of in Sensorweb by external inputs or by one or more of the members of the web itself. An embodiment of a SensorWeb (Lemmeran, et al., 2001; Delin & Jackson, 2001) may rely heavily on artificial intelligence, permit coordinated coincident observing from multiple perspectives, is driven by reconfigurable mission dependent software, may require advanced communication capabilities and protocols, and is enabled by real time “on-board” processing, analysis and decision-making.

What is unique about this system architecture is that, unlike present day weather observing systems, the SensorWeb (and by extension the sensors within it) will have access to knowledge beyond what individual sensors see in isolation. The SensorWeb will have access to information about the present state of the atmosphere globally and, most importantly to information about the probable future states of the atmosphere generated by the forecast model. This will enable observing strategies to be tailored to schedule critical observations of certain types at times and locations that will have highest impact on the ultimate forecast of the event. Observing requirements schedules may likewise be relaxed in areas where the atmosphere is known to be slowly evolving, in order to conserve resources.

Within the SensorWeb system must reside the intelligence to "understand" (and react to) the meteorological context against which observations are to be made and interpreted. The system will have the capability to depart from default observing scripts depending on location, situation, or phenomenon expected (from the forecast model) or that might be encountered unexpectedly during routine observing. During a single orbit, a LEO satellite will encounter day, night, winter, summer, land, ocean, convective, stable, tropical, mid-latitude conditions, etc. Flexible spatial sampling from spacecraft and instruments will enable observing to be reconfigured with changing needs along the orbit in response to new information provided by spacecraft in the SensorWeb, and in response to the modeling system on the ground.

There must also be sufficient on-board processing and storage so that individual spacecraft and instrument in the SensorWeb can autonomously recognize targets of opportunity, and alert other spacecraft and the model to meteorologically significant developments. Specifically driving on-board processing and storage requirements will be the need for on-board image processing, analysis, and pattern (change) recognition. A command and control function, the brain of the sensorweb, will coordinate communications among spacecraft system-wide and have the ability to autonomously re-direct observing system resources based on targets of opportunity and on pre-programmed priorities.

5.1.2 Modeling & Data Assimilation System (M/DAS)

The M/DAS (right side in figure 5) is comprised of the model that generates the weather forecast, and the assimilation process by which observations are incorporated into the model. Together they comprise a system whose essential feedback is reflected by the pink arrow in figure 4. These interactions are already well established in operational forecasting.

In the new framework, the M/DAS has an additional purpose. It will provide the sensorweb with predictions of what individual sensors should expect to see at a given time and place throughout their next orbit (in space) or other observing period (terrestrial systems). Model predictions and actual observations will be compared in near real time; and in response to such real-time feedback from the SensorWeb, the model may automatically reconfigure itself, for example by modifying its parameterizations, or by adapting its grid resolution in order to better capture what has been observed.

Similarly, based on its own predictions and assessment of observational needs, the M/DAS will able to automatically request operational / behavior changes within the Observing System and among observational network elements. The M/DAS will be able to direct the SensorWeb, through a command and control system, to schedule specific targeted, complementary, time sequential, multi-view observations whose assimilation will especially improve model depiction and forecast, or will facilitate ongoing assessment of model forecast performance.

5.2 Detailed Description of Architecture Components

As an expansion on figure 5, figure 6 attempts to illustrate in more detail, important functional elements of the overall architecture and the various interactions among them. The background image aims to convey that the proposed system is viewed as being "immersed" within a pervasive seamless Earth – Space – Airborne telecommunications network operating with unified protocols. The land-based terrestrial telecommunications backbone network will continue to evolve tremendously in terms of speed and bandwidth. Important technology gaps are probably not a consideration for land-based communications. However, the satellite-to-satellite communications (RF and potentially Laser), and downlink requirements will tax our ingenuity throughout the next twenty-five years. The burdens of Space-to-Earth communications will grow on one hand due to the numbers of satellites and volumes of data envisioned. On the other hand, doing more computational analysis and high-level information processing in space will alleviate some of this burden.

5.2.1 External Control System

Technically outside, an External Control System (**ECS**) performs regulatory functions. It provides the interfaces for humans in the loop, implements security, and provides overall monitoring and control for the combined observing and modeling systems. The ECS governs the implementation of human-directed policy regarding operation, prioritization and allocation of system resources. Given a set of policy priorities, it will have autonomy to adjust operation of all segments to optimize quality and throughput of products against those priorities. Priorities will change, for example with Hurricane season, during El Nino years, in response to natural disasters, or perhaps in deference to specific data requirements of important field research programs. It is through ECS that the science community would be authorized to address and interact with components of the SensorWeb for research purposes.

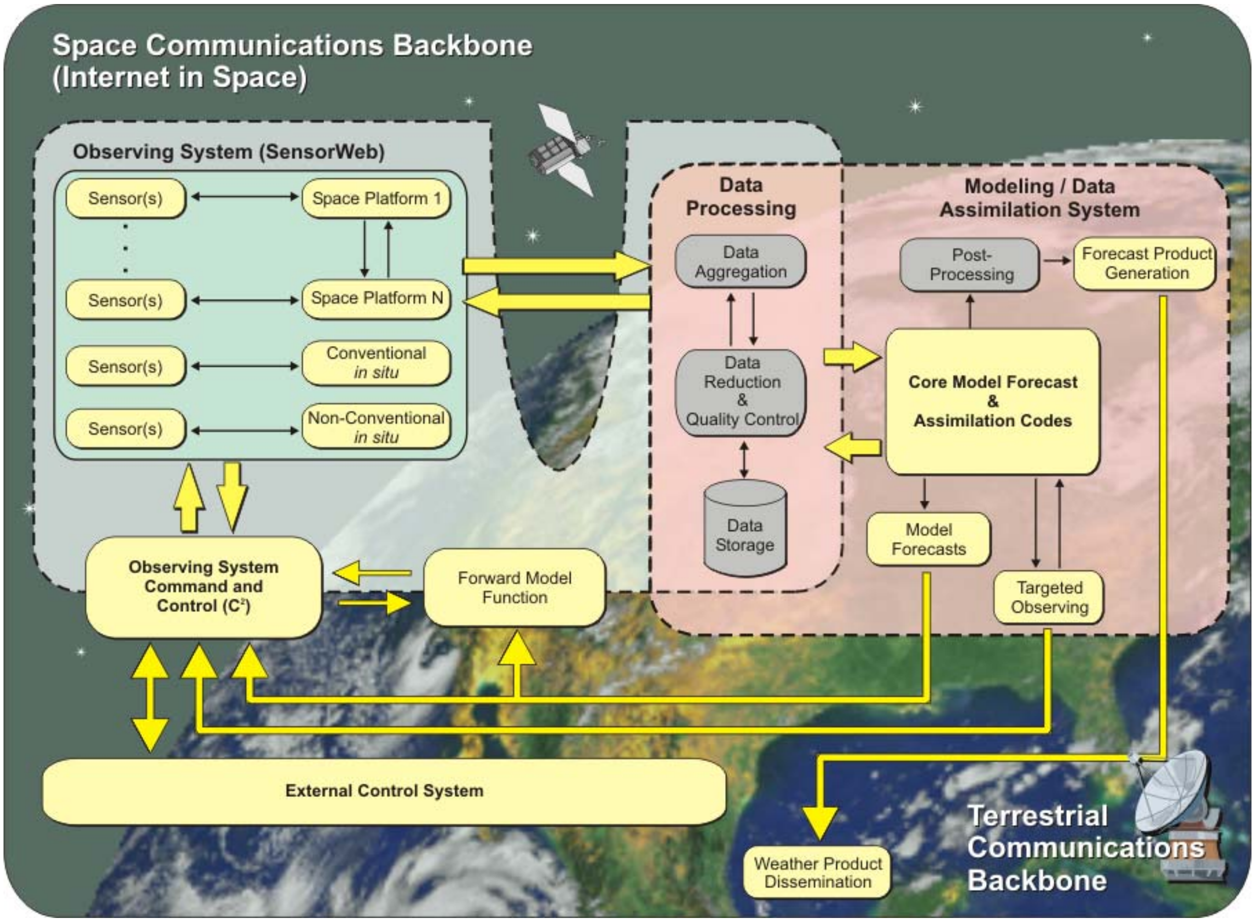


Figure 6. Architecture Concept for 2025 Weather Forecasting System

5.2.2 Required Observing System Capabilities

The observing system must have many advanced capabilities. Foremost, it must be able provide nearly continuous global coverage. As an operational system, it can have no single point of failure. So the architecture must be flexible, reconfigurable and able to adjust automatically to the addition or removal of individual spacecraft, instruments or other system components without compromising the operational mission. Thus, the system should be allow plug & play spacecraft replacement. The entire system must be supported by a robust telecommunication system in the form of a seamless space-terrestrial "internet", that can collect and support near-real-time (15 minutes) global delivery and staging of data.

The observing system must have the capacity to take advantage, on command, of multi-temporal sampling and multi-angle viewing opportunities of the same location using identical or different sensors, this being necessary to detect, confirm, interpret and retrieve certain types of geophysical measurements or processes. This implies a need for flexible, accurate, taskable sensor pointing and control, and potentially even spacecraft maneuverability. Within the observing system must reside a high degree of artificial intelligence and substantial on-board processing capability. The observing system will be required to reconfigure itself flexibly and quickly in response to changing missions.

The first functional mode of the observing system is to reliably collect, process and deliver the default routine global observations that the M/DAS needs to produce operational forecasts. Departures from this default mode will arise often based on a determinations from any accepted source that an unanticipated event/or departure (from model forecast) has begun, that a future event is anticipated at a certain time and location that requires additional observations, or that a change in observational priorities/policy has been directed from the ground (ECS).

The second functional mode of the observing system is executing measurement strategies in response to needs identified *by the modeling system*. If the modeling system determines that additional observations are needed in key locations (i.e., targeted data collection), those requirements are conveyed the Observing System C^2, which calculates how to optimally manage and schedule observing assets as needed, and then and elicits behaviors at the platform and sensor level. The third functional mode of the observing system mode is to execute measurement strategies in response to needs identified autonomously by elements of the Observing System itself; for example, in cases where some incipient phenomenon has been detected that bears special attention or confirmation from other sources

Just like space-based assets, ground-based observing systems are part of the Sensorweb, collecting in situ data, calibrating it, geo-locating it, quality-checking it, and reformatting it at the sensor or platform, and uplinking it via the global Earth-space communications network in near real-time to a collection point. For example, automated GPS radiosondes would be released both on a set schedule and with special releases as requested and coordinated through the Command & Control. Once released, a sonde would log into and dump its data real-time to the observing system network and (or radio data back to the release site, which is logged into the network). Other automated ground-based and airborne systems (such as acoustic or RF wind profilers or ACARS) would similarly transmit data on a routine basis, and in response to special requests.

To every extent possible, satellite and in situ data will be calibrated/corrected/quality-checked (e.g., on-board calibrations for reference measurements and corrections for antenna patterns and radiometer drift) and earth located before being forwarded to the [ground-based] aggregation point for any additional required processing, and staged for potential assimilation.

5.2.3 Observing System Command & Control (C^2)

Among those functions pictured as in the domain of the Observing System are the Command and Control (C^2) System, the Instruments and Platforms, the Forward Model, and Observing System Network. Observing platforms contain the instrument(s) and sensor(s) needed to make fundamental observations. The observing system includes both space and ground-based assets. A data collection function requires onboard processing and storage capability to allow new levels of data processing on orbit. The rationale is to help to offset the burden of downlinking tremendous volumes of raw radiances and other intermediate observations to the ground for processing.

A very important feature is the ability for the observing system to have resident within itself, a sufficient degree of intelligence and analysis capability to independently recognize and characterize change relative to model predictions or previous observations. This requires that a given sensor or platform is able receive and utilize information from the ground about what the model has predicted and also what other sensors/platforms saw on earlier overpasses of a particular area. Much of this communication will be coordinated by the **C^2** and facilitated by a seamless space/terrestrial “internet”.

Taking into account the data requests and future disposition and availability of various sensors, the observing system **C^2** determines what each of the sensors should measure in future orbits. As each sensor makes measurements (routine or otherwise), some level of processing will usually need to occur either on board or on the ground in order to cast space measurements and information from the model or other observational platforms into a common variable space for intercomparison. An example might be the need to compare a satellite 500 mb temperatures against actual radiosonde or model predicted temperatures. To compare on the basis of temperature, a satellite-based temperature must be retrieved from radiance observations. Given on-board knowledge of a model-based first guess and sufficient computing power, the retrieval code could be executed on-board the spacecraft. Alternately, the radiance data could be communicated to the ground, where the retrieval and intercomparison would be executed. Given the volume of data represented by radiances, there is clearly a trade to be examined between onboard processing and communications bandwidth. If the comparison was to be performed on-board in radiance space, the model forecast state at the satellite location would be converted into radiances. As another example, there may be a need to compare 500mb Doppler Lidar line-of-sight wind information with a model prediction. In this case, given precise knowledge of where the satellite is and where the laser is pointing, comparable winds might be computed from the model using the same line-of-sight.

Of great importance to the observing system is a seamless space-and ground-based network that allows very fast global communication between observing assets and **C^2**, between platforms, and with the central data collection point on a nearly continuous basis (i.e., assured downlink opportunities with less than 15 minute waiting).

Much of the intelligence of the overall system will reside in the **C^2**. Whether **C^2** functionality is provided for on the ground or in space, or is consolidated or distributed, was not determined. However, the magnitude and complexity of the **C^2** envisioned presents very complex challenges in the arena of software system engineering and artificial intelligence. More than all the technological challenges, this aspect of software engineering presents the greatest overall challenge in terms of scope, complexity and human labor investment.

The **C²** system manages and directs all Sensor Web assets based on inputs from the MDAS, other users, and from the SensorWeb itself to collect data non-routinely as opportunities are known. **C²** monitors the quality of the data that is being returned by the Sensor Web and automatically schedules additional or corroborating observations that might be needed to ensure high confidence in data quality. Based on requests, the **C²** tasks the observing system to take observations as needed. If the total of observing requests exceeds the capability of the Sensor Web, the **C²** will be able to prioritize and resolve conflicting requests.

In order to collect data to calibrate the space segment, the **C²** will automatically periodically schedule and carry out special measurement scripts that are coordinated with in situ measurements from buoys, aircraft, balloons etc.

5.2.4 Observing System Network (Integrated Space-Ground Backbone)

The network system provides the communications infrastructure for the observing system. The network will be both space-based (e.g., internet in space), ground based (series of ground stations to transmit/receive requests and data), seamlessly integrated. Platforms with various levels of processing capability will support the network communications overhead, including providing the ability to log into the observing system network, collect/calibrate observational data, and pass data to the collection point. Every sensor system can be a "node" on a large-scale network (e.g., could act as separate computers in a distributed processing network).

A pervasive communications network, including a space segment, will ensure seamless interoperability between space, airborne and terrestrial platforms. By 2025, the Network function will be expected to handle a combination of a MEO constellation of wide bandwidth satellites, terrestrial wide-band data links and various other communications systems. The MEO constellation may be shared with other users, but the weather data will need to have guaranteed communications capacity and the highest priority. The MEO constellation will perform 2-way communications with each LEO observing satellite using wideband RF or optical means. The MEO constellation will relay its data either directly to the ground or to another constellation satellite that is in communication with the ground. Other mobile platforms such as aircraft and balloons may use this same constellation or some other existing communications system. Data from geostationary satellites will be sent directly to ground stations and fixed observing instruments and will utilize terrestrial data networks. In all cases, once the data are received on the ground, they will be forwarded to the appropriate users using the address information contained in each packet of data.

Data integrity, reliability, security: Advanced communication protocols proven for use in stressed environments to will need to be developed, standardized, tested and implemented. The protocols will ensure virtually error-free transmission by using forward error correction and detection codes. When an uncorrectable error is detected, the communication system will request a retransmission of the packet of data that was in error. To ensure the security of the data it handles, all commands routed to and from the C^2 will need to be encrypted. If necessary, measurement data itself may be encrypted to prevent its unauthorized use.

5.2.5 Forward Modeling Function

An explicit "forward modeling function" will facilitate an apples-apples comparison of what a given satellite sensor (at a given place, time and viewing path) actually "sees", and the geophysical parameter the forecast model has projected. Most satellite-based measurements do not provide direct observations of a geophysical variable, but rather a radiometric or some other partial or indirect representation of the desired variable measurement. Making such comparisons often involves non-trivial calculations to convert the satellite measurement into a geophysical variable (retrieval process). The intercomparison may also involve converting a geophysical variable into the satellite radiance space (forward process) to be compared with the satellite radiance measurements. The differences between observations, whether viewed in geophysical parameter space or a radiance space are what ultimately get assimilated into the model.

In the forward process case, the forward model function will be able to transform MDAS' forecast atmosphere into model forecasts of satellite observations that each sensor on each platform should expect to see in its native sensor format throughout its upcoming orbit. This includes transforming model data to match any parameter space (e.g. radiance) and sensor viewing geometry. Because the modeling system "knows" the precise orbital parameters of each satellite, as new MDAS forecasts become available the current and forecast state information relevant to each satellite and sensor are delivered to each platform and instrument through the **C^2**. Each satellite measurement can be geo-located and calibrated on-board, and compared to the forecast of that same measurement. These model data delivered to the platform will be for change detection, quality control or for providing first guess information for an on-board geophysical retrieval. Quality flags may be assigned indicating differences as meteorologically real & significant, or suspect, before passing processed data back to modeling system through the **C^2** for later assimilation.

In Figure 6 the forward modeling function shown with the observing system. The most important issue is not whether a forward model function is needed (it isn't in all cases), but whether overall system efficiencies can be gained by moving these calculations from the ground to sensor platform. The trade involves consideration of the competing demands of doing geophysical retrievals or other calculations in space (requires significant on-board processing) and downlinking the processed observations, versus downlinking tens of thousands of raw uncorrected radiances for processing on the ground and placing greater demands on space communications infrastructure. Consider that the proposed GIFTS alone will produce nearly 10^{10} radiance measurement per day. With the trend toward hyperspectral remote sensing in general, it could be far less demanding to emphasize increased on-board processing than to downlink all radiometric data for ground processing.

5.2.6 Data Processing

Overlapping the SensorWeb and MDAS in figure 6 are functions categorized generally as "Data Processing". These functions do not involve significant technological challenges. They are discussed as part of the MDAS, because such processing has traditionally taken place on the ground. A Data Aggregation Function provides an aggregation point (real or virtual) for all Earth and Space-based data collected and transmitted via the Observing System Network. The data are binned and forwarded to the Data Reduction and Quality Control.

Data Reduction and Quality Control refers to the processes that convert sensor-specific data into environmental parameters and formats that can be assimilated into a model (e.g., temperature soundings, surface wind speeds, etc.). Data reduction includes geo-location, calibration, and correction, some of which will increasingly done on-board the observing platform. Quality control (QC) of observational data, and the correctness of a decision to keep or reject data is traditionally one of the largest identifiable sources of forecast error. Data may be rejected for a variety of valid reasons: transmission errors, instrument failure, or contamination from the atmosphere (e.g., cloud contaminated satellite temperature retrievals). Operational quality control algorithms reject as much as 10% of available data -- a consequence of the threshold and statistical techniques employed. However, there are instances in which bad data pass the quality control and good data do not. Intelligent systems and protocols can be developed that can better distinguish between "bad" measurements and "valid outliers". Based on a global continuous data collection capability involving many types of complementary data from multiple platforms and perspectives, additional resources can be quickly tasked to provide additional observations to help decide whether to keep, reject, or replace suspect flagged data.

Reduced and quality controlled data are staged to Data Storage to wait for the next assimilation. Some of these data must also be retained for some period of time to be used in comparison with data from future times for change detection.

Post Processing includes modules that continuously validate model forecasts and compile statistics that will ultimately feed back to the model development and forecast improvements. Post processing also includes application of (MOS-type) final statistical corrections to raw model output prior to creation and dissemination public-use forecast products.

5.2.7 Targeted Observing Function

Besides the global Forecast Model and Assimilation Processes themselves, other functions critical to the architecture are shown under the M/DAS side of figure 6. Among the more interesting of these is a Targeted Observing Function which contains the software and operations that determine, based on current evolution of the model atmosphere, where and what observations will be most important for updating the model in order to optimize future forecasts. The Targeted Observation Function tasks the SensorWeb through **C^2** to acquire the desired observations, if possible. Targeting as used here has two contexts. First, determining which observations will produce the best forecast as measured in an "overall" sense. The second context refers to identifying specific observations based on their potential positive impacts in a specific location or region. The two approaches may not always be simultaneously achievable. From the point of view of supporting (for example) military operations at a target site, the second approach would have considerable value.

The implementation of a 'targeted observation control loop' would direct changes in the variety and schedules of data collections, and engage additional assets / sensors to observe at locations where perceived needs are greatest (i.e. where greatest forecast impacts from those data are likely to be realized). The decision to execute a specific observing strategy implementation might be driven by where and when a model predicts rapid significant future development, by where the model forecast shows greatest uncertainty (as revealed in ensemble forecasts), or by where observations reported real-time from the sensorweb reveal deficiencies in model performance. The architecture proposed in this study is especially suited to the implementation of targeted observing strategies. The feed back between the observing system and modeling system enables targeting to actually be carried out!

The dispersion among an ensemble of forecasts over time and by region may offer a relatively simple means for determining where a model prediction is most sensitive or uncertain, and thereby provide clues as to where observations may be beneficial. But the atmosphere and models are highly non-linear; and in actual fact, the development of an effective targeting function may be quite complex. Initial state (analysis) errors at any given location, which derive from errors in the forecast model first guess or from observational errors, eventually propagate, decay or potentially grow, and evolve with time through the model domain in complex and non-intuitive ways … but none-the-less in ways that can be estimated using mathematical-statistical methods. Techniques for estimating where observations are most needed include the Ensemble Transform Kalman Filter (ETKF) (Bishop & Toth, 1996, 1999) that aims to predict the evolution of error covariances, and a Singular Vector (SV) method in which targeting is based on projecting initial errors (and correction thereto) onto rapidly growing modes identified by dominant singular vectors from an ensemble of model runs (Gelaro, et al, 1999).

The efficacy of model-guided targeted observing for synoptic weather systems was demonstrated in the FASTEX, NORPEX, CALJET, WSR99 and WRS00 field programs (Toth, et al, 1998, 1999; Gelaro, 1999; Szunyogh, et al, 1999, 2002). As a result, targeting strategies are being implemented operationally by the National Weather Service relative to Winter Storms (Toth, et al, 2001). The benefits of targeting observing in relation to hurricanes have also been operationally established (Burpee, et al, 1996).

5.3 ASPECTS OF MODEL IMPLEMENTATION

5.3.1 Model Configuration

The assumed atmospheric model configuration for this study was as follows. For year 2025, we assumed a global mesoscale model with 100 vertical levels from the surface up to 80 km or .01 mb. Expressed in height coordinates, this would provide 100 m vertical resolution in the lowest 2km (planetary boundary layer), 250 m upward to 12 km, 500 m resolution up to 15 km, 1km resolution up to 35 km, 2km resolution up to 50 km, and 6 km resolution to 80 km. The lower mesosphere would serve as a buffering transition to the top of the model. It is probably not necessary to represent mesospheric dynamics in any complex way since, except during very occasional episodes, middle atmospheric dynamics would have little impact on tropospheric dynamics over a two-week period. The 250 m vertical grid spacing in the free troposphere was assumed to be able to resolve major structures comparable to what one sees resolved in a *typical* radiosonde profile. The increase in the number of vertical levels only represents a factor of 2 increase over present day models. Of greater significance is the horizontal resolution, since required computing power increases as the square of the increase in resolution: increasing grid spacing (resolution) by a factor of 2 requires 4 times the computing; increasing grid resolution by a factor of 10 requires a hundred-fold increase in computing power.

5.3.2 Computational Demands

Figure 7 presents an analysis of improvements relative to year 2000 computing power that would be needed to operate current DAO (GDAS) and NCEP (eta) models with the configuration described above. It shows that to run current model formulations globally with 100 vertical levels and 1km horizontal resolution in an operational setting would require on the order of a million time increase in computing power. Extension of Moore's law extrapolated to 2025, would offer only a 10^5 increase in computing power.

This analysis is based only on increases in the total numbers of grid points over the entire globe, with a commensurate reduction of time step with increasing spatial resolution. It does not take into account computational demands of more sophisticated parameterizations, coupled surface modeling, advances in numerical techniques, or the multiplying demands of ensemble forecasting. It also does not take into account the much larger computational demands associated with data assimilation. Taking these factors into account, is not difficult to project that we would need 10^8 to 10^9 increase over year 2000 computing power to execute the assumed model and data assimilation system in a useful operational mode.

Model Attributes	**2000 Global Synoptic NOAA**	**2000 Global Mesoscale NOAA**	**2025 Global Mesoscale NOAA**	**2000 Global Synoptic NASA/DAO**	**2025 Global Mesoscale NASA/DAO**	**2025 Global Mesoscale NASA/DAO**	**2025 Global Mesoscale NASA/DAO**
Vertical Levels	42 Levels	50 Levels	100 Levels	55 Levels	100 Levels	100 Levels	100 Levels
Vertical Extent (km) (mb)	43 km 2.0 mb	25 km 25 mb	50 km 0.8 mb	80 km 0.01 mb	80 km 0.01 mb	80 km 0.01 mb	80 km 0.01 mb
Horizontal Extent	global	10% global	global	global	global	global	global
Horizontal Resolution (km)	80 (T170)	22 km	1 km	100 km	25 km	10 km	1 km
Wall-clock time to execute 1 day model forecast	14 day forecast in 98 minutes	14 day forecast in 80 days	14 day forecast in 220 years				
Time Step (min)				30 min	15 min	5 min	0.5 min
Computing Factor	1	1 e+03	1 e+06	1	5 e+01	1 e+03	1 e+06

Figure 7 – Weather Models

Because the European Center for Medium Range Weather Forecasting and the NASA Data Assimilation Office (DAO) has already begun experimenting with a 25 km horizontal resolution global model, this was the coarsest horizontal resolution considered in this study. At 25 km the hydrostatic assumption is still valid. At finer resolutions, the requirement for an additional prognostic equation for vertical motion significantly changes the physics of the system in ways that require the model to be integrated using shorter time steps, and resulting in up to factor of 10 increase in computing. Whether or not a 1km or 10km resolution will be needed is highly situation specific and will depend on the phenomena, location (tropics versus mid-latitudes, etc.). A brute-force approach would be to run the model at 1 km over the entire globe. Although the computing cost is high, the importance of scale interactions in overall development of the atmosphere argues for global 1 km resolution. An alternative would be to run the model globally with a 25 km resolution, but invest in development of adaptive grid techniques that can automatically increase (or decrease) resolution as circumstances warrant. The computational resource savings that might result from reduction (if any) in numbers of grid points would have to be weighed against the overhead of managing, assessing and monitoring adaptive grid processes. Although adaptive grid techniques have been employed extensively in aerodynamics and other engineering design work, the challenges for weather forecast models are much more complex and will require both significant investment in research on useful adaptive grid techniques, adaptive parameterization, and artificial intelligence.

Most meteorological phenomena must be resolved adequately in both horizontal and vertical dimensions in order to be accurately represented observationally and in a model. For a given phenomenon or situation, the choice of horizontal model resolution and the choice of vertical resolution are coupled (Lindzen and Fox-Rabbinovitz, 1989). Therefore, some consideration will need to be given to the intelligence of adaptive vertical grids as well as horizontal grids.

A great deal of the total computing requirement for the model itself is in parameterization of sub-grid scale processes. The formulations of many parameterizations of unresolved processes are both space (and time) scale dependent. By increasing the model resolution, some part or all of some process that may have been previously parameterized by one method at a coarser resolution may need to be extensively reformulated for a finer grid, or perhaps even explicitly calculated on a finer grid. Reformulating some parameterizations may entail a significant scientific research investment. Recognizing that a parameterization is by definition a simplification of a complex process, explicit calculation of a process will almost always come at greater computational price tag than its parameterization. In this study, we see inevitable movement in the direction of higher resolution models and more sophisticated parameterizations of key physical processes. The incremental computational cost of incorporating such improvements in model physics, independent of costs increasing the grid resolution, could levy a 5-fold to 10-fold increase in computational cost of forward integrating the model.

5.3.3 Ensemble Forecasting

An ensemble is collection model forecasts with the same base initial states, except that each has been deliberately perturbed in a manner intended to mimic, in a statistical sense, typical analysis errors based on data type and atmospheric structure. Forecasts from these initial states diverge with time; and the aim is to have selected each ensemble member and in sufficient numbers such that the spread of their forecast solutions gives a true representative sample of the full range (envelope) of possible model trajectories. The manner in which these perturbations are determined is itself an evolving science, and has great bearing on the numbers of ensembles that are required, and thus affects computing resource requirements. The full answer to how many ensemble members are required may depend on many factors including model physics, grid resolution, model domain, and particular configuration and condition of the atmosphere itself. This is an active area of research.

The information provided by ensembles serves a number of purposes. For example, the ensemble mean may be assumed to be the forecast that is most likely to be correct; and the spread about the mean a measure of confidence in the forecast. Statistics derived from the ensemble forecasts also provide measures of reliability of model forecast first guess fields relative to observations, and thus the relative weight given to the first guess in constructing the next initial state analysis. And as already discussed, statistical information derived from properly designed forecast ensembles is useful for carrying out targeted observing.

A good summary of current issues in ensemble forecasting approaches and observation targeting is presented in Hamill et al, (2000), and further discussion is well beyond the scope of this report. The primary point of this discussion is to impress the fact that in the future, operational forecasting and observing strategies will depend not on a single model forecast, but on many, perhaps even hundreds of model forecasts being run in ensemble batches every six to twelve hours, this being essential to obtain optimal forecast performance. These considerations alone suggest the need for up to two orders of magnitude increase in computing.

5.3.4 Assimilation Frequency

A possibility for improving on current operational systems would be to perform the data analysis and assimilation hourly. There would be several advantages.

An hourly assimilation cycle would take better advantage of the proposed continuous global satellite data collections. In current practice, only about 15% of all satellite data are assimilated operationally. While there are quality control issues, most satellite data are culled solely due to the inability of current assimilation (and computing) systems to accommodate the observations. Most operational forecast models are initialized at standard synoptic times -- every 12 hours – with asynoptic data queued in a 3 – 6 hour window up to assimilation time. This means that at least half of the satellite data is too old to be included, and even data that is 3 – 6 hours old may require correction for atmospheric state changes that have occurred during the several hour intervals between the observation time and initialization time, a process requiring expensive 4DVAR techniques.

Assimilated hourly, observational “errors” related to the difference between the assimilation time and actual observation time are bound to be smaller, therefore requiring smaller, less disruptive (model shock) corrections. It will also be easier to detect when and where the model forecast and observations diverge and thus to dispatch additional observations to such locations.

Since the computational cost of data assimilation is largely driven the number of observations being assimilated (K-F assimilation scales as the square of the number of observations), frequent analysis of small amounts of data may in the end be more computationally efficient than infrequent analyses with large amounts of data. There is, however, a counter-argument to the above analysis that will require additional trade studies: It may be that the increase in total numbers of observations available to be assimilated from a g*lobal continuous* space-based (and ground) observing system, will more than offset any expected reduction of computing due to more frequent assimilation.

The trend must be toward more frequent assimilation in order for the benefits of the proposed architecture to be realized. Ideally, a true time-continuous assimilation system will evolve, a concept whose feasibility and benefits have been demonstrated (Ghil, et al., 1979).

5.3.5 Forecast Model Development

For historical and practical reasons, the development of atmospheric prediction models have been formulated based on a set of partial differential equations expressed in eulerian form and solved on basically spherical coordinate system, using (most commonly) Cartesian finite difference methods. More sophisticated methods for solving Navier-Stokes-based equations have been in use in engineering disciplines (notably aeronautics) including finite element methods, particle methods, and cellular automata. Increasingly, numerical solution methods are developed in tandem with advancements in computing hardware architectures and operating systems, trending toward algorithms that operate efficiently in highly parallelized, highly distributed computing environments. Indeed, experts in the computational physics communities have been occasionally derided the primitive state of computational science as applied in atmospheric modeling. One factor behind the slow evolution in atmospheric modeling is the tremendous front-loaded investment the meteorological community has in heritage codes, which are the same basic codes that still underpin operational weather forecasting. So there are cost and risk issues.

At the same time, critics have been known to recant when on closer study, they finally comprehend the challenges involved in real-world numerical weather prediction. Among the complicating factors is the preponderance of non-continuous discrete point processes or step function processes such as convection and cloud formation; irreversible diabatic processes; and the reliance on parameterization. Advances are being made as the meteorological research and operational communities (NOAA, NASA, NCAR) are coming together to develop new climate and weather modeling frameworks, using best practices in model architectures and numerical formulations that will eventually find their way into operations well before 2025.

5.3.6 Operational Forecast Production Cycle

Figure 8 illustrates how the two-week forecast operations might be integrated with daily short-range and medium-range forecasting operations. All forecasts would be underpinned by an hourly assimilation cycle, compared with six to twelve hours in current operational forecast and assimilation systems. Initial states derived hourly would serve as the starting point for short, medium and long-range forecasts. What is essentially new is that a significant number of model ensembles would have to be calculated that would provide the data and statistics needed to a) support construction of statistically optimal hourly-updated initial states, b) to support adaptive and targeting observing operations, c) to provide a “statistically most likely” average forecast for final public dissemination, and d) to provide more useful confidence limits for the forecasts. Currently, NCEP and ECMWF use 17-20 and 50-100 ensemble members for their mid-range and long-range forecasts, respectively. We should anticipate 250 ensemble members as the minimum to support the 10-14 day long-range forecast in a 2025.

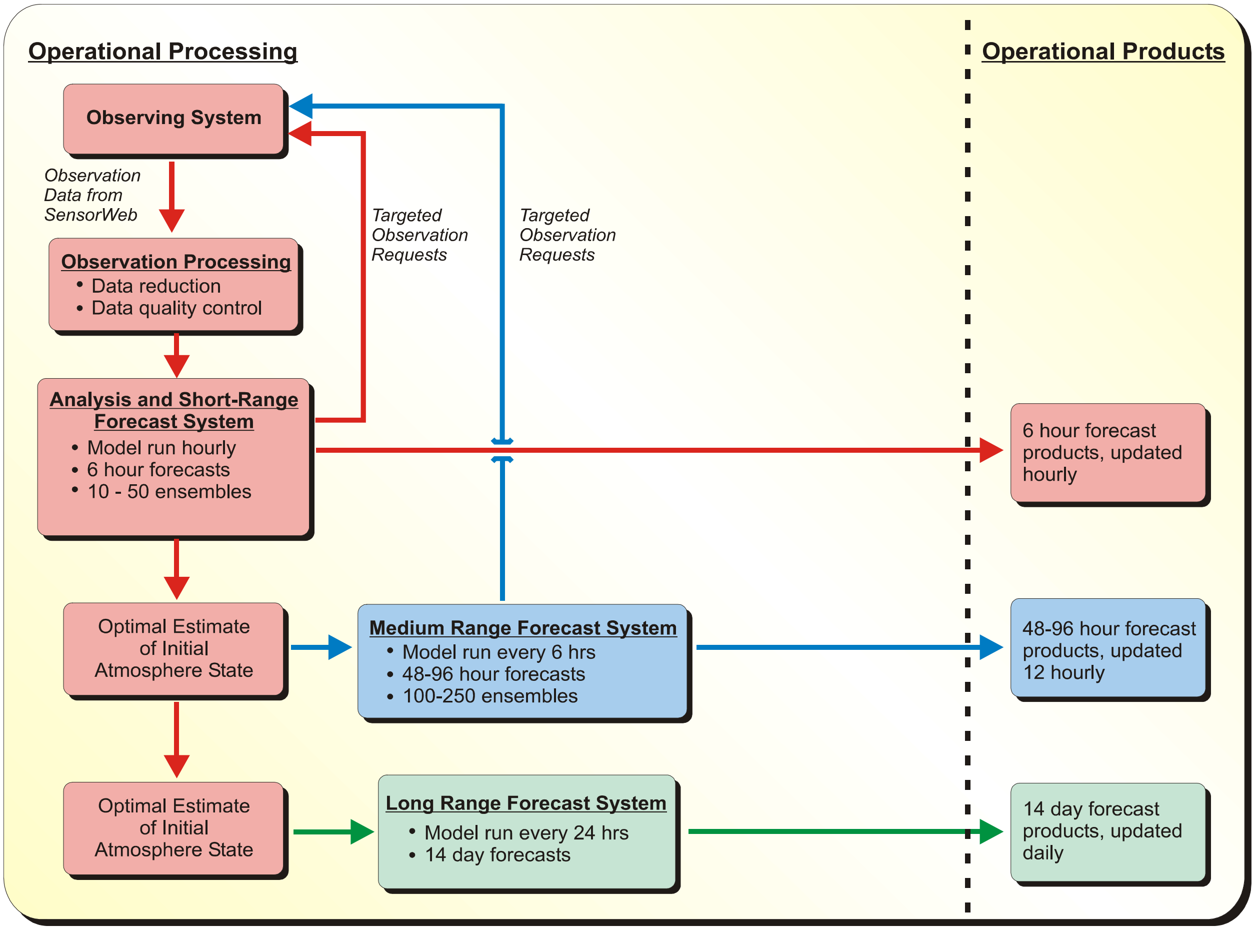


Figure 8 – Operational Forecast Production Cycle

The number of ensembles and the lengths of the ensemble integrations to support six-hour short-range forecasts may be fewer (10 – 50) and shorter (6-12 hours). The short-term (6-hour) forecast, whether run within a regionally nested model or potentially a global model at 1 – 10 km resolution, is intended to capture and keep the model current with respect to rapidly evolving local and regional events and processes. Therefore, the targeting observations and observing system feedbacks may be quite different from those needed to support the medium and long-range forecasts, and the construction of ensembles may also be quite different. Thus, fig. 8 shows two separate targeting feedback loops from the short range and medium range forecast models. The length of the (potentially hundreds of) ensemble forecasts that support the 10 – 14 day forecast requirement are probably integrated out at least ten-days.

5.3.7 Observing System Simulation Experiments (OSSEs)

Technology alone will not solve the long-range weather forecast problem. Equal investments in science and technology are essential. Bringing scientific research and technology together the right way will necessitate an organized and consistent means for assessing rigorously the probable impacts of future instruments on numerical forecasting. An infrastructure to support a centralized test-bed system for Observing System Simulation Experiments (OSSEs) is necessary in order to assess the impacts of various data types, data quality and availability against model configuration and assimilation method (Atlas, 1997). In a first step to develop that research infrastructure, NASA and NOAA, recently have embarked on joint endeavors to coordinate their data assimilation efforts (Einaudi, et al., 2001).

6.0 COMPUTING SYSTEMS

Today's GigaFLOPS capabilities are already saturated. Based on projections of complexity of future modeling and assimilation system and operations, it is almost certain that there will be significant gaps between the computing needs and the expected capabilities. Significant advancements in computer technology are required. The following sections consider the areas where IT capability advances may be gained. Evolutionary improvements to current technologies increase capabilities in a (mostly) steady and predictable way. Revolutionary improvements resulting from technology breakthroughs allow a stepwise (and often unpredictable) increase in capability.

6.1 Evolutionary Improvements

Moore's "Law" of increasing transistor density (doubling every 18 months, and often related to computing power) is anticipated to fail around 2015 due to limits having been reached in optical lithographic methods of microchip manufacture. However, new technologies now being developed may extend the applicability of Moore's law beyond 2020. EUV and Electron beam technologies overcome limitations of optical lithography (2) – these can create smaller features, well below the current limitations of 120-150 nm feature size. Nanotechnology (3) (4) will be able to produce extremely small features (tens of nm) by nanoassembly. IBM recently produced arrays of transistors based on carbon-nanotubes that are orders of magnitude smaller than current silicon-based transistors.

Since a minimum feature size will be achieved at some point (2020-2025), a way must be found to continue to provide a higher transistor density. "Stacking" layers of chips to achieve a higher density (6) has been considered, although design issues remain with regard to thermal/cooling difficulties and inter-layer communications. Advancements in this technology might provide aid in extending the applicability of Moore's law.

There is a caveat to Moore's (first) Law that must be recognized. Moore also observed that the cost of chip fabrication *facilities* also increases exponentially, doubling every 3 years or so (7). With current semiconductor manufacturing facilities costing upwards of $\$10^9$, a point may be reached at which it will become economically prohibitive to pursue commercial development and manufacture of next generation semiconductors, without comparable development of new markets.

The computational capability and performance of MPI (U.S.) versus vector (Japan) supercomputers are also being re-examined in particular for meeting the needs of global climate and weather prediction (5). A return to Vector Computing may result in performance enhancements if access to vector computers is available.

6.2 Revolutionary Improvements

Revolutionary changes are by definition difficult to predict. However, early developments in promising technologies may provide a glimpse of the future of high end computing.

In the next 10-15 years Hybrid Technology Multithreaded (HTMT) computing is expected to provide 10^2 – 10^3 improvements, well into petaflop range (8). A joint project between NASA/JPL, NSA, NSF, and DARPA, HTMT is likely to result in advances through several new technologies. Superconducting logic will provide clock speeds of 100 – 200 GHz (to as high as 700 GHz) while greatly reducing power requirements to 50 watts (for Petaflop performance) – of course, a disadvantage is the need for cryogenic cooling. Dynamic adaptive resource management and processor-in-memory technology to decrease contention/latency/overhead/starvation and improve load-balancing. Holographic Memory storage (for virtual memory) is 100-times faster than conventional disk drives and has a data density of 10 Gbits in a few cubic cm. Optical Communications has 100 times the per-channel bandwidth of conventional wire systems.

Beyond 2020, optical computer technology could provide even greater performance increases (9) (10) (11). Using these techniques, a potential of a factor of 10^5 – 10^8 improvement over current silicon technology is possible based on work being done at NASA/MSFC, University of Colorado, and Rome Laboratory. Optical computing gets its performance improvements due to three primary enabling technologies:

- Optical switching (logic), instead of using electrical signals in silicon or even superconductors, optical switching uses laser pulses in an optical medium to perform logic.
- Holographic Memory Storage similar to the HTMT
- Optical Communications similar to HTMT

Another long-range technology (20+ years) will be in the development of quantum computing (12). Improvements in computing speed should be drastic, but are extremely difficult to ascertain. Because of the unusual nature of quantum bits (which can take multiple values at the same time), quantum computing seems better suited to certain types of processing. Some problems (such as factoring large numbers) could see a 10^{15} – 10^{20} improvement or better. Other problems, such as numerical weather prediction modeling may or may not be suited for execution on this type of computer. Perhaps some portion of the forecast process (e.g., the analysis) may be adaptable to quantum computing, while the NWP portions might have to remain on other computing architectures.

One final technology to consider is that of molecular (sometimes called biological) computing (13). Also in the long range (20+ years), molecular computer technology will yield extremely small and cheap processors that could be self-assembling. While this technology will probably result in speed improvements over today's systems, it is not anticipated to be on the order needed for future weather modeling needs. Their chief advantages are size and cost. UCLA recently built a molecule-size switch, based on the ability to give or receive an electron. Such a switch (in conjunction with many more like it) could form the basis of a rice-grain size computer that can be manufactured using chemical processes.

6.3 Onboard Computing Versus Ground-Based Computing

Our working assumption, following an experiential rule of thumb is that space deployable computer processing will have 10% of the processing available on the ground. Given expected evolutionary developments and probable breakthroughs in computing, for our applications onboard computing needs will be met in the 2025 time frame. However, we also assume that whatever the computing technology, it will always be considerably more expensive to implement in space than on the ground, perhaps by a factor of ten.

7.0 SPACE SYSTEMS

7.1 Guidance, Navigation and Control

Multiple Guidance, Navigation and Control (GN&C) challenges must be addressed in order to implement the envisioned space platform elements of the next generation weather architecture described in this White Paper. The current State-of-the-Art GN&C systems will not satisfy many of the "next-generation" weather satellite spacecraft GN&C functions, including the associated weather instrument stabilization, pointing and tracking functions. Preliminary analysis indicates that the potential enabling GN&C technologies required will include, but are not limited to, the following:

- Instrument stabilization, pointing and tracking
- Low power, low cost, compact GPS-based systems and components for autonomous spacecraft time, position and attitude determination
- Innovative multifunction systems and components that consolidate the capabilities of two or more traditional spacecraft subsystems such as flywheel energy storage systems for the integrated management of spacecraft momentum and electric power states.
- "Drag Free" orbit control system technology for autonomous LEO platform orbit maintenance

7.2 Distributed Space Systems

The essence of the SensorWeb concept is a Distributed Space System (DSS) which time-synchronously covers a substantial area of the Earth and atmosphere, and processes and disseminates the data as necessary to ensure quality information available to the user. Among the very substantial and unique engineering and technology capabilities required of a DSS are: (1) formation sensing and control, (2) inter-satellite communications, (3) constellation management and mission operations, (4) data fusion, processing, and analysis, (5) mission synthesis, design, and validation, and (6) miniaturized spacecraft technology. The manner in which each of these technologies is brought to bear, will be defined to various degrees by the problem-specific implementation of the DSS. In terms of enabling system through-put, the DSS must possess the means to sense (measure) and maintain relative positions and orientations of observing platforms and sensors, communicate and transmit science and engineering data, perform with the necessary level of autonomy, convert data from multiple sources into useful information, and perform the end-to-end analysis, mission, definition, and systems engineering, and to enable launch of all of the DSS elements within reasonable budgets. The drivers will be requirements for: timeliness of the data, level of coordination among spacecraft, mix between on-board vs. ground processing, maintaining "synthetic distributed apertures", and knowing relative positions and orientations for science processing, to name a few. For example, if hyperspectral data with 256 or more bands is required, and such must be shared among the spacecraft or transmitted to the ground in raw form, Gigabit-order communications may be required. If stereographic imaging or interferometry is required, measurement and in some cases, control, of vehicles relative to one another may be required at the sub-meter down to sub-micron level. If constant baselines between spacecraft must be maintained for substantial durations, then specialized formation control algorithms will be needed along with high-specific impulse thrusters. If decision-making is required at the top-level without human-intervention, for the sake of cost or timeliness, then autonomous constellation management and fault-detection algorithms will be needed. The key to refining the specific requirements will be in performing end-to-end mission design, simulation, and analysis, and requirements allocation from the fundamental scientific needs.

7.3 Maintenance of Orbital Data Sources

A significant portion of the SensorWeb -- the space-based segment -- proposed for collecting data for the M/DAS would consist of a large number (~30) of spacecraft deployed in a Low Earth Orbit (LEO) constellation spread around the Earth. The architecture of the space assets has multiple copies of an instrument on different spacecraft orbiting simultaneously to provide the coverage required.

This space system needs to be robust enough to gracefully degrade if these data sources are, at any given time, less than 100% operational; and the architecture needs to be such that when any one instrument ceases to provide data, the coverage from the others allows a degraded system operation. Also, the architecture should be such that the "cost-effective" replacement of failed instruments to restore 100% system operation is possible. Methods to accomplish this include: (1) inexpensive, throw away spacecraft carrying only one instrument so that when that instrument fails, the spacecraft can be de-orbited and replaced; (2) "permanent" platforms carrying more than one type of instrument and so designed to be serviced in orbit to replace a failed instrument or a platform subsystem. More study remains to be done as to which of these two options would be more cost effective. Much work is already being done in this area across several organizations (NASA, DoD) and should be leveraged if this area is further explored.

8.0 APPLICATIONS AND KNOWLEDGE DISSEMINATION SYSTEMS

The potential applications will be as wide and varied as the users. However, all users have a common interest in risk reduction, whether in business or personal activities. For example, in the agricultural arena the ability to reasonably anticipate precipitation, humidity, and temperature extremes 10 – 14 day forecast in advance would have immeasurable value to farm operations, from planting, to chemical applications, to herd management. Longer-range forecasts would re-shape agricultural commodities and futures markets. Based on more reliable long-range model forecasts and attendant improvements in accuracy of shorter range forecasts, transportation concerns from airlines, to land and ocean-going shipping would realize tremendous benefits by having more confidence and lead time to manage logistical operations (e.g. routing) in anticipation of potential weather related disruptions. In all cases, forecasts derived from model outputs might be tailored to the customer by any number of commercial value-added vendors and disseminated via a satellite direct broadcast system under a subscription service.

Products must be distributed to the users. An open architecture wireless network capable of reaching multiple users would easily meet this need. Figure 9 shows a notional depiction of such a system. The model data from the M/DAS is provided to a commercial data production facility where the user-specific weather knowledge products are created. Using the global wireless network, they routed to their intended users, who would receive them using a variety of means ranging from aircraft navigation computers down to handheld devices.

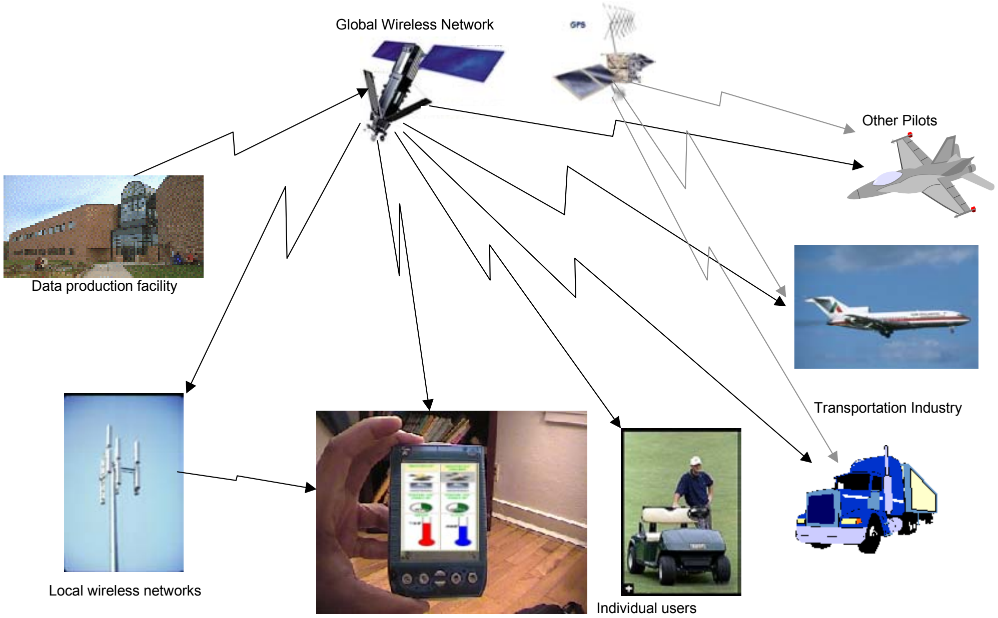


Figure 9 Open Architecture Wireless Network for Distribution of Weather Knowledge Products

REFERENCES

Atlas, David and C. Laurence Korb, 1981: Weather and Climate Needs for Lidar Observations from Space and Concepts for Their Realization. *Bull. Amer. Meteor. Soc.*, **62**, 1270 - 1285.

Atlas, Robert, R.N. Hoffman, S.M Leidner, J. Sienkowicz, T.-W. Yu, S.C. Bloom, E. Brin, J. Ardizzone, J. Terry, D. Bungato, and J.C. Jusem, 2001: The Effects of Marine Winds from Scatterometer Data on Weather Analysis and Forecasting. *Bull. Amer. Meteor. Soc.,***82**, 1965 – 1990.

Atlas, Robert, 1997: Atmospheric Observations and Experiments to Assess Their Usefulness in Data Assimilation. *J. Meteor. Soc. Japan*, **75**, 111 – 130.

Baker, W., G. Emmitt, F. Robertson, R. Atlas, J. Molinari, D. Bowdle, J. Paegle, R. Hardesty, R. Menzies, T. Krishnamurti, R. Brown, M. Post, J. Anderson, A. Lorenc and J. McElroy, 1995: Lidar-Measured Winds from Space: A Key Component for Weather and Climate Prediction. *Bull. Amer. Meteor. Soc.,* **76**, pp 869 – 888.

Benjamin, S.G. and B.E., Schwartz, 1999: Accuracy of ACARS Wind and Temperature Observations Determined by Collocation. *Weather Forecasting*, **14**, 1032 – 1038.

Bishop, C. H. and Z. Toth, 1996: Using ensembles to identify observations likely to improve forecasts . *Preprints* , 11th AMS Conference on Numerical Weather Prediction, Norfolk, Virginia.

Bishop, C.H. and Z. Toth, 1999: Ensemble Transformation and Adaptive Observations. J. Atmos. Sci., **56**, 1748 – 1765.

Burpee, R., J. Franklin, S. Lord, R. Tuleya and S. Aberson, 1996: The Impact of Omega Dropwinsondes on Operational Hurricane Track Forecast Models. *Bull. Am. Meteor. Soc.*, **77**, pp 925 – 933.

Davis, Shannon, Mark Bourassa, Robert Atlas, Joe Ardizzone, Eugenia Brin, James O'Brien and David Zierden, 2002: Sea Surface Pressure Fields from NASA's Sea Winds Scatterometer and Their Impact on NWP., Preprint Volume of Symposium on Observations, Data Assimilation and Probabalistic Prediction, American Meteorological Society, Orlando, FL . pp. 14 – 17.

Delin, K.A. and S.P. Jackson, 2001: The Sensor Web: A New Instrument Concept. SPIE Symposium on Integrated Optics, San Jose, CA, January 2001. pp **X-X**

Einaudi, F., L. Uccellini, J. Purdom, D. Rodgers, R. Gelaro, J. Dodge, R. Atlas and S. Lord, 2001: Weather Prediction Improvement Using Advanced Satellite Technology. IEEE Geoscience and Remote Sensing Symposium (IGARSS), Sydney, Australia, June 2001.

Emmitt, G.D: A Geosynchronous Lidar System for Atmospheric Winds and Moisture Measurements. IEEE Geoscience and Remote Sensing Symposium (IGARSS), Sydney, Australia, June 2001.

Flesia, C. and C.L. Korb, 1999: Theory of Double-Edge Molecular Technique for Doppler Lidar Wind Measurement. *Applied Optics*, **38**, 432 – 440.

Gelaro, R., R.H. Langland, G.D. Rohaly and T.E. Rossmond, 1999: An Assessment of the Singular-Vector Approach to Target Observations Using the FASTEX Dataset. *Quart. J. Roy. Met. Soc.*, **125**, 3299 – 3328.

Ghil M., M. Halem, and R. Atlas, 1979: Time-Continuous Assimilation of Remote-Sounding Data and Its Effect on Weather Forecasting. *Mon. Wea. Rev.*, **107**, 140 – 171.

Hamill, Steven Mullen, Chris Snyder, Zoltan Toth and David Baumhefner, 2000: Ensemble Forecasting in the Short to Medium Range: Report from a Workshop. *Bull. Amer. Meteor. Soc.*, **81**, 2653 – 2664.

Ingman, Paul, 2001: Lidar Wind Measurements – The Atmospheric Dynamics Mission, CGMS Workshop on Long-Term Future of the Basic Sounding and Imaging Missions, 23-24 April, WMO, Geneva, Switzerland.

Jackson, T., 1993: “Measuring surface soil moisture using passive microwave remote sensing,” Hydrological Processes, 7:139-152.

Korb, C L. and C.Y. Weng, 1982: A Theoretical Study of a Two Wavelength Lidar Technique for the Measurement of Atmospheric Temperature Profiles*. J. Appl. Meteor.*, **21**, 1346 – 1355.

Korb, C. Laurence and Chi Y. Weng, 1983: Differential Absorption Lidar Technique for Measurement of the Atmospheric Pressure Profile. *Applied Optics*, **22**, 3759- 3770.

Korb, C. L., G. Schwemmer, M. Dombrowski, and C. Y. Weng, 1989: Airborne and Ground-Based Lidar Measurements of the Atmospheric Pressure Profile. *Applied Optics*, **28**, 3015 – 3020.

Kuo, Y. H., S. V. Sokolovskiy, R. A. Anthes and F. Vandenberghe, 2000: Assimilation of GPS Radio Occultation Data for Numerical Weather Prediction, *Terr. Atmos. Ocean. Sci.*, **11**, 157 – 186.

Kursinski, E.R., G.A. Hajij, J.T. Schofield, R.P. Linfield and K.R. Hardy, 1997: Observing Earth’s KAtmosphere with Radio Occultation Measurements using the Global Positioning System. *J. Geophys. Res.*, **102**, 429 – 465.

Lemmeran, L., K. Delin, F. Hadaegh, M. Lou, K. Bhasin, J. Bristow, R. Connerton and M. Pasciuto: “Earth Science Vision: Platform Technology Challenges”. IEEE Geoscience and Remote Sensing Symposium (IGARSS), Sydney, Australia, June 2001.

Lindzen, R.S. and M. Fox-Rabbinovitz, 1989: Consistent Vertical and Horizontal Resolution. *Mon. Wea. Rev.*, **117**, 2575 – 2583.

Lorenz, E.N., 1963: Deterministic Non-Periodic Flow. *J. Atmos. Sci.*, **20**, 130-141.

Lorenz, E.N., 1969: The Predictability of a Flow Which Possesses Many Scales of Motion. *Tellus*, **21**, 289 - 307.

Miller, S.D., and G.L. Stephens, 2001: Cloudsat Instrument Requirements As Determined From ECMWF Forecasts Of Global Cloudiness. *J. Geophys. Res.*, **106**, 713-733.

O' C. Starr, D., C.L. Korb, G.K. Schwemmer, and C.Y. Weng, 1992: Observations of Height-Dependent Pressure Perturbation Structure of a Strong Mesoscale Gravity Wave. *Mon. Wea. Rev*., **120**, 2808 – 2820.

O'Neill, P., N. Chauhan, and T. Jackson, 1996: "Use of active and passive microwave remote sensing for soil moisture estimation through corn," Int. Journal of Remote Sensing, 17: 1851-1865.

Palmer, P.I., J.J. Barnett, J.R. Eyre, and S.B. Healy, 2000: A Non-linear Optimal Estimation Inverse method for Radio Occultation Measurements of Temperature, Humidity and Surface Pressure. *J. Geophys. Res.,* **105**, 513 – 526.

Poli, Paul, Joanna Joiner, and E. Robert Kursinski, 2002: 1DVAR Analysis of Temperature and Humidity using GPS Radio Occultation Refractivity Data. *Submitted to JGR*.

Poli, Paul and Joanna Joiner, 2002: Assimilation Experiments of One-Dimensional Variational Analyses with GSP/MET Refractivity. Proceedings 1st CHAMP Science Meeting, Potsdam, Germany, January 21 – 25, 2002.

Schmugge, T., T. Jackson, W. Kustas, and J. Wang, 1992. "Passive microwave remote sensing of soil moisture: results from HAPEX, FIFE, and MONSOON '90," ISPRS Journal of Photogrammetry and Remote Sensing, 47: 127-143.

Shue, R.-S., L. Yubao, L.-N. Simon and L. Carson, 2002: The Impact of Assimilating ACARS Data on the Performance of a Real-Time FDDA Weather Analysis and Forecasting System. Preprint Volume of Symposium on Observations, Data Assimilation and Probabalistic Prediction, American Meteorological Society, Orlando, FL . pp. 26 – 28.

Soil Moisture Mission Working Group, 2000. "Soil Moisture Mission Workshop Report," workshop held September 12-13, 2000, Las Vegas, NV, NASA Land Surface Hydrology Program (now Terrestrial Hydrology Program), http://lshp.gsfc.nasa.gov.

Stephens, G.L., G. V. Vane and S.J. Walker, 1998: The Cloudsat Mission: A New Dimension to Space-Based Observations of Cloud in the Coming Millennium. GCSS-WGNE Workshop on Cloud Processes and Feedbacks in Large Scale Models. November 9-13, 1998. (from http://cloudsat.atmos.colostate.edu)

Stephens, G.L. D.G. Vane, R. Boain, G. Mace, K.Sassen, Z. Wang, A. Illingworth, E. O'Connor, W. Rossow, S.L. Durden, S. Miller, R. Austin, A. Benedetti, C. Mitrescu and the CloudSat Science Team. "The CloudSat Mission and the EOS Constellation: A New Dimension Of Space-Based Observations Of Clouds And Precipitation. *Bull. Amer. Meteor. Soc.*, *Submitted.*

Szunyogh, I., Z. Toth, S. Majumdar, R. Morss, C. Bishop, and S. Lord, 1999:Ensemble-based targeted observations during NORPEX. Preprints of the 3rd Symposium on Integrated Observing Systems,10-15 January 1999, Dallas, Texas, in print.

Szunyogh, I., Z. Toth, S. J. Majumdar, and A. Persson: On the propagation of the effect of targeted observations: The 2000 Winter Storm Reconnaissance Program (Monthly Weather Review, under review).

Toth, Zoltan, I. Szunyogh, K. Emanuel, C. Snyder, J. Woollen, W. Wu, T. Marchok, R. Morss and C. Bishop, 1998: Ensemble-based targeted observations during FASTEX. *Preprints* , 12th AMS Conference on Numerical Weather Prediction, Phoenix, Arizona.

Toth, Z., I. Szunyogh, S. Majumdar, R. Morss, B. Etherton, C. Bishop, S. Lord, M. Ralph, O. Persson, and Z.-X. Pu, 2000: Targeted observations at NCEP: Toward an operational implementation. Preprints of the 4th Symposium on Integrated Observing Systems, 10-14 January 2000, Long Beach, Ca, in print.

Toth, Zoltan, Istvan Szunyogh, Craig Bishop, Sharan Majumdar, Rebecca Morss and Stephen Lord, 2001: The Use of Targeted Observations in Operational Numerical Weather Forecasting. American Meteorological Society. Preprints of the 5th Symposium on Integrated Observing Systems. January, 2001, Albuquerque, New Mexico, pp 72 – 79.

Whiteman, DS.N., S.H. Melfi and R.A. Ferrare, 1992: Raman Lidar Systems for the Measurement of Water Vapor and Aerosols in the Earth’s Atmosphere. *Applied Optics*, **31**, 3068 – 3082.

Turner, D. D. , W.F. Feltz, and R.A. Ferrare, 2000: Continuous Water Vapor Profiles from Operational Ground-Based Active and Passive Remote Sensors. *Bull. Amer. Meteor. Soc*., **81**, 1301 – 1317.

World Meteorological Technical Document (WMO/TD) No. 1052, 2001: “Statement of Guidance Regarding how well Satellite and in situ Sensor Capabilities meet WMO user Requirements in Several Application Areas”.

REFERENCES from Section 6.

(1) ESTO Improving Weather Prediction Study: Analysis Approach Presentation, Study Lead – Mark Steiner, SESAC Code 731, July 25 2001

(2) "MPR's Technology of the Year Award (EUV Safeguards Moore's Law for Another Decade)", Microdesign Resources, 22 Jan 2001

(3) "Beyond Silicon", Boomernomics, http://www.boomernomics.net/VitaminB/may0201.htm, 2 May 2001

(4) "Nanotechnology and the Next 50 Years", Presentation to University of Dallas – Board of Concilors, 7 Dec 1995

(5) "Climate Modeling", White Paper by R. Rood, undated

(6) "Power, Cost Inhibit Nanometer-Scale Designs", EEZine, 22 Jun 2001

(7) "VLSI Development", Electronic News, 27 Sep 1993

(8) "A Hybrid Technology Multithreaded (HTMT) Computer Architecture for Petaflops Computing", JPL Paper, 31 Jan 1997

(9) "The Growing Role of Photonics in the Computer Revolution", The Photonics Design and Applications Handbook, 1998

(10) "Optoelectronic Architecture for High-Speed Switching and Processing Applications", The Photonics Design and Applications Handbook, 1998

(11) Recent Advances in Photonic Devices for Optical Computing, NASA/Marshall Paper, Apr 2000

(12) "IBM Quantum Computer Solves Code Cracking Problem", EEZine, 22 Aug 2000

(13) "Biochips: Pushing Moore's Law Towards a New Horizon", University of Texas Southwest Medical Center Paper, undated

(14) "Challenges of High-End Computing, Lawrence Berkeley Laboratory Paper, undated 1997

(15) "Central Satellite Data Processing", NOAA/NESDIS briefing November 5, 1999.

(16) P.M. Lyster. "The Computational Complexity of Atmospheric Data Assimilation", July 4, 2000. http://dao.gsfc.nasa.gov/DAO_people/lys/complexity/

(17) Rood Climate Modeling Paper

(18) C. Staton, "NCEP High Performance Computing System: Strategic Goals"

(19) L. Lemmerman, *et al.*, "Earth Science Vision: Platform Technology Challenges,"

(20) Golden White Paper, "Computational Climate Modeling," May 25, 2001

(21) Discussions with Science Experts…. These will be documented and listed here.

(22) Central Satellite Data Processing, NOAA/NESDIS briefing, November 5, 1999

APPENDIX A

SATELLITE CONSTELLATION ARCHITECTURE – Providing Global Coverage

Swath Width for 1 Hour Coverage By Various Instruments

To provide global 1-hour coverage would require multiple satellites in phased circular orbits with identical altitude and inclination but with different ascending nodes. The ascending nodes and phasing would be such that each satellite follows another satellite along the same ground track but one hour later, ensuring that all areas of the earth are sampled at least hourly. Over high-latitude Polar Regions where orbits overlap, satellite data sampling would be coordinated to avoid excessive redundancy. The swath width of the instruments and the orbit determines how many satellites are needed to provide a global 1-hour data set.

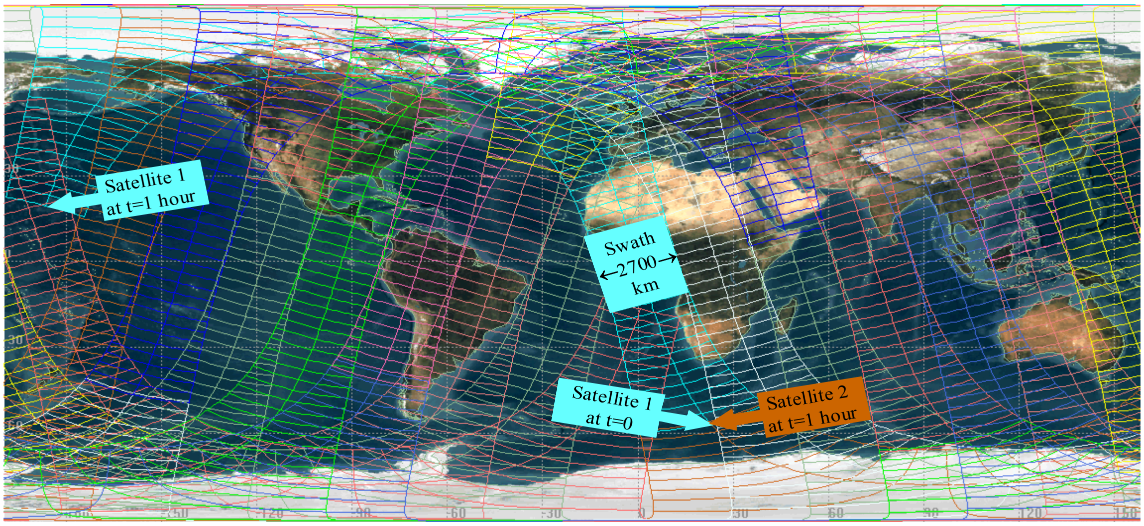


Figure 1 – 13 Satellites With 2700 km Swaths Provide 1-Hour, Global Coverage

In the present analysis, the inclination of the orbit is selected so that the earth's poles can be seen. There are two solutions for this – posigrade (inclination <90 degrees) and retrograde (inclination >90 degrees). The advantage of the posigrade orbit is a propulsion advantage since the earth is rotating in the same direction. The advantage of the retrograde orbit is that a given instrument swath will cover more longitude along the equator. For some swath widths this reduces the number of satellites required to provide the desired coverage. See Figure 2.

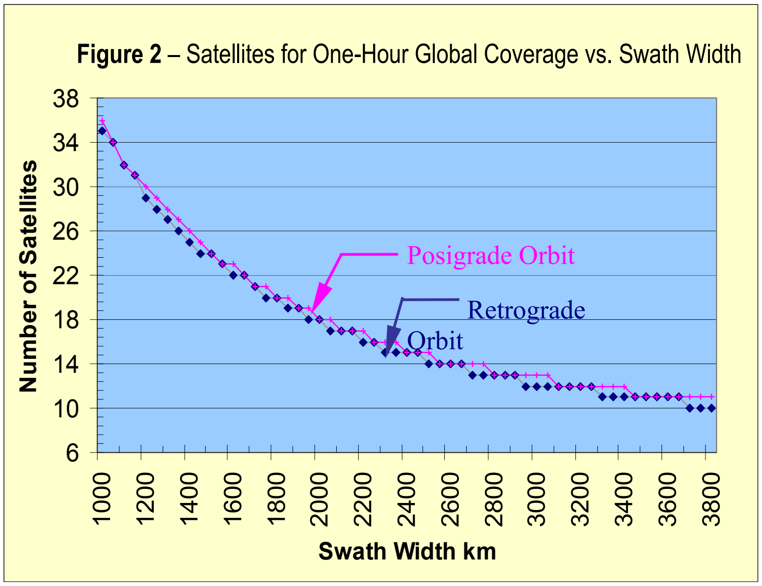


Figure 2 – Satellites for One-Hour Global Coverage vs. Swath Width

Swath Width for 3 Hour Coverage By Laser Wind Measurements

A Laser wind profiler concept determines atmospheric wind by transmitting range-gated pulses of laser light and then processing a return backscattered signal (from aerosols, clouds or air molecules) for a Doppler shift. The locations of the wind measurements are along the line-of-sight of the laser beam. The time between transmission and the return determine the distance to the sample. The velocity of the wind at each location along the laser beam is determined by comparing the laser frequency that was transmitted with the laser frequency that was received at the time delay corresponding to the distance to volume of atmosphere to be sampled. The frequency difference (or “Doppler shift”) is proportional to the relative velocity between the satellite and the volume of the atmosphere being sampled.

The Doppler shift caused by the atmospheric wind is actually only one part of the apparent relative velocity. To determine the actual wind velocity, both the component of the satellite’s velocity and the component of the earth’s rotational velocity along the line-of-sight must be subtracted. The calculations requires that several parameters be known:

- The satellite position and velocity must be accurately known in earth–fixed coordinates. This can be accomplished using a multi-frequency GPS receiver.
- The off-nadir angles of the transmitted and received line-of-sight for the Laser Wind Instrument must also be known to high accuracy. There are three major components to the line-of-sight determination:
    - The satellite’s attitude system will measure and control the satellite’s attitude using startrackers and gyros.
    - The attitude bias error between the instrument telescope line-of-sight and the attitude control system can be determined by occasionally pointing the telescope at a star. This line-of-sight calibration would be scheduled to occur when the instrument is at high latitudes and another Laser Wind Instrument on another nearby satellite can provide the wind data.
    - The instrument pointing gimbals will measure their angles very accurately. The satellite control computer will provide the instrument with angle commands that it will use to accurately point the instrument.

To determine the horizontal wind vector, a volume of atmosphere must be lased from multiple view angles -- ideally from two orthogonal directions. (The vertical component of wind is very small and can be assumed to be zero.) This can be approximated from a single satellite by taking several measurements of the same volume as the satellite passes. For instance, a Laser Wind Instrument could take a measurement at some range fore and 45 degrees to the side of the satellite track, then again at shorter range as the satellite passes the target sampled volume, and then again at some range aft of the satellite and 45 degrees to the side. The nearly orthogonal fore and aft measurements would determine the two-dimensional horizontal wind vector.

This technique has limitations. An atmospheric volume on the satellite’s ground track can measure wind components only along the direction of the ground track, but no cross-track wind component. Fortunately, flying Laser Wind Instruments on other satellites in nearby orbits can rectify this deficiency. Consider a pair of satellites flying identical high inclination orbits and both crossing the equator together, but spaced apart along the equator (See Figure 3). The Laser Wind Instrument on each satellite will measure the along track wind in a volume of atmosphere directly ahead and below and then, looking to the side, directly under the other satellite. Thus, pairs of satellites observing nearly orthogonal to each other in coordinated fashion, could accurately resolve wind vectors located along and between their two ground tracks.

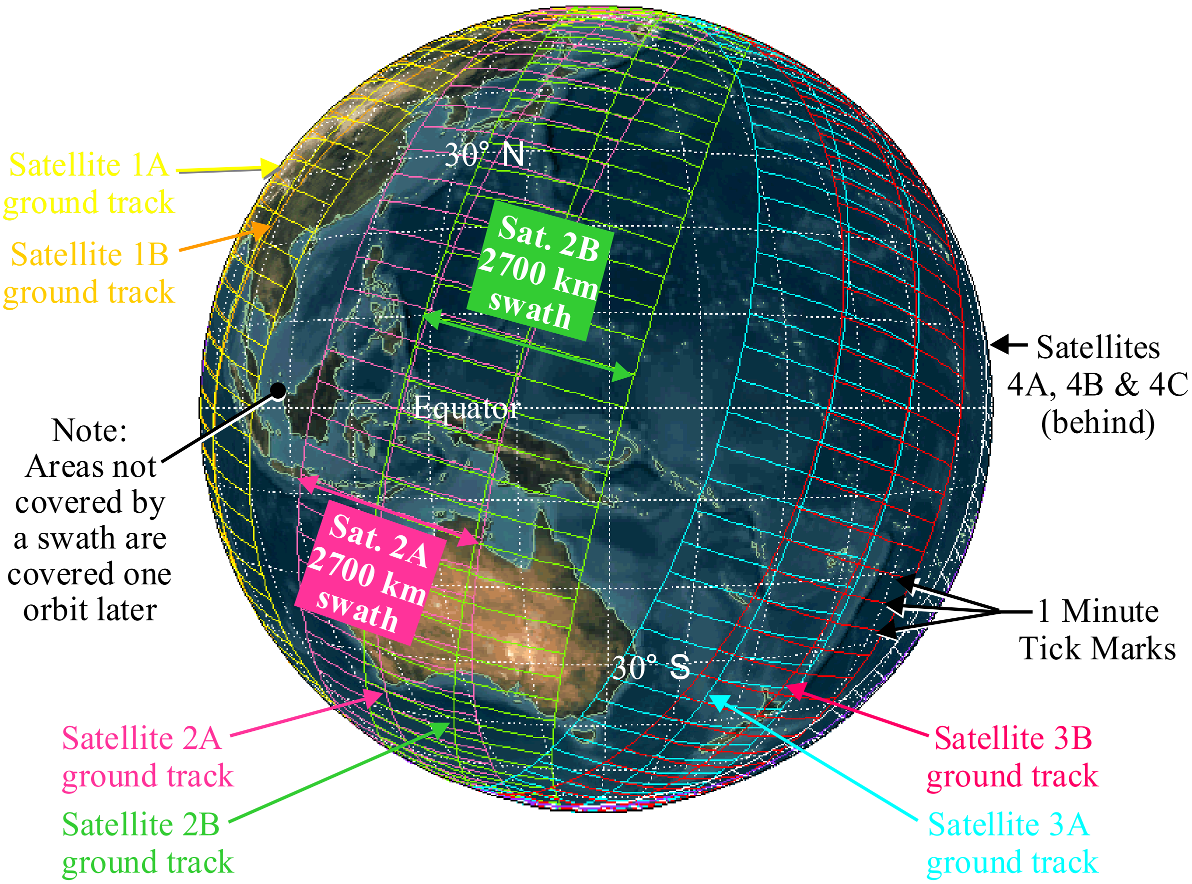


Figure 3 – Satellite-A Sees Satellite-B's Ground Track and Vice Versa (Part of One Orbit Shown)

Each instrument would still take single line-of-sight winds in the direction away from its paired satellite. Such wind measurements would be either assimilated as line-of-sight winds, or combined with orthogonal laser shots from earlier or later laser wind instrument passes from a neighboring pair of satellites. The assimilation process will be able to optimally combine all observed wind information inputs with model fields to arrive at accurate 2-dimensional wind field.

At latitudes above ~63N and below ~63S the worst case time separation between the earlier and later measurements becomes just a few minutes because there is nearly complete overlapping coverage among all the satellites. This is illustrated in Figure 4. Where overlapping coverage exists, the Laser Wind Instruments on all the satellites will be programmed to gather the needed wind information as efficiently and accurately as possible while minimizing the number of laser shots.

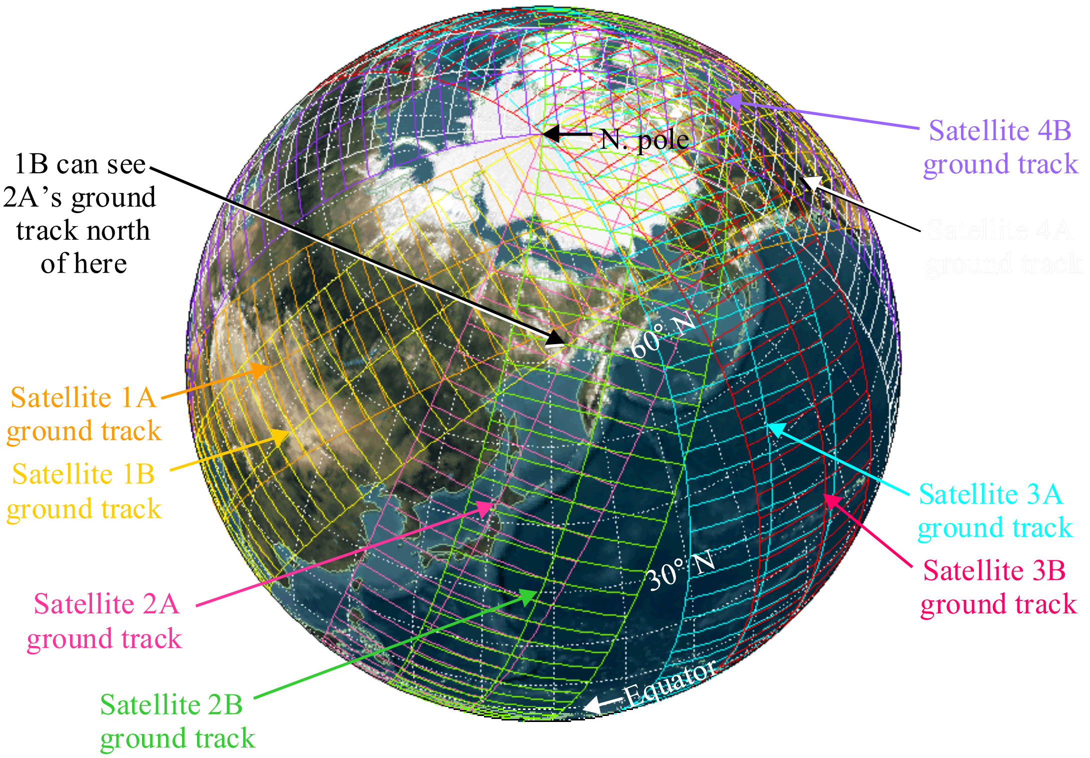


Figure 4 – Redundant Coverage in Polar Regions by 4 Pairs of Satellites

Assuming a 2700 km swath (as in Figures 1 and 2), four pairs of Laser Wind Instrument satellites (8 total) are required to provide global 2-dimensional winds every 3 hours. All satellites cross the equator at essentially the same instant of time, but their ascending nodes would be spaced around the earth to give the global data set that is needed.

If a Laser Wind Instrument with 2700 km swath is not practical, more satellites will be required to achieve global coverage. For instance, with an 1800 km swath, 3-hour global coverage would require 12 satellites flying in groups of three as illustrated in Figures 5 and 6.

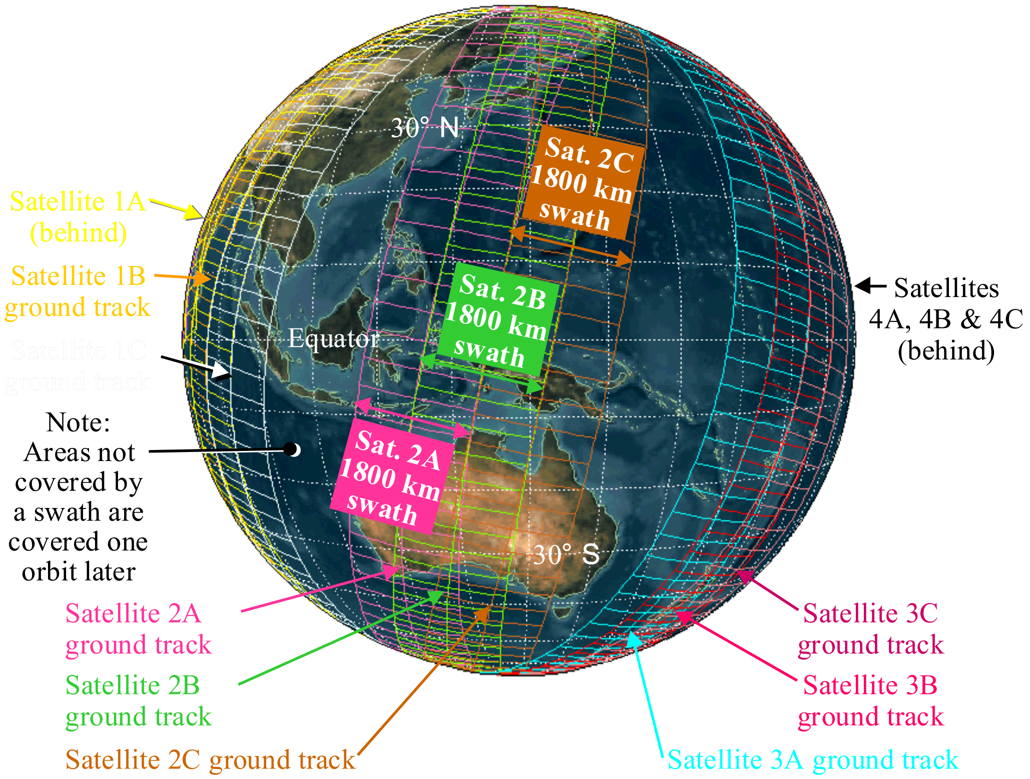


Figure 5 -- Sat-A sees Sat-B's Ground Track & Vice Versa
also, Sat-B sees Sat-C's Ground Track & Vice Versa

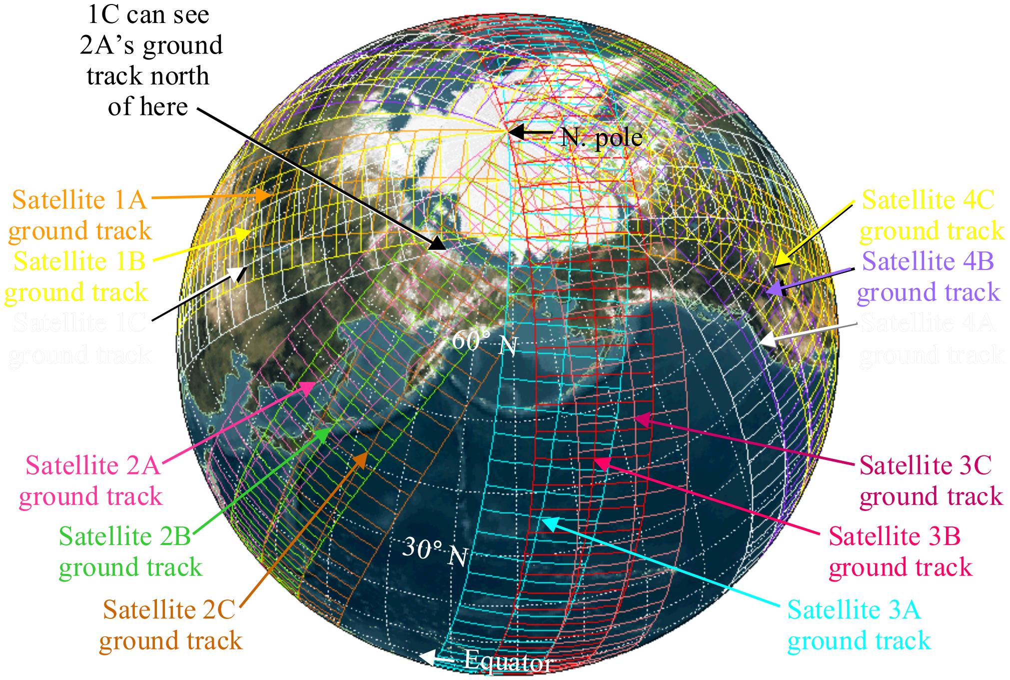


Figure 6 -- Redundant Coverage Near North Pole - 4 Groups of 3 Satellites

Figure 7 shows the number of satellites required for global 3-hour wind measurements as a function of swath width.

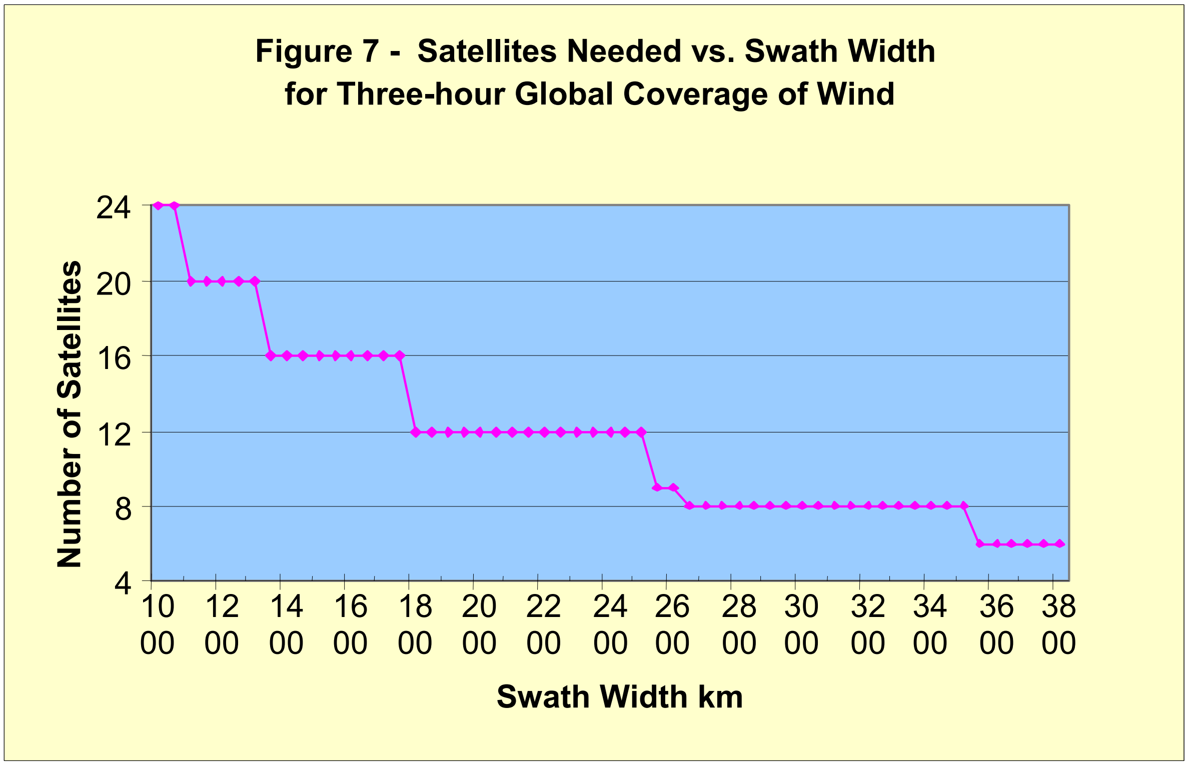


Figure 7 - Satellites Needed vs. Swath Width for Three-hour Global Coverage of Wind

GODDARD SPACE FLIGHT CENTER

Advanced Weather Prediction Technologies: NASA's Contribution to the Operational Agencies

Gap Analysis Appendix

May 31, 2002

Greenbelt, Md. 20771

1. Introduction

This appendix describes the gap analysis performed in support of the notional vision architecture described in the study report. The architecture was used as a starting point to perform these gap analyses.

The first portion of the appendix is the matrix of possible technologies that the study team determined to be candidates for the 2025 architecture. The matrix identifies the anticipated maturity levels, and the analyses were performed for key areas where significant gaps were flagged.

The second portion of the appendix contains the details of the gap analyses.

Notes:

This appendix is not meant to stand alone from the study report; very little explanation of the technologies' roles in the architecture is given. To get the complete picture this appendix must be read after the study report, with frequent references to the study report likely.

2. Summary of Potential Technology Gap Areas

The table in Attachment 1 lists the technologies identified as possible candidates to be used in the notional vision architecture described in the main body of the Weather Prediction Technology Investment Study report. It is structured to follow the top level of the architecture, namely the Observing Systems, expansions on that Observing System, the Model and Data Assimilation System (M/DAS), and the Knowledge Delivery system. The next level down in the chart is organized by function or subsystem, with the next levels being technologies within that function or subsystem.

This chart represents the best estimates of the team, but was not exhaustively worked for completeness of technologies listed nor their anticipated readiness levels. It was not meant to be as formal as TRLs but rather as a starting point for future, in-depth studies. The levels of anticipated readiness are described at the end and are listed by categories across the top. The engineers and technologists on the study team researched technologies in their area of expertise and extrapolated developing technologies to the 2025 timeframe, a difficult task given its highly subjective nature.

The column definitions are as follows:

- Performance Objective – what the technology needs to be able to provide in 2025; wherever possible a number was developed, but that was not always the case given the nature of advanced concepts.
- Capability Maturity – current status
- Challenges and Investments – given investment levels, how likely is it that this technology will be ready to meet the performance objectives in 2025? The higher the number, the less likely, with red being the color that identifies a gap.
- Technology columns – this identifies the technology nature of the gaps
- Science columns – this identifies whether a science gap is in the area of research or infrastructure

3. Observing System Gap Descriptions

As stated in the main body of the report, the success of a weather forecast depends heavily on how well the initial conditions are portrayed. The "goodness" of this portrayal is determined both by the accuracy and by the "representativeness" of the measurements as well as their timeliness. This section discusses the technology gaps in both the means to collect representative measurements as well as the means deliver them to their destination in a timely fashion.

Collection technologies include:

- Sensor technology needed to measure the various parameters
- On-board computing needs to process the data
- Guidance and navigation technologies needed to precisely determine the collection location
- Collection management and control to orchestrate the data collection

Delivery technologies primarily involve the communications and networking capabilities needed to get the collected data to the users in the required timelines.

3.1. Wind Remote Sensing - Anticipated Wind Lidar Technology Needs

The assessment of technology requirements for a wind lidar is complicated because there are two techniques proposed for making this measurement. The basic idea of both is to measure the Doppler shift of light scattered by molecules and/or particles carried by the wind. The direct detection Doppler lidar method uses a high spectral resolution optical filter (often a Fabry-Perot interferometer) to measure this shift using the atmospheric backscattered laser energy from either molecules or aerosols. The coherent or heterodyne Doppler lidar converts laser light backscattered from aerosols or clouds from optical to radio frequencies and uses RF spectral analysis techniques to measure this frequency shift. Current direct detection approaches use near UV wavelengths (λ~350 nm) for the molecular Doppler wind measurements while the heterodyne technique proposes to operate in the near IR (λ-2 microns). Both of these approaches have been demonstrated using ground based lidars and have been studied extensively for spaceborne operations.

In general the capability to measure winds of both lidar approaches is a function of the signal-to-noise ratio (SNR) of the signal detected from the atmosphere. In both cases the SNR will be a function of instrumental characteristics (e.g. laser energy and repetition rate; number of laser shots averaged; telescope collection area; detector quantum efficiency; optical throughput), spacecraft related characteristics (e.g. orbital height; nadir angle; pointing accuracy and control) and atmospheric effects (e.g. spatial distribution (horizontal and vertical) of the target particles (aerosols and/or molecules); wavelength dependent molecular and aerosol backscatter coefficient; two way atmospheric transmission; cloud distribution, height and optical properties). It becomes clear that any detailed analysis of the technology trade space will be highly dependent on the specifics of the implementation. To complicate this further, the details of how the SNR relates to the characteristics of desired wind product (accuracy, vertical and horizontal resolution) are different for coherent detection and direct detection Doppler lidars. Finally, the scaling of technologies for the individual approaches to larger sizes (with thereby improved capabilities) is not directly comparable. For example the heterodyne approach requires a telescope with diffraction limited performance while the direct detection approach can use a much lower quality telescope. On the other hand, heterodyne detection has high out of band noise rejection and so will be immune to solar background noise even in daylight while direct detection signals must be determined in the presence of background noise. In any case the measurement by either technique is extremely challenging.

Fortunately, recent engineering studies of both approaches have established reference baselines for both coherent and direct detection approaches, and for the purposes of evaluating technology needs and gap analysis we can use the results of those studies to extrapolate the needs of the future. Consider the following analysis:

The measurement requirements proposed for this study are as follows:

1. Global measurement of 2-D winds with precision of 1 m/sec.
2. Horizontal resolution 25 km x 25 km.
3. Vertical resolution 0.25 km
4. Temporal resolution 3 hours.

We define a "measurement" to be an altitude profile measured by the lidar viewing in a single direction. This means to get a 2-D wind determination in a single horizontal resolution element, but at all required altitudes, requires 2 "measurements." We calculate the number of "measurements" required per day as follows:

Area of a resolution element is 25 x 25 = 625 km^2.

Area of the Earth is $4\pi r^2$ where $r = 6.36 \times 10^3$ km. This equals 5.1×10^8 km^2.

Therefore, the number of horizontal resolutions elements on the earth is about 816,000.

Each of these elements must be visited twice every 3 hours to yield the required 2-D temporal resolution, so the total number of daily altitude profiles is

816,000 x 16 = 1.3×10^7 per day.

A recent engineering study of wind lidar capabilities presumed that a single lidar had the ability to make about eight line-of-sight wind measurements per minute. If that lidar system is baselined then it could make

8 x 24 x 60 = 11520 line of sight measurements per day or

4 x 24 x 60 = 5760 horizontal wind measurements per day

It would therefore require roughly 2200 such lidars operating continuously to meet the horizontal wind measurement requirements.

This factor of 2200 could be made up in various ways (eg. 22 platforms with lidars of 100 times more capability or say 100 platforms having lidars with 22 times greater capacity.)

However, in the engineering studies the lidar system only had the ability to resolve vertically at 1 km resolution and had a precision of 3 m/sec rather than 1 m/sec. Scaling up the performance of the lidar to meet the more stringent requirements assumed in this study implies that we need about another order of magnitude improvement in the lidar capability above the 2200 already discussed.

If we assume that 100 platforms could be utilized, then we need to achieve about a factor of 100 improvement in sensitivity of the lidar over the next 25 years. Improving lidar sensitivity can come about in several ways:

1. Increase in laser power (pulse energy X repetition rate).
2. Increase in the collection area of the receiver optical system.
3. Increase in the efficiency (quantum efficiency or throughput) of the detection system.

Detector quantum efficiency is already relatively high for both the near IR coherent and near UV direct detection systems. It might be possible to achieve some level of improvement, although it is difficult to imagine that this increase could be more than about a factor of 2.

Current telescope systems for spaceborne lidars have an aperture of about 1 meter. If this could be scaled up by a factor of 3 then we would have found about an order of magnitude in sensitivity, as the SNR scales as the area. (Note: Background noise would also increase for a direct detection lidar and pointing knowledge and control requirements would increase for a coherent system). A premium here is placed on increasing aperture without significantly increasing mass. In addition, single satellite lidar systems may be required to obtain multiple perspectives by slewing the FOV to different azimuth angles by rotating the telescope or an external scanning optic.

The remaining factor of 5-10 would require improvements in many areas. Among these are the following:

- Laser efficiency (conversion of electrical power from spacecraft into light energy).
- Solar power conversion and storage efficiency (for high power drain lasers on spacecraft).
- Pointing system performance and efficiency.
- Frequency conversion efficiency.

3.2. Microwave Remote Sensing Technology

Microwave measurements have become an important input into the weather forecast models of today. Many of these measurements will need to be performed on a daily or even hourly basis on a global scale. In particular this section will try to address some of the technologies, which will allow measurement of these parameters on a global scale from space.

There are several passive and active microwave measurements which are useful for weather forecasting. This section will attempt to detail the technology improvements necessary to meet the requirements of this future forecasting study. Measurements that are particularly well suited to microwave remote sensing include soil moisture (and other similar surface parameters), atmospheric temperature, and atmospheric moisture.

3.2.1. Soil Moisture

Currently the 6.9 GHz radiometer on AMSR is the only high-resolution passive microwave measurement of soil moisture available. This frequency is not the optimum for measuring soil moisture. A better choice would be the 1.4 GHz remote sensing band. This frequency is much more sensitive to soil moisture and can penetrate through more vegetation than higher frequencies. The difficulty with using 1.4 GHz to measure soil moisture is that the wavelength is 5 times longer than 6.9. This results in an antenna whose diameter is 5 times as large. The most recent proposals for measuring soil moisture from space would produce resolutions on the order of 40 km. These missions have antennas that are about 6m in diameter. To move to a measurement of 1 km would require increasing the antenna size to approximately 240 m. This seams like a tall order since that antenna would also have to spin at about 120 RPM (2 Hz). Given these assumptions, potential solutions can be imagined.

3.2.1.1. High Resolution Array Technology

The technique of Synthetic Thinned Array Radiometry (STAR) similar to what was proposed for HYDROSTAR or the European Space Agencies' Soil Moisture Ocean Salinity (SMOS) missions can be scaled to higher resolutions than are being proposed today. Within 10 years at current levels of funding a 10 km soil moisture measurement could be implemented. The configurations that have been proposed are probably limited to 10 km resolution. Above this resolution the system becomes so large that decorrelation becomes a problem for the available bandwidth. The alternative is something that has been called the Doppler radiometer. It is an interferometer made of three or more radiometers, which are phase locked to each other. These radiometers would fly in a formation separated by the maximum diameter of the antenna required to achieve the desired resolution. A great deal of further study is needed to determine if this configuration can actually achieve the sort of sensitivity that is useful for a soil moisture measurement. A modest increase in funds for the study of this concept and methods of formation flying could lead to a space demonstration within 10 to 15 years.

3.2.1.2. Active/Passive Combined Algorithm

An alternative approach is one that has been suggested by Ulaby and others. This involves a combination of an active and passive measurement of soil moisture. The advantage of this approach is that it potentially takes advantage of the strengths of each measurement. The passive measurement would provide the high accuracy low-resolution soil moisture to use as a reference for the less accurate but potentially very high-resolution active measurement. One could envision a measurement, which included a 10 km STAR radiometer imaging primarily on each side of the spacecraft and an unfocused Synthetic Aperture Radar (SAR) with 1 km spatial resolution. This set of measurements would then be combined in a statistical way to provide a measurement, which has 1 km resolution and high sensitivity to soil moisture in the presence of vegetation. It is again unclear if this method will ever be useful. It currently has not been demonstrated in any field experiment. If it can be accomplished, the development of a field instrument and funding for the development of the algorithm will be required to prove that it can work.

3.2.1.3. Large Real Aperture Approach

Perhaps the least elegant but simplest approach to the electrical design is the large single aperture. This aperture would have to grow to several hundred meters to make a 1 km measurement possible. This certainly seems to be unlikely to happen in the immediate short term. The technology that would make something like this possible is thin film inflatable antennas. This technology can produce very lightweight antenna structures with reasonable antenna characteristics.

3.2.2. Atmospheric Temperature and Moisture

Currently, atmospheric temperature and moisture are measured reasonably well, but at course horizontal and vertical resolutions. The frequencies of interest for these parameters typically range from approximately 19 – 85 GHz. These higher frequencies will require much smaller antenna systems than those for measurement of soil moisture. In all likelihood, there will not be a technology gap for these measurements.

3.3. On-Board Processing Technology

As our concept of a SensorWeb evolves, more of the processing needed to support its intelligence will need to be moved from the ground to space platforms. Some of this processing, such as data calibration and reduction, will be relatively simple (in terms of computing costs) while other functions, such as automated event recognition to enable SensorWeb reconfiguration, may be somewhat computationally expensive.

3.3.1. Anticipated On-Board Processing Needs

Although it is impossible to estimate the exact needs without a full concept development for how the SensorWeb will operate, we can at least make some broad generalizations. Data correction and reduction are currently done on workstation-class computers with capabilities in the low 100's of MFLOPS range. Other functions, such as initial forays into event detection, are done on higher-end workstation- or mainframe-class computers with capabilities in the high 100's of MFLOPS to GFLOPS range in a research mode. The anticipated needs in an operational mode are not yet known.

3.3.2. Anticipated On-Board Processing Capabilities

Currently, NASA's most powerful radiation-hardened computer matches the capability of a 80486 processor. However, non-hardened, higher end computer processors could be flown in space using various software and hardware techniques to overcome radiation effects. Other technologies being flown commercially are already putting high-end computing capabilities into space.

Boeing's recently launched Thuraya communications satellite has a digital signal processing power equivalent to 3000 Pentium III computers. Unfortunately, the power requirements of this system necessitate solar panels nearly 35 meters across for electrical power generation and a 7.4 square meter radiator to dissipate the heat, certainly beyond practical limits for a constellation of LEO platforms.

If a practical application of these computing capabilities is to come about, especially for a constellation of LEO platforms, a means of reducing the power requirements is needed. Studies by the Space Telescope Science Institute and the Jet Propulsion Laboratory (JPL) are examining the potential for high performance computing in space with lower power requirements. Experiments underway are running prototype systems capable of 30 MFLOPS per watt. It is hoped that computing capabilities of 300 – 1000 MFLOPS per watt will be available to support the Next Generation Space Telescope expected to fly as early as 2008.

If Moore's Law is applied to these numbers, we can estimate the expected on-board computing capability over time. Figure 3-2 shows this projection of on-board computing capabilities through 2025. The lower line (in red) shows the expected growth of computing capability per watt of power. Because systems of 20 such processors are envisioned, the computing capability of a 20-node system is shown by the upper line (in green). Thus, it seems reasonable that an on-board computing capability of 1000 GFLOPS – or just a little bit better than today's ground-based super computers – can be expected by 2025.

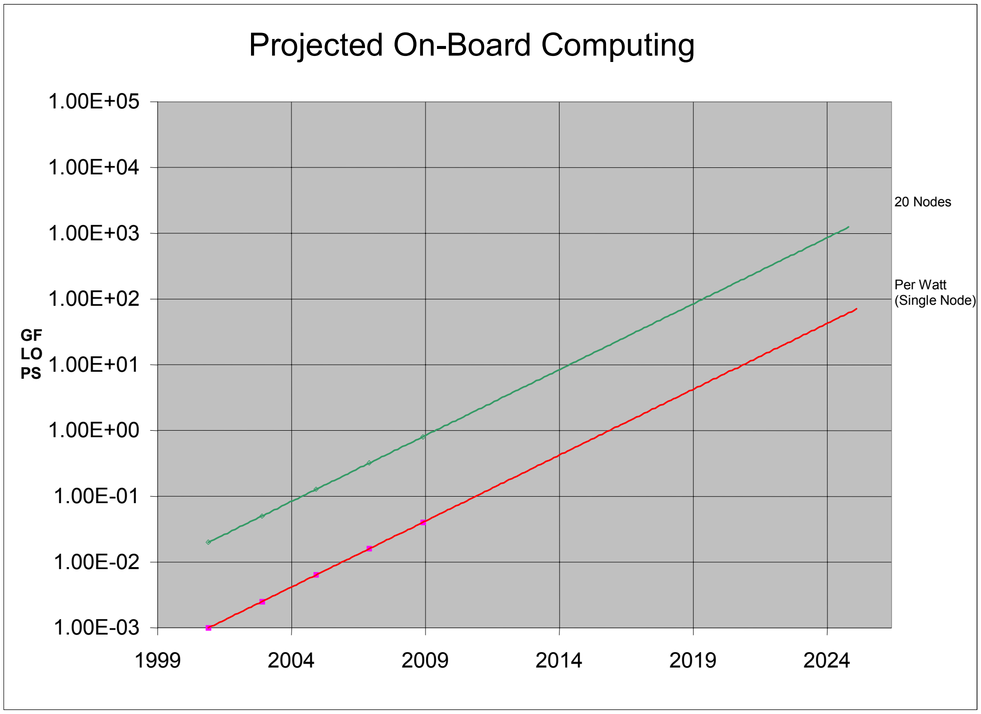


Figure 3-1 Projected On-Board Computing Capabilities

3.3.3. Gap Analysis

It is difficult to assess whether a technology gap exists based on the lack of a firm concept of operations. However, based on the initial concepts outlined in the study report (where only data calibration/reduction and initial QC are accomplished on board), it appears as if there will be no gap in on-board computing capabilities.

However, event detection and recognition algorithms (whose computational complexity is not yet known) might tax the expected capabilities. Further studies into potential science applications and their computational costs are needed to fully understand these needs.

3.3.3.1. Technology Shortfalls

As Boeing's Thuraya communications satellite demonstrates, high performance computing in space is possible, even with today's technology. However, it comes at great costs in terms of weight, power, and thermal considerations. If these computing capabilities are to become a possibility for a LEO constellation, the computing capability per watt is a critical factor. Although current prototype efforts are promising, it remains to be seen if these technologies will be scalable for future needs.

3.3.3.2. Trade Areas

The key trade area to be investigated relates to functions to be processed on board vs. on the ground. These trades must be weighed against how much communications bandwidth is expected to be available. If greater communications capacity can bring data to the ground in near real time, then processing can be done on the ground, reducing the need for on-board processing requirements. However, if communications bandwidth is limited, or if the science of event detection and recognition dictates an immediate response, then high performance computing in space must be considered.

If the vision of low-power computing does not come to fruition, then trades must be made between computing capability and power/thermal considerations.

3.3.3.3. Future Technology Needs

The future weather architecture outlined in the study report will certainly require increased computational resources on-board the space platforms. With the power requirements of the instruments, especially the active sensors, power considerations will be a limiting factor. Thus, increased computing capability with reduced power costs is going to be crucial. If these power reductions cannot be realized, increases in power generation capabilities (e.g., more efficient solar panels) and better thermal management will be a must.

3.3.3.4. Recommendations

As in the ground-based computing portion of this study, we must keep an eye on the computer industry. Furthermore, we must maintain an open dialog with the computer research community (such as the REE project) so they remain aware of our future computational needs.

ESTO should also support research into developing more efficient computational systems and algorithms to make better use of the available computational resources and support research into smarter analysis and forecast algorithms.

ESTO should support a follow-on effort to flesh out a concept of operations in order to more fully identify what processing requirements are needed in space. In the current version of the notional architecture, only minimal data processing is accomplished on board the spacecraft conducting remote sensing measurements. It is conceivable that some portion of the data processing should be done on the spacecraft that would benefit either the efficiency of the system or quality of the collected data and forecast products.

3.4. Guidance, Navigation, and Control (GN&C)

In addition to the sensor and computing technologies needed, improvements in spacecraft GN&C will also be needed. As higher resolution measurements are made, it is increasingly more important to have a better understanding of exactly where the measurements are being made to ensure that they are representative of the true state of the atmosphere. With the larger number of space platforms envisioned, the cost and complexity of managing the operation and control of the constellation could become prohibitive. The next sections discuss the technologies needed to address these concerns.

3.4.1. Global Positioning System (GPS) Navigation

Today's GPS can provide all the LEO weather satellites with latitude, longitude and altitude to an accuracy that is better than the minimum needed to locate the data collected.

This is not true for the GEO satellites. These satellites are above the GPS constellation and can only receive GPS signals from those GPS satellites that are close to setting (or rising) behind the earth. There is every theoretical reason to believe that satellites in GEO orbit will be able to use these signals to achieve adequate positional accuracy, but so far, no mission has demonstrated this. The GEO weather satellites will need this capability because accurate satellite location is needed for accurate, automatic location of images and other data produced by these satellites.

No current efforts are planned to explore this technology. Figure 3-2 shows the relative GEO GPS levels of performance (or technology readiness) and the approximate time each level could be obtained. Trades on performance (and benefits) of obtaining this capability compared to their costs must be conducted to determine if the appropriate performance level can be reached.

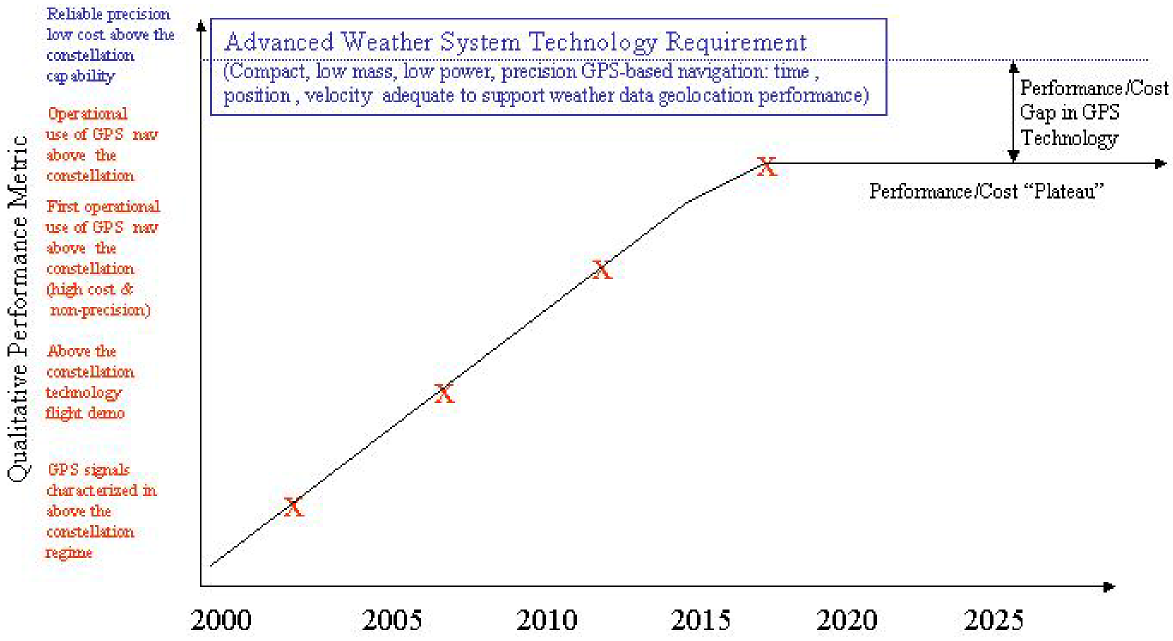


Figure 3-2 GPS Navigation Gap Analysis

3.4.2. Drag-free Control

In typical ESE missions significant ground resources are applied to the orbit determination and orbit prediction tasks. It is well known that the most uncertain part of solving the Low Earth orbiting flight dynamics problem is the prediction of atmospheric drag levels. The drag force from the Earth's atmosphere not only tends to decay spacecraft orbits, but also can vary significantly from day to day. This uncertainty drives the need for increased spacecraft tracking and detailed orbit modeling, determination and analysis to measure the vehicle positions as well as control them with periodic propulsive maneuvers.

The effective elimination of drag from the spacecraft flight dynamics problem reduces any orbital decay to purely gravitational perturbations which are known quite well and which can be compensated for with appropriate analysis. Through the use of an integrated accelerometer package on a spacecraft, most likely consisting of a floating proof mass in an internal chamber with electrostatic (capacitive) sensing and actuation, a high specific impulse (Isp) thruster, and a low-cost processor with appropriate filtering/control algorithms, a closed loop drag free control system can be synthesized. Such a drag free system will:

- Eliminate the effect of drag on each spacecraft to prevent decay of the orbit (using virtually insignificant continuous and non-interfering thrust).
- Continuously maintain the constellation elements within their boxes to avoid undesirable interactions
- Avert the need to shut down the mission every 1-4 weeks to perform a delta-vee orbit correction.
- Maintain precise knowledge of the orbital position of the vehicles continuously without sensitivity to upsets, bit-flips, etc. and without the need for expensive sensors. This will enable vastly improved geolocation performance to enable us to meet specs for such tasks as wind speed measurement.
- Avoid the need for complex algorithms for collision avoidance and large scale constellation maintenance.

Figure 3-3 shows the expected performance capabilities and the requirements of the system outlined in the report. Two performance metrics – acceleration cancellation levels and the number of spacecraft for system level application – are presented over time. Acceleration cancellation technologies already planned are anticipated to meet the requirements of the system envisioned. The second metric, however, is dependent upon two technologies, (1) low-cost, moderate performance, drag free (floating proof-mass) sensor development and (2) algorithms which enable us to use this technology at the constellation level, rather than at the single spacecraft level. Currently, no systems are planned to implement these technologies in constellations with the number of spacecraft envisioned. Significant development in this area is required.

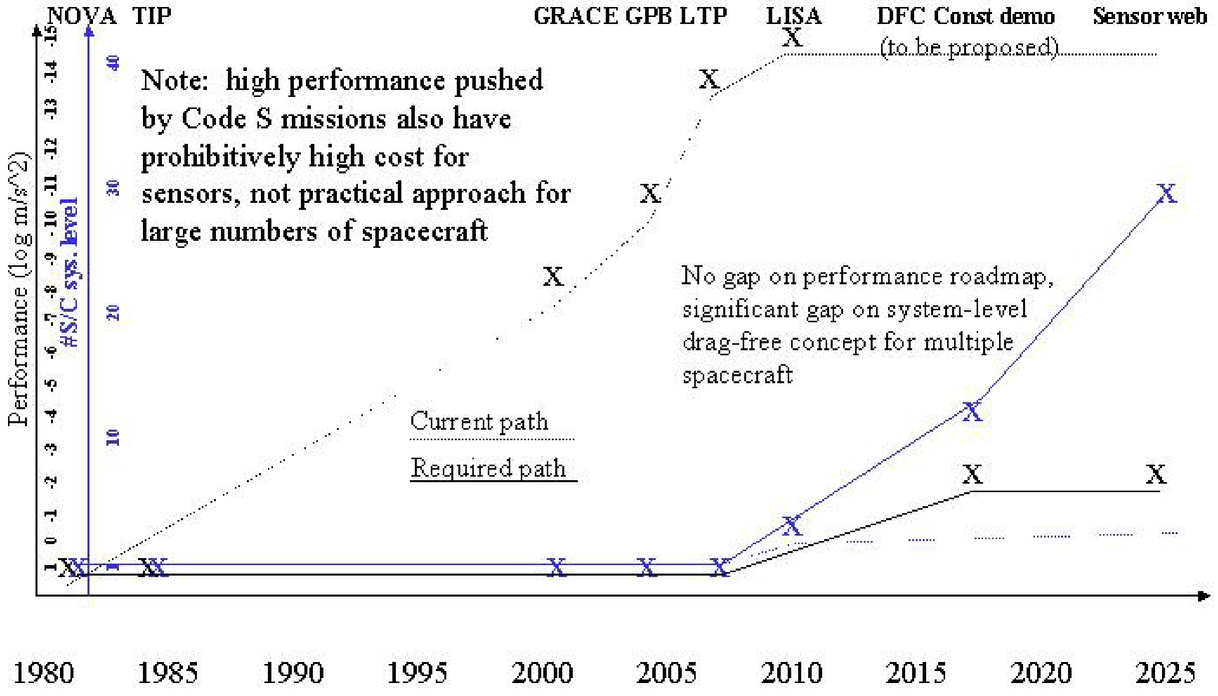


Figure 3-3 Drag Free Control Gap Analysis

3.5. SensorWeb Management/Control (SWM/C) Technology Gap

As envisioned, the SensorWeb will require rapid, nearly seamless communication between assets located in space, in the air, on the ground, and at sea. An overarching “intelligence”, referred to here as the SensorWeb Management/Control (SWM/C), would manage the assets to make regularly scheduled data collections and to optimize the scientific targets of opportunity. The SWM/C would provide coordination between command/control for widely disparate collection platforms and complex dynamic planning and scheduling. Because of the unprecedented configuration of the proposed SensorWeb and the complexity of scheduling the assets technology gaps are evident.

3.5.1. SensorWeb Command/Control

The SensorWeb Command/Control will require all of the standard operating components seen in today’s satellite systems. It will also require the services provided by such components to be extended across all observing systems assets (e.g. aircraft, ships, ground-based sensors, etc.). Such components include:

- Data architecture to identify the major components of the overall observing system
- System management architectures that provide for the organizational and management of the operations environment of the assets
- Control interfaces that provide a mechanism to operate and manage the assets
- Decision support components to operate the assets and process commands

3.5.1.1. SensorWeb Dynamic Planning and Scheduling

Planning and scheduling of the SensorWeb assets addresses the problem of formulating a sequence of commands that will result in achieving a desired scientific goal. A possible scenario for SensorWeb operations would be three "operating modes":

- *Normal operating mode* would schedule regular collections of operations that would satisfy the basic requirements of the data assimilation system. This would include making measurements of temperature, moisture, wind, etc. at the appropriate temporal, spatial, and spectral resolutions. This mode would also address the scheduling of data points that must be re-sampled due to initial flagging by the meteorological quality control. The main driver of this mode of operations would be the data assimilation system.
- *Opportunistic science mode* would be used to capture events of specified scientific interest. The observing system elements and/or the data assimilation system would alert the planning/scheduling algorithm to perform intensive data collections focused on specific locations. For example, if a satellite detects conditions favorable for severe weather (perhaps by using on-board event-detection algorithms) the planning/scheduling component would interact with the data assimilation system to predict the location of the event over the next several hours and schedule high-resolution data collections accordingly. The entire lifecycle of the severe weather outbreak could then be captured.
- *Field experiment mode* would be used to manually select regions for intensive observations. This mode would be particularly useful for research studies that require higher resolution data over a specified location over a period of time.

It is unlikely the SensorWeb would ever operate in a single mode. Rather, to maximize the scientific benefit, an optimal combination of the three modes is necessary. Defining the optimal combination of such a dynamic system is a grand challenge of building the SensorWeb.

3.5.2. Anticipated Technology Capabilities

NASA and other government agencies are now formulating the roadmap to develop intelligent Distributed Spacecraft Systems (DSS). In recent studies, DSS is defined as a spatially distributed intelligent network of multiple space assets, collaborating as a collective unit, and exhibiting a common system-wide capability to accomplish shared objectives. This work is significant to the current study because there is considerable overlap in the command/control system requirements for DSS missions and for SensorWebs. Proposed Earth Science DSS missions that may require enhanced command/control capabilities over the next 10-15 years include:

- Global Precipitation Mission (currently scheduled for launch in 2007)
- Leonardo
- Topography and Surface Deformation
- GPS Atmospheric Sounding

There are now investigative activities underway to prepare for DSS missions. For example, two recent investments by the NASA Office of Aerospace Technology are funding studies into the development of discrete event controllers for autonomous, distributed spacecraft command/control, and for autonomous command/control for formation flying. There is also a significant amount of work underway in Space Sciences, with much of the research for distributed spacecraft problem solving being performed by the Jet Propulsion Laboratory. At Goddard Space Flight Center work is underway to develop the so-called "goal-oriented commanding" that is designed to perform high-level tasking of a constellation of satellites with a minimal amount of human intervention.

Although the DSS studies will likely provide some benefit in designing a SensorWeb command/control system, they are limited because they address only on-orbit assets. There are no studies currently investigating the design of a command/control system that manages a diverse suite of assets that would be needed by the proposed SensorWeb.

Figure 3-4 displays a likely technology capability roadmap for command/control capabilities of distributed spacecraft systems. The data contained in the figure is based upon an Earth Science Enterprise planning workshop conducted in 2001. The analysis suggests significant strides will be made in autonomous spacecraft control and scheduling over the next five years, with demonstrated mission capability likely by the end of the decade.

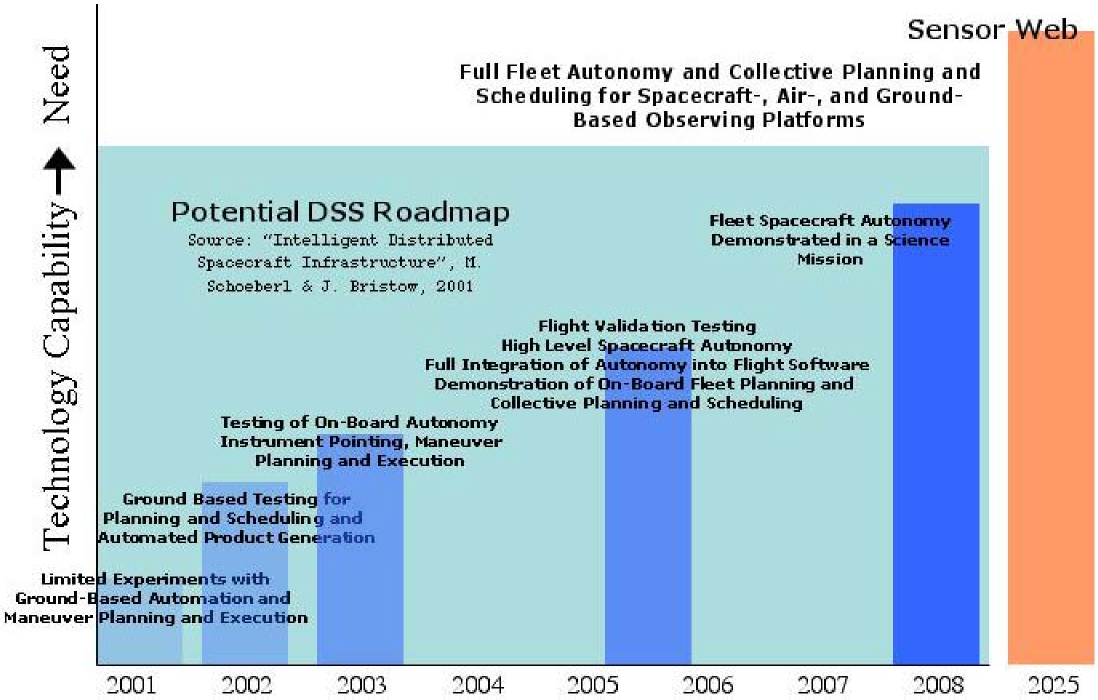


Figure 3-4 Anticipated technology roadmap for addressing command/control and planning/scheduling of DSS missions.

3.5.3. Gap Analysis

Although not entirely quantifiable, a technology gap appears to exist for the successful development of the SWM/C components of the SensorWeb. The gap is related to the diversity of the assets that must be managed, and the time constraints placed on the highly complex scheduling necessary.

3.5.3.1. Technology Shortfalls / Future Needs

Based upon the data contained in Figure 3-4, command/control and planning/scheduling algorithms for satellite constellations will likely be available by the end of the decade. Studies associated with the requirements analysis and the design of such algorithms will benefit the design of similar SensorWeb components. However there does not appear to be current or planned studies involved with linking services for satellite-based observing systems with air- and ground-based observation networks.

For the SWM/C, “asset awareness” and the complexity of the optimization problem appear to be the most significant gaps. In order to optimize data collection to maximize scientific return, the scheduling algorithm would be required to identify accurately the current and future locations of all assets, as well as deployment times and overall availability of rawinsondes, unattended aircraft, drifting buoys, ships, and current and anticipated states of many other resources. Therefore, the scheduling algorithm must have detailed, up-to-the-minute knowledge of perhaps tens of thousands of assets, and must perform scheduling decisions within a matter of seconds. Such decisions must be based upon weighing requests made under the three operating modes of the SensorWeb and rapidly formulating the “best” decision. Although similar algorithms exist today (goal-oriented commanding of spacecraft and even computer chess games that anticipate and score future moves of the chess pieces are relevant examples), the requirements of the SensorWeb require a significant step forward in both hardware and software technology.

3.5.3.2. Recommendations

Future studies should serve to bridge the apparent gap between relevant research on DSS and the requirements of the proposed SensorWeb, and should attempt to quantify at a low level the requirements for planning and scheduling within the SWM/C. Simulation of the SensorWeb environment would be one approach to understand the magnitude of the technology gap and to assist researchers in addressing the challenges presented by a complex observing system.

3.6. Communications Technology -- GEO Satellites

The GEO satellites would downlink their data directly to ground stations and the commands would be uplinked from the same stations. On the ground, commercial communications links would carry the data and the commands between the ground stations and the weather system command and control center(s). There are no technology concerns with any of these communications links.

3.7. Communications Technology -- LEO Satellites

3.7.1. Requirements that drive the communications System

A key requirement highlighted in the main report is a data latency of no greater than 15 minutes. This requirement drives the system architecture to use a space-based communications system similar to the present Tracking and Data Relay Satellite System (TDRSS). The capability of the satellites that compose the present TDRSS could do this mission.

Since the LEO weather satellites could have coverage at locations around the earth at a given time, a minimum space-based communication system requires three full-capability communications satellites located approximately 120 degrees apart. The ground terminals for each of these satellites must have adequate communications with the users of the weather data and the Command and Control Center. (Today's TDRSS has satellites located to the East and West of the United States with adequate capacity, but there is only one satellite with partial capability over the Indian Ocean. It is expected that by 2025, global TDRS capability will be available.)

3.7.2. Description of the Tracking and Data Relay Satellites (TDRSs)

The latest generation of TDRSs has three modes of receiving data from LEO satellites. They each include a Multiple Access (MA) service, a K-band Single Access (KSA) service and an S-band Single Access (SSA) service. The MA is an S-band phased array that is able to receive up to 3 Mbps from five LEO satellites simultaneously. Each TDRS has two Single Access Antennas and each antenna includes a KSA service that can receive up to 300 Mbps and a SSA service that can receive up to 6 Mbps. The KSA service can receive data from one LEO satellite and simultaneously can receive SSA data from the same satellite or from a second, nearby LEO satellite. The dead time between the end of one SA contact and the start of the next is typically 1.5 minutes for the newest TDRSs, but was only about 30 seconds for the original TDRSs. For this paper, it is assumed that future TDRSs will have dead times of less than 30 seconds.

3.7.3. Low Data Rate Spacecraft/Sensors (Multi-Access)

Table 3-1 gives the estimated data rates for the satellites that would make up the LEO constellation of weather satellites. It is seen that only the 8 LIDAR satellites have data rates suitable for the MA service. The three TDRSs would each handle 2 or 3 LIDAR satellites simultaneously. The data rate would be about 300 kbps or about 25% larger than the 241 kbps shown in the table. This is to allow for the dead time while the LEO satellite switch between one TDRS and another and to allow for the occasional data rate surges that are higher than average.

Table 3-1 Estimated LEO Weather Satellite Data Rates

		These instruments may be grouped on one satellite				LIDAR Satellite	RADAR Satellite
		EO Imager	MW Sounder	MW Imager	IR Sounder		
Number of Instruments or Number of Satellites		13	13	13	13	8	13
Time to Achieve Global Coverage hours		1	1	1	1	3	1
Swath Width (km)		2700	2700	2700	2700	2700	2700
Spatial Resolution (km) (nominal)		1	25	25	25	25	25
Spatial Resolution (km) (surge)		0.5	5	5	10	10	1
Period (min)		98	98	98	98	98	98
Data size per observation (bits)		16	10	10	16	10	16
Metadata / observation (bits)		1	40	40	512	512	40
NOMIN	# Channels or Vertical Samples	12	7	10	200	320	20
	Observations / swath	2700	108	108	108	108	108
	Swaths / sec	6.808	0.272	0.272	0.272	0.272	0.272
	Nominal Data Rate Mbps	3.383	0.003	0.004	0.104	0.104	0.010
SUR	# Channels or Vertical Samples	12	7	10	200	320	80
	Observations / swath	5400	540	540	270	270	2700
	Swaths / sec	13.616	1.362	1.362	0.681	0.681	6.808
	Peak Data Rate Mbps	13.533	0.077	0.098	0.651	0.651	23.139
	% Time In Surge Mode	15%	15%	15%	15%	25%	50%
	Average Surge Data Rate Mbps	2.030	0.012	0.015	0.098	0.163	11.570
SURGE+NOMI	Data Rate per Instrument Mbps	4.906	0.014	0.018	0.186	0.241	11.575
	Data Rate per Satellite Mbps *	3.20				0.15	7.23
	TDRS link required	Single Access				Multiple Access	Single Access
	Total Data Rate By Satellite Class Mbps *	42				1.2	94
	Total Data From All LEO Satellites Mbps *	137					
* Note: Includes 2X Compression plus 25% overhead for forward error correction code and formatting							

3.7.4. High Data Rate Spacecraft/Sensors (KSA)

The data rates from the Imager/Sounder satellites and the RADAR satellites are high enough that they must use the KSA service. At any one time, there could be 8 or 9 satellites in view of each TDRS and they each would have to cycle through all these 8 or 9 satellites every 15 minutes. Given that 0.5 minutes is lost between LEO satellite contacts, one SA antenna would lose 4.5 minutes every 15 minutes while slewing among 9 satellites. This leaves only 10.5 minutes for data collection from 9 satellites giving each LEO satellite only 1.1 minutes to transmit its previous 15 minutes of data. To accommodate some higher than average surges of data, we use a contact time of 1 minute or a data speed up of 15X. This would increase the data rate from the Imager/Sounder satellites to 48 Mbps and from the RADAR satellites to 109 Mbps. These data rates are well within today's communication capability.

3.7.5. Sensitivity to the Selected Data Parameters

It is believed that the parameters of Table 3-1 can be achieved with appropriate technology investment.

The number of each type of satellite and its swath width is driven by the time to achieve global coverage requirement. If the 2700 km swaths assumed in Table 3-1 cannot be achieved, more satellites will be required. The total data to be communicated will not change significantly, but the additional satellites will cause more wasted dead time for the KSA Service.

The data rates shown in Table 3-1 assume that the instruments continuously sample. This is appropriate when the satellites are in tropical latitudes, but the earth's polar regions would be oversampled. It is expected that the sampling in the polar regions would be reduced, but this was not factored into the data rate calculations because there can be 15 minute periods when most of the satellites will be in tropical latitudes.

Lossless compression of a factor of 2 was used in Table 3-1. The compression that will be achievable may be more or less than this number. However, it is seen that even if no compression were possible, the satellite data rates would still be within a TDRS's capability.

Table 3-1 assumed that the imager and sounder instruments were flown on the same satellite. If they were flown on separate satellites, the EO Imager would still require KSA service, but the others could be handled on MA. If flown on separate satellites, they would then likely fly in a close formation to achieve nearly simultaneous area coverage. In this case, while the EO Imager data was being transmitted on KSA, the data from the other three satellites could be sequentially transmitted using the SSA capability.

4. Modeling and Data Assimilation Gaps

4.1. Computing Technology

Even today, numerical weather prediction is one of the most computationally taxing functions performed. Indeed, many of the current limitations in weather prediction are imposed not by uncertainties in the science, but rather by the inability to perform the necessary calculations in time. Although advances in computer technology will lead to faster computers, the needs of the future weather forecasting architecture will also increase tremendously.

4.1.1. Anticipated Computing Technology Needs

Clearly, the future computing needs of a weather forecasting system will increase as the architecture discussed in the main part of this study report comes to fruition. A quantitative assessment identifies several aspects that will greatly impact the ultimate computational complexity of this future system. Three key elements driving the increase in computational needs are related to:

- Increases in the resolution of the analysis and model functions
- Increased complexity of algorithms contained within these functions
- Increased numbers of observational data collections providing an input to the models

4.1.1.1. Resolution Increases

As the model resolution increases by some factor in the horizontal, the number of calculations required increases by the square of that factor. A quick, qualitative assessment concludes that one would expect huge increases in processing needs as the analysis and model resolutions go from the current 1x1° globally (about 111 km resolution) to a resolution of 25 km or better. When the increases in the number of vertical layers represented by the models is also considered, these increases become even greater.

4.1.1.2. Algorithm Complexity Increases

Analysis Complexity

Current analysis schemes used by various agencies range from 3-D variational (3Dvar) analyses, spectral/statistical interpolation, or other variations of optimum interpolation schemes. The near-term future of analyses will likely progress to 4-D variational analyses, where observations are brought into the model by analyzing the data with respect to time in addition to the 3 spatial dimensions. Of course, the additional analysis dimension adds considerably to computing complexity. The computational complexity of these schemes tends to scale as the square of the number of observational data points being brought into the analysis. Although the future will likely cause an increase in observational data of two to three orders of magnitude, these increases can be offset conducting analyses at more frequent intervals (decreasing the numbers of observations ingested at each step).

Further into the future, analysis schemes will likely include Kalman filtering as a principle component. Although well-understood, Kalman filtering techniques for global analyses are computationally very expensive – prohibitively so at present. The complexity of these schemes scales as the number of observational points times the square of the number of model gridpoints. Again, the number of observations processed for a given analysis can be decreased by more frequent analyses, however, the number of model gridpoints cannot be easily reduced.

Model Complexity

The complexity of atmospheric models is subject to great variability. Because the resolution of today's models is such that certain smaller-scale features cannot be accurately modeled, they are parameterized. In many cases, the complexity of these parameterizations is greater than the explicit modeling. As the resolution of models increases, however, these features could be modeled explicitly, perhaps bringing a general decrease in the actual complexity of the NWP algorithms. However, with increasing resolution comes a greater number of model gridpoints. To what extent the competing effects will weigh is uncertain, although it is almost certain that the general trend will be towards overall increases in complexity. A linear increase with the number of modeling points does not seem unreasonable.

4.1.1.3. Observational Data Increases

With today's satellite remote sensing, many more observational data points are available than there were even 10 years ago. Typical estimates for the number of observations used by today's models center around approximate 10^6 observations per day. Even by the most conservative estimates, this number will increase by two orders of magnitude by 2025. Even with larger numbers, however, it seems reasonable to expect that no more than 10^8 observations per day will be used by the models once redundant and/or low-quality data are filtered out. Still, with analysis complexity scaling as the square of the number of observations, this will result in a huge increase in computational costs.

So far, much of the discussion has centered on qualitative assessments of the increases in computational complexity. In order to obtain a quantitative assessment, it is necessary to use estimates of future model specifications (resolution, numbers of observations, etc.) to calculate model complexity. These numbers can then be used to determine how much computing capability will be needed.

Many of the calculations used to determine the computational costs of future systems were based on Lyster, July 2000. In this paper, Lyster presents a methodology for calculating the complexity of various analysis and modeling schemes based on specifications such as number of analysis points, number of observations, time step, etc. The results of the Lyster calculations are a total number of floating point operations needed to perform the stated function. Based on assumptions of the amount of time needed to complete a given task and an estimated computational efficiency, an estimate of the required sustained computing power is obtained.

Complexity Calculator

The calculations described by Lyster have been included in a simple spreadsheet. By changing forecast system specifications (such as analysis/model resolution, number of observations, runtime, etc., an estimate of computing resources in GigaFLOPS (10^9 Floating Point Operations per second) is returned for various computational algorithms.

Input Variables

The key input variables for the calculations (and their initial values) are shown below:

- Horizontal resolution (25 km)
- Number of analysis levels (100)
- Number of model levels (100)
- Analysis interval (1 hour)
- Number of upper air analysis/prognostic variables (4)
- Number of surface analysis/prognostic variables (1)
- Number of observations per day (10^8)
- Analysis run time (10 minutes)
- Quality Control run time (10 minutes)
- 24 hour forecast run time (5 minutes)
- Targeted observation run time (10 minutes)

Based on these inputs, various portions of the Modeling and Data Assimilation System will require anywhere from 10^7 to 10^{13} GFLOPS of computational resources. In other words, the range of computational resources needed is 10^{16} to 10^{21} Floating Point Operations per Second. For the curious, the range can also be stated as 10 PetaFLOPS to 1 ZettaFLOPS.

4.1.2. Anticipated Computing Technology Capabilities

At first glance, the numbers discussed in the previous section appear so high as to be impossibly ludicrous. However, with the expected growth in computing capabilities, the lower end of this spectrum actually falls within the domain of possibility.

When Gordon Moore first observed the growth in transistor density on computer chips, he found that it doubled roughly every 18 months. Although he was not necessarily referring to computing speed, transistor density typically relates linearly to it. Thus, the assessment of a doubling of computing speed every 18 months is now widely known as Moore's Law.

There is some concern that today's conventional computing systems (e.g., silicon or CMOS-type chips) will reach a size barrier in anywhere from 15 to 20 years. However, if one extrapolates computing speed back in time before solid state computing, it becomes apparent that the computing speeds of the earlier tube computers is consistent with Moore's Law. Thus, it is not unreasonable (at least for this study) to assume that some future technology (e.g., optical or quantum computing) is likely to pick up where silicon leaves off. Thus, the future computing capabilities expected in 2025 (over the course of normal evolution) are based on application of Moore's Law to some current computing capability.

Today's state of the practice systems boast speeds on the order of hundreds of GFLOPS. Although systems capable of sustained speeds in the TFLOPS range are in use, they are still considered state of the art and not readily available for operational centers. For this study, a current capability of 500 GFLOPS was used as a baseline for application of Moore's Law. Projecting forward, a sustained computing speed in the range of 10^7 – 10^8 GFLOPS was obtained.

4.1.3. Gap Analysis

With expected computing capabilities and a tool to estimate computing resource needs, we can now examine the technology gaps.

4.1.3.1. Technology Shortfalls

Figure 3-1 illustrates the expected growth of computing capabilities (red line) over time against the estimated computational needs of the future weather forecast system with specifications as stated in Section 4.1.1.

As discussed in the main body of the study report , the MDAS consists of several functions. These include quality control of the input data, analysis of the observed data onto a regular grid, the global forecast model, and the targeted observation selection. The spreadsheet tool estimates computation complexity of each of these functions. For the most part, the Kalman filtering analysis function is the most computationally expensive part of the future system. To explore options to reduce the system's computational requirements, several additional analysis schemes were also explored. The four threshold lines (in blue) indicate the computational needs (in GFLOPS) for the full MDAS obtained for each of these analysis options.

The topmost line indicates the processing needs for an analysis and forecast system using a full Kalman filtering analysis scheme. Because the computational needs of this algorithm scale as the square of the analysis gridpoints times the number of observations, this turns out to be computationally very expensive. This analysis places the computing needs at more than five orders of magnitude greater than the expected capabilities.

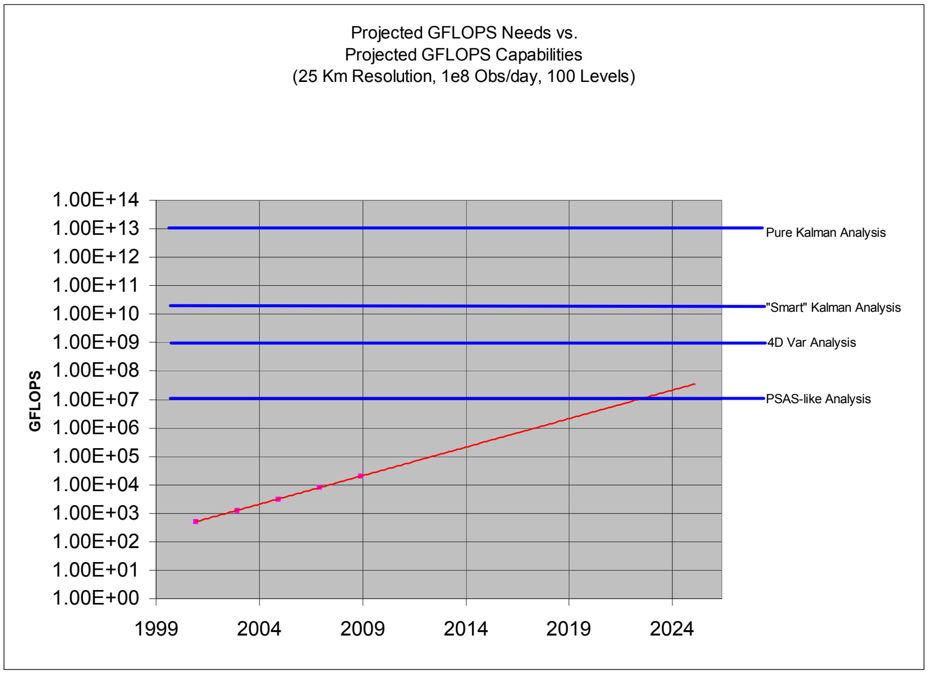


Figure 4-1 Projected Computing Capability Gap

4.1.3.2. Trade Areas

Obviously, a full Kalman filtering is most likely well beyond the capabilities of projected future technologies. However, discussions with analysis experts from DAO have led to other options. While a full Kalman filter would provide the best analysis, a partial or "smart" Kalman filter might yield a solution that is "good enough." Such an analysis might scale as the number of gridpoints raised to the power of 1.6 or 1.7 (vice 2.0). Such a scheme would realize great savings in computational costs. The value of this factor would be related to how well the Kalman analysis performs. Thus, the next lower line (labeled "Smart" Kalman Analysis) marks the threshold for a factor of 1.7. Although nearly 2½ orders of magnitude higher than expected capabilities, it is still significantly lower than a full Kalman analysis. Using a factor of 1.6, the requirements drop to $4x10^9$ GFLOPS, less than 1½ orders of magnitude above the projected capabilities.

Another scheme that has a lower computational costs is the 4 dimensional variational (4D Var) analysis. As of now, 4D Var analysis schemes are seeing use in either small scale analysis of all observational data types or for global analysis of limited data types. As the implementation of this scheme improves and resources allow for the somewhat high computing costs, 4D Var will most likely see use as a full global analysis scheme in the near future (at current spatial resolutions). The estimates of the 4D Var computational costs were made for the same specifications as the Kalman filtering scheme and are indicated by the third blue line. The costs, while lower than all the previous analysis schemes, are still 1½ orders of magnitude above the expected capabilities. Additionally, the computing cost of the analysis portion of the MDAS is now within the realm of the other functions within the system, especially the global forecast model.

The final analysis scheme examined to lower computational costs is a 3 dimensional Variational (3D Var) analysis. A simpler version of the 4D Var, the complexity of this scheme is on par with today's analysis tools used operationally. When the complexity of the 3D Var is estimated using the same specifications (lower blue line on Figure 4-1), computing needs are found to be on the order of 10^7 GFLOPS – well within the projected capabilities of future computing systems. Furthermore, the computing costs of the analysis are found to be of the same order of magnitude as those of the global forecast model.

4.1.3.3. Future Technology Needs

From this examination, we see a large gap between expected computing needs and resources. There are two ways to close the gap – raise the available computing speed or lower the computational requirements. Raising computing speed can only be done through the technology advances of the computer industry. Reducing the computational requirements, however, is within the realm of the earth science community by way of computation that is more efficient or smarter data analysis algorithms.

4.1.3.4. Recommendations

ESTO should keep an eye on the computer industry and examine technologies that could lead to the capabilities needed, such as reconfigurable computers. Furthermore, NASA's earth science community should maintain an open dialog with the computer research community so they remain aware of the future computational needs

ESTO should also support research into developing more efficient computational systems and algorithms to make better use of the available computational resources. As was noted earlier, current weather codes have an efficiency of perhaps 10%. Significantly increasing this efficiency would be a good start at closing the gap. In addition to efficiency, ESTO should support research into smarter analysis and forecast algorithms such as the "smart" Kalman filter. Although related to the science of numerical weather prediction, these are actually computing technology advances that need to come about in order to close the anticipated gaps.

4.2. Meteorological Science

In addition to the technologies discussed above, significant development will have to occur in key areas of the numerical modeling arena in order to support the concepts discussed in the study report. These areas include selection of targeted observations, continuous assimilation and model self-assessment.

5. Areas for Further Study

As this study was conducted, certain assumptions were made in order to complete the study in a limited timeframe. The following sections discuss several areas in which additional study would be beneficial.

5.1. Study Detail Refinements

5.1.1. On Board Processing Trades

In the current version of the notional architecture described in the study report, only minimal data processing is accomplished on board the spacecraft conducting remote sensing measurements. Instead, the calibrated, earth located data are transmitted to the ground to be reduced for ingest into the analysis models.

It is conceivable that some portion of the data processing should be done on the spacecraft that would benefit either the efficiency of the system or quality of the forecast product. For instance, data that are reduced from the raw (sensor-based) measurements to the required parameters and spatial resolution needed for the model might be far less voluminous, decreasing the bandwidth requirements needed for downlink. As another example, data that are reduced to their desired environmental parameters might be better suited for the automated event detection needed for a rapid reconfiguring of the SensorWeb.

However, moving processing from the ground to the space platform entails its own difficulties. Some reduction schemes require supplemental data that will have to be uplinked to the spacecraft -- will the potential bandwidth savings and quality improvements be worth the additional uplink? Additionally, data reduction schemes could be computationally expensive -- can these tasks be accomplished on board with expected spacecraft capabilities? Would the benefits be worth the costs of providing these capabilities?

This portion of the follow-on study would identify several key areas where processing could be moved to the spacecraft, develop a concept of operations for the processing, and discuss the benefits versus costs of these changes.

As an example, temperature and humidity data are currently derived from infrared and microwave radiances transmitted to the ground. If the data reduction were to be performed in space, the amounts of data downlinked could be greatly reduced. However, some analysis schemes are being used that directly ingest these measured radiances rather than the derived environmental parameters. Would eliminating the availability of radiance data at the ground adversely impact the quality of the analysis? Even now, some organizations are moving away from radiance assimilation.

Other areas of data processing to be considered could include (but are not necessarily be limited to):

- Data Quality Control
- Rapid Event Detection
- Calculation of the forward model results (needed for QC)
- Data Analysis (reduction of sensor-based coordinates to model grid coordinates)

5.1.2. Assimilation and Forecast CONOPS

In the original gap study, computational requirements were estimated using an assumed concept of operations (CONOPS) for data assimilation and global forecast generation. Among the variables for which values were chosen (and that could have large impacts on the MDAS computational resource requirements) are:

- Frequency of assimilation runs: Hourly was chosen, but other intervals might produce better quality products -- even a continuous assimilation process has been suggested in the community.
- Number and types of ensemble forecasts: For current long-range forecasts, numbers of ensemble members range from 10s to 100s. Furthermore, other ensemble approaches (e.g., Monte Carlo suites) have been suggested that could greatly impact the computing resources required.
- Targeted observation methodologies: Currently proposed techniques for selecting observations to be collected revolve around calculating the adjoint of the models. Other techniques might be available that would have different levels of computational cost and produce better results.

The follow-on study should gather information from domain experts and literature review to generate options for an assimilation and forecast CONOPS. Using this updated information, computational resource requirements will be re-estimated to provide a more accurate range of values needed.

5.1.3. SensorWeb Management and Monitoring

Much of the intelligence surrounding the management and monitoring of the SensorWeb was not fully developed in the original study. Some of the aspects not fully explored include:

- Architecture requirements (e.g., communications and computing needs)
- Timing requirements for SensorWeb responses (e.g., how quickly does the SensorWeb need to respond to events detected by other portions of the SensorWeb? How does communications latency affect the response?)

The follow-on study should develop a CONOPS specific to the SensorWeb that will address the intelligence required (both distributed and integrated), define various options for the location of various portions of the intelligence, and examine how these options will affect the architecture needs. Updates to these needs will be examined to see how they might affect the gap analysis.

5.1.4. Architecture Management and Monitoring

In the original gap study, some aspects of the overall system monitoring and management were given cursory examination. This portion of the architecture is responsible for monitoring the performance of the system as a whole. These functions also provide for such things as the setting of "policy" items (e.g., forecast production schedules), approval of science community requests for data collection, or human update/intervention into system operations.

Although these functions were discussed in the study report, very little detail was provided about what impacts to the overall architecture these functions might have. The follow-on study should provide additional detail of the overall architecture.

5.2. Three to Five Day Forecast Study

This follow-on study would build on the same high level system architecture concept that emerged from this study, that was non-specific in some areas since it necessarily involved [educated] speculation on almost every relevant future technological capability from constellation management, to computing technologies, to communications. In order to bore down into the deeper meaning of the two way interaction, we think it is essential now to hold some variables constant so that we may focus on the system architecture question in more concrete terms. This increase in detail can be obtained by focusing on a well-controlled scenario that is known to be tractable (1-5 day forecasting).

For control, this new study would start by assuming only the capabilities of research and operational space-based observing systems that are being planned now for deployment in the 2010 – 2020 time frame (e.g. GPM, NPOESS, GIFTS), and/or technologies that are fully expected to have reached a prototype demonstration stage of maturity the 2010 – 2025 time frame. The basic observing characteristics of these future systems are more or less given. Also, to sustain focus on the system architecture, this follow-on study should use as science scenarios a finite set of weather phenomena whose evolution and prediction would be encompassed over time scales ranging from 24 hours to 120 hours. For example, localized severe thunderstorm forecast 24 hours in advance, prediction of East Coast snowstorms 4-5 days in advance, or prediction of devastating winter low-pressure systems that impact California in El Nino years. By naming the scenarios and phenomena of interest up front, and knowing the class of observations that will or should be available, a higher level of specificity can be obtained. This will permit a better focus on how the entire system must operate and be designed in order to provide the needed coordination between and among space and terrestrial based observing systems, and operational weather modeling systems.

Given more concrete notions based on realistic use-case scenarios of the data flows and desired interactions among the system components, it will be possible then to consider the system logic, architectures and technologies, as well as advances in system theory, communications and that could provide the necessary interactivity and results from a highly intelligent, highly integrated operational weather forecast system not only for short to medium range forecasting but out to 10 – 14 days.

2025 Wx Study - Gap Analysis Chart (5/31/2002)				Technology					Science	
TECHNOLOGY REQUIREMENTS	**Performance Objective**	Capability Maturity	Challenges Investments	Computing Technology Challenges	Communication Technology Challenges	Detector Technology Challenges	Space System Technology Challenges	Software Engineering Challenges	Scientific Research Challenges	Scientific Infrastructure Challenges
		1-8	1-8	1-8	1-8	1-8	1-8	1-8	1-8	1-8
OBSERVING SYSTEM										
Command and Control										
Artificial Intelligence (event detection)	**Semi-autonomous control of 30 spacecraft**	6.5	6.5	6.5	5.5	N/A	6.5	6.5	6.5	
System-wide Distributed Intelligence	**Semi-autonomous control of 30 spacecraft**	3.5	6.5	6.5	5.5	N/A	5.5	6.5	6.5	
Dynamically Reconfigurable Mission	**Semi-autonomous control of 30 spacecraft**	7.0	6.5	6.5	5.5	N/A	6.5	6.5	6.5	
Spacecraft & Instrument Tasking	**Semi-autonomous control of 30 spacecraft**	4.0	4.0	3.0	5.0	N/A	4.5	3.5	3.0	
Coordinated Multi-Spacecraft Observing	**Semi-autonomous control of 30 spacecraft**	4.0	4.0	3.0	5.0	N/A	4.5	3.5	3.0	
Coordinated Space & Ground Observing	**Semi-autonomous control of 30 spacecraft**	4.0	4.0	3.0	5.0	N/A	4.5	3.5	3.0	
Mission / Situation Dependent Observing	**Semi-autonomous control of 30 spacecraft**	4.0	4.0	3.0	5.0	N/A	4.5	3.5	3.0	
Dynamic Asset Re-assignment / Optimization	**Semi-autonomous control of 30 spacecraft**	4.0	4.0	3.0	5.0	N/A	4.5	3.5	3.0	
Communication, External Interfaces	**No Gap**									
Via Terrestrial Backbone Network Infrastructure	**No Gap**									
M/DAS	**No Gap**									
"Guiding Intelligence"	**No Gap**									
Computing	**No Gap**									
Observing System Network										
Communication, Internal		4.5								
Real-time Global Communication and Data Exchange		4.5								
RF Bandwidth		4.5								
Space System Architecture		4.5								
Seamless Space Ground Communications		4.5								
Internet in Space		4.5								
Protocol Development		4.5								
Data Security & Encryption		4.5								
Laser Communication (Space to Space)		4.5								
Forward Model										
Ground-based computing		2	2	1	NA	NA	NA	1	2	2
On-board computing		3	3	3	3	NA	2	3	2	3
Space-Based and In-Situ Observations										
Space-based Precipitation (Active)	**Global 1-hr @ 1km resolution**	4.5	4.5	2	2	3	6	3	4.5	4.5
Multi-frequency/polarized Precipitation & Cloud Radar		5.5		2	N/A	3				
Multple Phased-Array Antennae Technology		4.5		2	N/A	3	6			
Large Deployable Antennae	**1 km driver**	6		2	N/A	N/A	6			
On-board Power Generation (RF)		3		2	N/A	N/A	3			
Pointing Control Requirements				2	N/A	N/A	3			
On-board Processing				2	N/A	N/A				
Science Algorithm Processing				2	N/A	N/A				
Artificial Intelligence Overhead				2	N/A	N/A				
Calibration & QC				2	N/A	N/A				
Space-based Laser Winds (Active)		6	7	2	N/A		6	N/A		
Telescopy (heterodyne vs		5.5		2	N/A		6	N/A		
On-board Power Generation (Laser & SC control)	**minimum: 4kw avg, 7kw peak**	Schnurr		2	N/A		6	N/A		
UV Laser Technology & Techniques				2	N/A		6	N/A		
IR Laser Technology & Techniques				2	N/A		6	N/A		
Frequency Agile Receivers transmitters				2			6			
On-board Processing				2	N/A		N/A	N/A		
Doppler Signal Processing				2	N/A		N/A	N/A		
Artificial Intelligence Overhead				2	N/A		N/A	N/A		
Calibration & QC				2	N/A		N/A	N/A		
Pointing & Alignment				2	N/A					
Space-based Soil Moisture	**Global 1-hr at 1 km horiz. resolution**	6.5	5							
Active										

2025 Wx Study - Gap Analysis Chart (5/31/2002)				Technology					Science	
TECHNOLOGY REQUIREMENTS	**Performance Objective**	Capability Maturity	Challenges Investments	Computing Technology Challenges	Communication Technology Challenges	Detector Technology Challenges	Space System Technology Challenges	Software Engineering Challenges	Scientific Research Challenges	Scientific Infrastructure Challenges
		1-8	1-8	1-8	1-8	1-8	1-8	1-8	1-8	1-8
Passive										
Hyperspectral Sensing (Passive)	IR Vertical Profiles vs. Land Characterization	2.5	5							
Detector Technology										
On-board Processing										
Signal processing										
Onboard image analysis & pattern recognition										
Science Algorithm Processing										
Artificial Intelligence Overhead										
On-board Data Storage										
Space-based 3D Temperature & Humidity	**Global 3-hr @ .25 km vert. & 25 km horiz. res.**	3	5							
Interferometric Sounding Methods		4	4.5							
GPS Refraction / Limb Occultation Methods		3.5	5.5							
Spacecraft to Spacecraft		6	5							
GPS to Spacecraft		3.5	5.5							
Differential Absorption LIDAR for humidity									2	4
Raman LIDAR for Pressure and Chemical Contituents										
Surface Ocean Winds	**Global 3-hr winds @ 25km vertical resolution**	1	4					2		
GPS Reflection		6	4.5						5.5	4
Scatterometry		1.5	4						2.5	
MW Polarimetry		3								
Conventional Terrestrial-based Observations		2	4							
Surface Shelter Observations (T,P,V,Q,C,R)		1	2	1	2	1		1	1	2
Aircraft (T,Q,V)		2.5	4	1	2	4		1		
Automated reporting			2							
Directed Dropsonde Release			2							
Radiosondes (T,Q,V)		1	2	1	2	2		1	2	2
Automated Remote Release		2.5	4							
Unmanned Aerial Vehicles		5	4	3	3	N/A	N/A	3	N/A	N/A
EXPANSIONS										
Guidence, Navigation, and Control		6.5	4.5	5	4	4	6.5	6.5		
Pointing, Stabilization, Tracking Control, and geolocation performance	sub-mas knowledge, mas ctrl	6	4	5			6	4		
Low Cost/Power GPS time, position, attitude determin.	high=allt, < $50k, < 5 W	6	4	6			7	7		
"Drag-Free Orbit Control System Technology	sub-meter accuracy	5	4	1			4	1		
Innovative Multi-Function S/C Systems	highly integrated bus	5	5	4	4	4	5	3		
Spacecraft Autonomous Orbit Maintenance	30 S/C Management, no operator in loop	6	4	3	5		4	4		
On-Board Computing and Storage		6.5	6.5	6	6.5	N/A	5.5	6.5	6.5	
Image Processing & Analysis				6	6.5	N/A	5.5	6.5		
Data Processing & Analysis		6.5	6.5	6	6.5	N/A	5.5	6.5	6.5	
Geolocation		5	5	4	N/A	N/A	2	2		
Calibration Functions		4	4	4	3	N/A	4	5		
Quality Control Functions		6.5	6.5	6	6.5	N/A	4	6.5	6.5	
Geophysical Algorithm processing		6.5	6.5	6	6.5	N/A	4	6.5	6.5	
Processor Speed and Capacity		3	3	3	6	N/A	3	3	3	
On-board Data Storage		3	3	3	3	N/A	3	3	3	
Radiation Hardening		3	3	3	4	N/A	5.5	3	3	
Distributed, Dynamically Allocable Computing		6	6	5	5.5	N/A	5.5	6	5	
Reconfigurable, Mission-Dependent Software (FPGA)		6	6	5	5.5	N/A	5.5	6	5	
Artificial Intelligence Overhead		6	4	3	5	N/A	4	4	5	
Computing (High Performance)										
Data Storage & Archival	1.0 e+15 Bytes/day	3	3	3	3	N/A	N/A	4	4	
System Architectures & Data Systems		6	6	6	6	N/A	N/A	5	5	

2025 Wx Study - Gap Analysis Chart (5/31/2002)				Technology					Science	
TECHNOLOGY REQUIREMENTS	**Performance Objective**	Capability Maturity	Challenges Investments	Computing Technology Challenges	Communication Technology Challenges	Detector Technology Challenges	Space System Technology Challenges	Software Engineering Challenges	Scientific Research Challenges	Scientific Infrastructure Challenges
		1-8	1-8	1-8	1-8	1-8	1-8	1-8	1-8	1-8
Processing Capacity	1.0 e+10 GFLOPS	7	7	7	5	N/A	N/A	6	6	
MODEL & DATA ASSIMILATION SYSTEM										
Data Ingest and Preprocessing	Ingest and Pre-processing of >> 1 Tbyte / day	1	3	2	2	N/A	N/A	2	N/A	N/A
Data Aggregation and Reduction		1	3	2	2	N/A	N/A	2	N/A	N/A
Communication		1	4	2	N/A	N/A	N/A	2	N/A	N/A
Artificial Intelligence		1	3 - 4	2	N/A	N/A	N/A	3 - 4	3	N/A
Computing Capacity		2	1 - 2	3	N/A	N/A	N/A	3	N/A	N/A
Quality Control Function		1	1	1	N/A	N/A	N/A	1	2	2
QC Methodologies		1	1	1	N/A	N/A	N/A	1	2	2
Computing Speed & Capacity		1	1	1	N/A	N/A	N/A	1	2	1
Analysis and Assimilation Function	**Hourly analysis of 1.0 e+8 obs on 25 km grid**	7	5	6 - 7	N/A	N/A	N/A	2	4	4
4DVAR Methodologies		3	5	6	N/A	N/A	N/A	2	4	4
Kalman Filtering Methodologies		6 - 7	5	7	N/A	N/A	N/A	2	4	4
Computing Speed & Capacity		7 - 8	7	7	N/A	N/A	N/A	5 - 6	4	4
Targeted Observation Methodologies		5.5	3	3	N/A	N/A	N/A	3	4	4
Error Growth Estimation / Prediction		5.5	3	3	N/A	N/A	N/A	3	4	4
Stichastic / Ensemble Predictions		5.5	3	3	N/A	N/A	N/A	3	4	4
Computing Speed & Capacity		7	7	7	N/A	N/A	N/A	3	4	4
Global Mesoscale Model	**1-10km Resolution Global Atmospheric Model**	7	6	4	N/A	N/A	N/A	4	6 - 7	5
Numerical Solutions & Techniques		5 - 6	5	3	N/A	N/A	N/A	3	5	4
Adaptive Grid Techniques		5 - 6	5	4	N/A	N/A	N/A	5	5	4
Targeted Observation Methodologies		4	5 - 6		N/A	N/A	N/A	4	5	4
Parameterization Development		6 - 7	6 - 7	5	N/A	N/A	N/A	4	5 - 6	4
Computing Speed and Capacity		3	3	3	N/A	N/A	N/A	4	N/A	N/A
SCIENCE INFRASTRUCTURE		4		2	N/A	N/A	N/A	2	2	2
Observing System Simulation Testbed		4		2	N/A	N/A	N/A	2	2	2
Scientific Community Interfaces	**Internet Addressable Sensor Web**	4		2	N/A	N/A	N/A	2	2	2
KNOWLEDGE DELIVERY	Product generation and dissemenation to multiple users	2	2	2	3	N/A	N/A	2	N/A	N/A
Post Processing	**NO GAP**	2	2	2	3	N/A	N/A	2	N/A	N/A
Dissemination	**NO GAP**	2	2	2	3	N/A	N/A	2	N/A	N/A
Intelligent Data Archives	**NO GAP**	2	2	2	3	N/A	N/A	2	N/A	N/A
Wireless	**NO GAP**	2	2	2	3	N/A	N/A	2	N/A	N/A
System Monitoring and Policy Mgt.	**NO GAP**	2	2	2	2	N/A	N/A	2	N/A	N/A

	Capability Maturity			**Challenges & Investments**
(Difficult) 8	Concept Postulated	**8**	**High**	Useful progress unlikely no matter the investment
7	Principles Understood	**7**		Useful progress likely with unprecedented investment
6	Feasibility Established	**6**		Complete success likely with unprecedented investment
5	Design Proposed	**5**		Useful progress achievable with new precedented Investment
4	Design Prototyped	**4**	**Medium**	Complete success achievable with new precedented investment
3	Operational Demonstration	**3**		Useful progress likely based on evolutionary advances
2	Operational Implementation Proposed	**2**		Complete success likely based on evolutionary advances
(Easy) 1	Operational Implementation Planned/Exists	**1**	**Low**	Complete success achievable with minimal development / resources