SPACE-BASED DOPPLER LIDAR SAMPLING STRATEGIES -- ALGORITHM DEVELOPMENT AND SIMULATED OBSERVATION EXPERIMENTS

Final Report
Under
NASA Contract NAS8-38559

Covering the Period
June 27, 1990 to December 1, 1993

Submitted by

Simpson Weather Associates, Inc.
809 E. Jefferson Street
Charlottesville, VA 22902

1 December 1993

<u>TABLE OF CONTENTS</u>

PAGE

1.0 Introduction... 1

2.0 LAWS Simulations for the OSSEs.......................... 1

3.0 LAWS Trade Studies..................................... 2

4.0 Ground-Based Efforts................................... 3

5.0 GLOBE Data in HDF Format................................ 3

APPENDIX A: LAWS SIMULATION MODEL (GLOBAL VERSION).......... A-1

APPENDIX B: 1992 OBSERVING SYSTEM SIMULATION EXPERIMENTS.... B-1

APPENDIX C: 1991 OBSERVING SYSTEM SIMULATION EXPERIMENTS.... C-1

APPENDIX D: GROUND-BASED DOPPLER LIDAR STUDIES.............. D-1

APPENDIX E: ADDITIONAL STUDIES............................. E-1

APPENDIX F: HDF FORMAT FOR GLOBE DATA....................... F-1

APPENDIX G: CONFERENCE PRESENTATIONS AND PAPERS............. G-1
REFERENCES... R-1

SUMMARY

1.0 Introduction

Simpson Weather Associates, Inc. (SWA) has been funded, under contract NAS8-38559, to research the feasibility and optimum functionality of the NASA proposed space-based Doppler lidar, the Laser Atmospheric Wind Sounder (LAWS), to measure global winds in the earth's troposphere. SWA has developed a LAWS Simulation Model (LSM) to evaluate the potential impacts of LAWS measurements on the prediction skills of current climate and forecast models and to address critical LAWS design issues such as orbit selection, scanning patterns, and power budgets. This final report covers the period from June 27, 1990 to December 1, 1993, where SWA's efforts were largely concentrated on addressing the impacts of several orbit configurations on LAWS global coverage, measurement accuracy, wind information, trade studies and ground based Doppler lidar data processing. A complete technical write up of the LSM is included in Appendix A. Following is a brief overview of the contract activities.

2.0 LAWS Simulations for the OSSEs

In the first year of this contract, SWA was tasked to conduct a series of LAWS simulations to address the selection of orbit (altitude and inclination) and system design criteria for LAWS (see Appendix C.). The European Center for Medium Range Weather Forecasting (ECMWF) provided 25 days of sequentially consistent atmospheric fields referred to as the Nature Run. The global LAWS Simulation Model (LSMg) was run for several sets of five Nature Run days for several orbits. The simulated lidar data sets, i.e., Line-of-Sight (LOS) and horizontal components, were given to the ongoing Observing System Simulation Experiments (OSSE) being run at NASA/Goddard Space Flight Center (GSFC) (Atlas and Emmitt, 1991) and Florida State University (FSU) (Krishnamurti et al., 1991; Rohaly and Krishnamurti, 1993). The initial set of LAWS OSSEs were designed to bracket the performance of LAWS. The first series of LAWS simulations were considered to be "perfect" LAWS measurements (i.e., the best any optical wind sensor could do with 10^9 shots in 5 years). Only opaque clouds and terrain obscurations prevented LAWS from making a measurement at a pressure level. The second series of LAWS simulations were considered to be more "realistic" by incorporating the effects of variables such as aerosol backscatter, molecular attenuation, the presence of cirrus clouds and atmospheric turbulence in the calculations of the horizontal wind components. As these experiments were being run, the choice of orbit configuration was refined. Table 1 summarizes the LAWS orbital configurations that were simulated in 1990-1991 period.

Table 1. LAWS Orbital Configuration Experiments for 1991

System Configuration	Orbit	Simulation
20 JOULE/1.50 M	98 DEG 705 KM	PERFECT
20 JOULE/1.50 M	98 DEG 450 KM	PERFECT
20 JOULE/1.50 M	55 DEG 450 KM	PERFECT
20 JOULE/1.50 M	98 DEG 525 KM	PERFECT
20 JOULE/1.50 M	55 DEG 575 KM	PERFECT
20 JOULE/1.50 M	55 DEG 575 KM	REALISTIC
20 JOULE/1.50 M	98 DEG 525 KM	REALISTIC

In the second year of our contract, SWA was tasked to upgrade the LSMg to support the new ECMWF T106 (~ 100x100 km grid) Nature Run data, incorporate aerosol information from the NASA Global Backscatter Experiment (GLOBE) and incorporate the Slingo cloud prediction model (see Appendix A.). A series of realistic LAWS simulations were performed in cooperation with the OSSE efforts ongoing at GSFC and FSU (see Appendix B) to address the performance of LAWS under various system configurations ranging from a low power LAWS to the "full up" LAWS. Table 2 lists the LAWS system configurations that were simulated in 1992-1993 period.

Table 2. LAWS System Configuration Experiments

System Configuration	Orbit	Simulation
20 JOULE/1.50 M	98 DEG 525 KM	5 days
5 JOULE/0.75 M	98 DEG 525 KM	5 days
.1 JOULE/0.50 M	98 DEG 200 KM	5 days
.2 JOULE/0.60 M	98 DEG 300 KM	5 days
2. JOULE/0.75 M	98 DEG 300 KM	3 hours

3.0 Trade Studies

Prior to generating extended (several days) LAWS simulated data sets for use in OSSEs, an engineering version (developed with SWA R&D funds) of the LSM was used to define optimal system and sub-system parameters within specified technological or resource constraints. The system trade studies included the specification of:

- the optimal nadir scan angle for a conically scanned lidar;

- the optimal pattern of sampling for both conically scanned and multi-telescoped unscanned systems;

- the optimal combination of energy/pulse and pulse repetition frequency for a given platform power budget; and

- the optimal management of lidar shots given fixed laser-lifetimes and platform resources.

Since the trade studies were normally against some reference atmosphere, SWA developed a new LAWS Baseline Atmosphere Model (BAM) to be used for LAWS trade studies to NASA/Marshall Space Flight Center (MSFC) (Wood and Emmitt, 1990; Wood and Emmitt, 1991). BAM gives the probability of a given aerosol backscatter as a function of altitude. The presence of thin cirrus clouds and molecular attenuation is also available in the model. The original backscatter profile was based on GLOBE data. The new version has added an expected maritime backscatter profile based on the LAWS Science Team's research.

4.0 Ground-Based Studies

The development of an optimal design for a space-based lidar wind sounder has been an iterative process as the science community provides better data and product generation algorithms to the LAWS system simulation/impact studies. Without a space-based heritage, the Doppler lidar programs have had to rely on ground-based and airborne experience to provide some sense of the reasonableness of model simulations. As part of this effort, SWA developed a series of computer tools to process and analyze NASA/MSFC ground-based Doppler lidar polypulse-pair and single shot data. Appendix D describes the models and a series of ground-based experiments in system stability, accuracy versus SNR, winds from cirrus decks, single shot pairing and detailed error analyses.

5.0 GLOBE Data in HDF Format

The GLOBE database project involves the development of a system for archiving, distributing and retrieving aerosol data obtained from seven different sites. Over the past several months, the archival strategy was completed for one site (GSFC) with the goal of developing a storage format that could be generalized to data from all sites. The Hierarchical Database Format (HDF) routines called from C programs were used to archive the data. The use of HDF allowed data and meta-data to be stored in the same HDF file. This strategy will make it possible in the future to distribute GLOBE database project files to interested users through the DAACs or the Earth Observing System Data Information System (EOSDIS).

APPENDIX A
LAWS SIMULATION MODEL (GLOBAL VERSION)

1.0 Introduction

The LAWS Simulation Model (LSMg) simulates a space-based Doppler lidar wind sounder providing global or regional three-dimensional simulated lidar winds and corresponding errors. The major modules of the LSMg include the platform, orbit, scanner, atmospheric library, line of sight, horizontal wind component and error models.

1.1 **Past and Current LSM Applications**

Over the past 8 years, NASA has supported SWA to develop LAWS Simulation Models that evaluate the potential impact of direct global wind observations on global climate models and current forecast models (GCM and regional scale). Under previous contracts, SWA developed two basic algorithms for use with simulated Doppler lidar wind profilers (Emmitt and Wood, 1988). A Shot Management Algorithm (SMA) controls timing and placement of lidar pulses and a Multi-Paired Algorithm (MPA) extracts horizontal wind components from the lidar radial velocity observations. These algorithms evolved into a fully integrated 'top to bottom' LAWS Simulation Model (Emmitt et al., 1990) that modeled the platform, orbit, scanner, signal processor along with an atmospheric library that incorporated the effects of atmospheric aerosols, water vapor, clouds, terrain and atmospheric turbulence. SWA used the LSM to address key LAWS issues and trades such as global coverage, accuracy and representativeness of LAWS measurements, data density, signal strength, cloud obscuration and temporal resolution.

Under the current contract, the LSM has evolved into a set of computer models that is schematically depicted in Fig. A.1. The current LAWS Simulation Model (LSMg) was designed to address questions ranging from the feasibility and optimal functionality of a space-based Doppler lidar system to the impact of global winds observations on the predictive ability of GCM forecast models. The LSMg assesses the impact of global winds by providing simulated lidar winds to institutions such as NASA/GSFC and FSU for Observing System Simulation Experiments (OSSEs) as described in Fig. A.2 (Atlas and Emmitt, 1991; Krishnamurti et al., 1991; Rohaly and Krishnamurti, 1993).

1.2 **LAWS Simulation Model Inputs and Outputs**

A new methodology for using SWA's LAWS Simulation Model (LSMg) to provide simulated LAWS winds and errors to OSSEs was introduced in this contract. Due to the cost of OSSEs and the growing need to expedite the turn around time for impact assessment, the LSMg was recoded. The LSMg was redesigned to run

on the NASA/GSFC CRAY YMP. The LSMg operates via three data file: an operations data file (LSMINP), an orbital coverage data set (COVDAT) and a Global Atmospheric Data Set (GADS). The LSMg operational inputs are read in from LSMINP that define the LAWS simulation run, system configuration, orbital parameters and laser wavelength. Current system and platform configurations are discussed in Section A2.0. A COVDAT file contains satellite and shot locations for a given orbit and scanner configuration and is discussed in Section A3.0. The LSMg shot management algorithms are discussed in Section A4.0. The GADS file contains ECMWF T106 nature run profiles, aerosol and water vapor optical properties, clouds and terrain and is discussed in Section A5.0. Signal processing models are discussed in Section A6.0. The LSMg provides three output files; level 1 LAWS products (LEVEL1), level 2 LAWS products (LEVEL2) and a performance matrix (MATRIX). The level 1 products are defined as simulated lidar line-of-sight winds and associated signal processing variables (Section A7.0). The level 2 products are defined as simulated lidar horizontal wind components and associated observational errors (Section A8.0). The performance display model and graphic algorithms are discussed in Section A9.0.

2.0 System Configurations

The LSMg supports a broad range of laser/optic system configurations. Table A.1 lists the primary parameters used in the LSMg.

Table A.1. System Configuration Parameters

LASER WAVELENGTH (µm)
TRANSMITTED ENERGY (Joules)
TELESCOPE DIAMETER (m)
NADIR SCAN ANGLE (DEG)
PULSE REPETITION FREQUENCY
OPTICAL EFFICIENCY
DETECTOR QUANTUM EFFICIENCY
HETERDYNE MIXING EFFICIENCY
PULSE DURATION

3.0 The Orbit/Scanner Model

The Orbit model provides the satellite latitude and longitude and the inclination angle of the orbit as a function of orbital time. The orbit model solves an oblique spherical triangle algorithm to compute the satellites location. The model assumes a spherical earth. Figure A.3 demonstrates an example of a 13 hour LAWS coverage orbit.

Equation 3.1 gives the speed of the satellite as a function of satellite altitude.

$$V_{sat} = ((GM/((R_e + Z_s)/1000))^{1/2})/1E06 \qquad (3.1)$$

where

V_{sat} (km/s)	satellite velocity
GM	Earth constant, 3.991e14
Z_s (km)	satellite altitude
R_e (km)	radius of earth.

The scanner models computes the latitude and longitude of each lidar shot as a function of atmospheric level and azimuth scan angle. The Eqs. 3.2-3.5 involve the nadir scan geometry angles and swath width, respectively, as shown in Fig. A.4.

$$= - SIN^{-1}((Z_s+R_e)/(R_e \quad SIN( )) \qquad (3.2)$$

$$= - - \qquad (3.3)$$

$$= /2 - - \qquad (3.4)$$

$$SW = 2 \quad /360 * 2 \quad R_e \qquad (3.5)$$

where

(rad) satellite to shot to center of the earth angle
a constant, 3.14159
Z_s (km) satellite altitude
R_e (km) radius of earth
(deg) nadir scan angle
(rad) satellite to earth's center to shot angle
(rad) slant path elevation angle
SW (km) swath distance.

The scanner model uses the longitude and latitude of the satellite and the azimuth scan angle of the laser along with the orbital parameters (satellite altitude and velocity, scanner period, etc.) to compute laser shot locations. The model uses an oblique spherical triangle algorithm (Kells, et al, 1940) that solves for a spherical triangle defined by the north pole, the position of the satellite and the position of the shot.
The scanner model supports conical and fixed beam patterns as demonstrated in Fig. A.5. The conical scanner gives the latitude and longitude of the lidar shot and the azimuth scan angle for a counterclockwise scanning lidar. Figure A.6 shows the conical scanning geometry convention.

While conical scanning provides the best coverage/accuracy performance, the physical complexity of the large rotating telescope makes a set of fixed telescopes a reasonable alternative. The French BEST program proposed a quad-beam arrangement illustrated in Fig. A.7. The LSMg was upgraded to include a quad-beam simulation. The quad-beam scanner gives the latitude and longitude of the lidar shot for four telescopes with fixed azimuth and nadir look angles.

4.0 Shot Management

Unlike most passive sensors in space, active laser based systems have limited lifetimes (pulses) and are ultimately constrained by available platform power. Such conditions call for some form(s) of resource management that will optimize the number of useful observations and the potential impact on the primary mission objective - i.e., improved understanding of the global circulations and transports.

Management of the lidar pulses as become a core focus of reserach under this contract. The objectives of shot management include:

1) extend mission lifetime;

2) optimize within a scan distribution of shots to obtain best wind measurements; and

3) optimize global distribution of shots within an orbit to favor regions of high ageostrophic (i.e., tropics, jet streams, major mountain ranges, etc.).

To meet these objectives, seven modes of shot management have been defined (Table A.2). While the detailed options of scheduling lidar pulses are unlimited, the general sense of the management is to use a finite number of shots to achieve the best set of data for a given mission objective. For example, if the mission objective is to provide full global coverage every 12 hours, then a combination of modes 2 and 3 is in order. If the mission objective is to provide direct measurements of winds in regions of ageostrophic flows, then a combination of modes 4 and 5 may be proper. If the mission objective is to provide data preferentially in regions where a forecasting model is having difficulties, them mode 7 would be employed.

The LSM is designed to invoke modes 1 through 5. The most common mode combination is 2 plus 3. Mode 2 applies only to the conical scanner. Given the shot number in a no cosine modification scan (Eq. 3.6) and the rotation rate, the model computes a new scan azimuth angle (Eq. 3.7) to back out a new time for the shot (Eq. 3.8). Figure A.8 gives a comparison of

Table A.2. LAWS Shot Management Modes

MODE	DESCRIPTION	RATIO[1]
1	Constant PRF at 100% duty	1
2	Cosine modulation of PRF within a scan period	1
3	12-hour polar redundancy suppression	.7
4	Tropical enhanced	.7
5	Ageostrophic priority	.1-.5
6	Condition recognition (onboard)	.7- 1
7	Condition recognition (up-linked)	.7- 1

Note 1: Ratio of shots taken per orbit for each mode compared to Mode 1.

the conical scan with and without the cosine modifier.

N_{sht} = NINT((T/(V_{rot}/$\#_{sht}$) - (INT(T/V_{rot}) $\#_{sht}$))   (3.6)

IF N_{sht} 1 AND N_{sht} (0.5 $\#_{sht}$)

$ß^{'}$ = COS^{-1}(1.0 - (4 N_{sht})/(($\#_{sht}$/V_{rot}) V_{rot}))   (3.7)

IF N_{sht} (0.5 $\#_{sht}$)   AND N_{sht} $\#_{sht}$

$ß^{'}$ = 2 - COS^{-1}((4 N_{sht})/(($\#_{sht}$/V_{rot}) V_{rot}) - 3)   (3.8)

TIME = V_{rot} $ß^{'}$ /2 + T

where

T   unmodified time,
T(s)   change in time between shots,
$\#_{sht}$ number of shots in one scan,
N_{sht} shot number in scan,
V_{sat} (km/s)   satellite velocity,
V_{rot} (s^{-1})   scanner rotation,
$ß^{'}$ (deg)   modified azimuth scan angle,
TIME (s)   modified time.

5.0 <u>Atmosphere Generator Model for the Global Atmospheric Data Set</u>

The Atmosphere Generator Model (AGMg) creates Global Atmospheric Data Sets (GADS) that the LSMg uses to obtain atmospheric profiles as a function of latitude and longitude. Each profile contains an ECMWF T106 Nature Run profile, aerosol and water vapor optical properties, clouds and terrain. GADS is a direct access file created by using the T106 data in liaison with the LSMg optical property models, cloud models, and terrain data set. The spacial resolution of the data set is 111.11 km by 111.11 km and the temporal resolution is 3 hours. It contains 32,412 grid area profiles (records) per time period. A record's content is defined below.

Table A.3. GADS Record Content

- center of grid area latitude (deg)
- center of grid area longitude (deg)
- surface pressure (mb)
- surface altitude (km)
- cloud cover (0-1)
- level index for base level above the terrain
- tropopause height level index
- surface temperature (k)
- surface dewpoint temperature (k)
- surface horizontal wind component U (m/s)
- surface horizontal wind component V (m/s)
- convective precipe (mm/day)
- height profile (m)
- temperature profile (k)
- relative humidity profile (%)
- u horizontal wind profile (m/s)
- v horizontal wind profile (m/s)
- vertical velocity profile (pa/s)
- cloud percentage profile (%)
- integrated cloud percentage profile (%)
- integrated cloud percentage profile as seen by LAWS (%)
- incremental cloud amount (%)
- layer index for a cirrus cloud
- cloud type index
    - 0 - opaque cloud or no cloud
    - 1 - thick cirrus cloud
    - 2 - thin cirrus cloud
- aerosol backscatter profile at 2.1 µm
- aerosol backscatter profile at 9.11 µm
- aerosol backscatter profile at 10.59 µm
- molecular attenuation profile at 2.1 µm
- molecular attenuation profile at 9.11 µm
- molecular attenuation profile at 10.59 µm

- cloud backscatter profile
- cloud attenuation profile

The T106 profile data has data for the following pressure levels: 1013, 1000, 850, 700, 500, 400, 300, 250, 200, 150, 100, 70, 50, 30, 10 mb. Figures A.9-A.25 are global maps of the GADS on January 16, 1987, 0300Z, for temperature, relative humidity and horizontal wind components at the surface and at the 500 mb level.

5.1 **Optical Property Models**

The AGMg has several options for estimating aerosol backscatter and molecular (water vapor) attenuation coefficients. The first option uses AFGL optical property models (LOWTRAN 7) in conjunction with the ECMWF global circulation data sets. A second method uses Global Backscatter Experiment (GLOBE) data sets in conjunction with global circulation model data sets. A third option uses the baseline atmosphere median profiles (Wood and Emmitt, 1991) in conjunction with the global circulation model data sets.

5.1.1 AFGL Optical Property Models

The AGMg contains a "trimmed down" version of the LOWTRAN 7 aerosol models (Kneizys et al., 1983) and FASCODE (Gallery et al., 1983) molecular attenuation models. This algorithm coupled with a global data base of LOWTRAN inputs and ECMWF meteorological profiles provides aerosol backscatter and water vapor attenuation for the laser wavelengths, 9.11, 10.59 and 2.1 µm. The optical property's natural variability due to altitude, location, seasons, and meteorological conditions are taken into consideration in the model.

The AGMg optical property model reads in an input index from a FORTRAN direct access data file, OPTTOP. Each index is stored as a function of latitude and longitude (-90S to 90N for -180W to 180E) at a 1 X 1 resolution. The index is used as a pointer for 5 pre-defined LOWTRAN inputs. The inputs are location profile model, haze model, coastal influence parameter, stratosphere model and upper atmosphere model. Given the ECMWF meteorological profile and shot location, the AGMg can estimate the optical properties as shown in Table A.4.

The OPTTOP data base is considered to be a baseline nominal data base. The model location index is a function of latitude. The haze model is a function of continent vs. desert vs. ocean location. The current version of the data set does not use the urban, ocean or fog models, but the data bases are stored in the AGMg for future upgrades. The coastal influence is a function of distance from a land mass (only three of the AFGL LOWTRAN options

Table A.4. Optical Property Model Options

Seasonal Models Location	Aerosol Haze Models	Coastal Influence Parameter	Stratospheric Model	Upper Atm. Model
Tropical	Rural 23 km vis	Open Ocean	Background Stratospheric	Normal Upper
Subtropical Summer	Rural 5 km vis	Midway to Continent	Moderate Aged Aerosol	Extreme Upper
Subtropical Winter	NAVY Maritime	Close to Continent	Moderate Fresh Aerosol	Volcanic to Normal
Midlatitude Summer	Ocean		High Aged Aerosol	Normal to Volcanic
Midlatitude Winter	Urban		High Fresh Aerosol	

Sub-Arctic Summer	Tropo-spheric		Extreme Aged Aerosol	
Sub-Arctic Winter	Desert		Extreme Fresh Aerosol	
U.S. Stndrd Summer	Fog 1			
U.S. Stndrd Winter	Fog 2			

are used). The stratosphere and upper atmosphere variables are set to clean background mode. Any variations to the computed baseline optical properties are made in the LSMg. For example, if one wanted to include a volcanic stratospheric dust level or an advection of maritime aerosols over Europe, one would use the LSMg to modify the baseline inputs as a function of latitude and longitude.

Given location and LOWTRAN pre-defined inputs along with the vertical profile of meteorological data from the ECMWF data set, the AGMg computes the scattering coefficient, backscatter phase function and the aerosol scaling parameter as a function of altitude. Aerosol backscatter is computed by Eq. 5.1 with consideration of aerosol natural variability due to local wind speed, temperature, relative humidity and standard visual range (Hanel, 1972). The AGMg computes the attenuation coefficient, shown in Eq. 5.2, from the molecular (water vapor) absorption, aerosol absorption, molecular scattering and aerosol scattering.

ß = ($_s$ P SCL)/1000 (5.1)

= $_m$ + $_m$ + ($_a$ + $_a$) SCL + RAY (5.2)

where

ß ($m^{-1}sr^{-1}$) aerosol backscatter coefficient
P (sr^{-1}) backscatter phase function
SCL aerosol scaling function,
(km^{-1}) attenuation coefficient
$_m$ (km^{-1}) molecular absorption coefficient
$_a$ (km^{-1}) aerosol absorption coefficient
$_m$ (km^{-1}) molecular scattering coefficient
$_s$ (km^{-1}) aerosol scattering coefficient
RAY (km^{-1}) Rayleigh scattering coefficient.

5.1.2 Aerosol Backscatter and Molecular Attenuation Distributions

Figures A.26-27 are illustrations of global distributions of aerosol backscatter coefficients at three levels for single time period of the ECMWF nature run. Below 850 mb, LOWTRAN/FASCODE is used to generate the backscatter. Above 850 mb to the tropopause, one of the three median GLOBE profiles is used, depending upon latitude and season, for backscatter. The three profiles obtained during GLOBE II were provided by Robert Menzies (Jet Propulsion Laboratory) and are shown in Fig. A28. Variability around the median values is slewed to variations in the relative humidity profiles provided by the ECMWF nature run. All the attenuation coefficients are computed from LOWTRAN/FASCODE models.

5.2 **Cloud Model**

The AGMg cloud model is based on the Slingo Cloud parameterization scheme (Slingo, 1987). The Slingo approach provides distinctions between high and mid-tropospheric stratiform clouds, convective clouds with and without anvil cirrus, low level clouds driven by weak vertical motion or inversion capped moist boundary layers.

5.2.1 Convective cloud

The convective cloud is inferred from 3 hour integrated precipitable water from the T106 meteorological profiles. A critical threshold value of 0.14 mm/day must be met for a convective cloud to be present. If the threshold is met, the convective cloud base amount is empirically derived. A restrictive limit of 80% is set for the convective cloud amount.

$$cc = 0.2473 + 0.1258 * cppt \qquad (5.3)$$

where

cc	is the base layer convective cloud amount (%)
cppt	is the integrated precipitable water (mm/day).

The top of the convective cloud layer is a function of the base layer convective cloud amount and the tropopause height. The cloud top is limited at the tropopause.

$$cc_{top} = (cc + 0.2) * TH \qquad (5.4)$$

where

cc_{top}	is the convective cloud top (km)
cc	is the base layer convective cloud amount (%)
TH	is the tropopause height (km).

The convective cloud coverage between cloud base and cloud top is defined as 25% of the base layer convective cloud amount.

If the top of the convective cloud is above the 400 mb layer and the integrated precipitable water more than 3.4 mm/day, then an anvil is defined. All anvil clouds are considered to be thick cirrus layers. The anvil cloud amount (%) is defined as

$$cc_{anv} = 2\ (cc - 0.3) \qquad (5.5)$$

5.2.2 High Non-Convective Clouds

All non-convective high clouds are derived as a function of relative humidity from the ECMWF T106 meteorological profile. A high layer cloud is only derived when the tropopause height is higher than the 400 mb layer. The AGMg evaluates the T106 relative humidity profile to find the highest value which is used to compute a relative humidity threshold.

$$RH_{thr} = (RH_{hgh} - 0.8)/(1.0 - 0.8) \qquad (5.6)$$

where

RH_{thr} is the high relative humidity threshold (%)
RH_{hgh} is the high relative humidity (%).

If the relative humidity threshold is greater than 0%, then high level cloud cover (%) is estimated as follows

$$HC = (RH_{thr})^2 \qquad (5.7)$$

The cloud is considered to be thin cirrus. See Section A5.2.6 for discussion on how this percent high cloud is used to provide variability in cloud optical depths.

5.2.3 Middle Non-Convective Clouds

All non-convective middle clouds are derived as a function of relative humidity from the ECMWF T106 meteorological profile. If there was a convective cloud or a high layer cloud, the AGMg drys out the T106 relative humidity profile.

$$RH = RH * (1.0 - CC) \qquad (5.8)$$

where

CC is either the convective or high layer cloud cover (%).

Like the high cloud algorithm, the AGMg finds the highest relative humidity in the profile and computes the relative humidity threshold (Eq. 5.6). If the relative humidity threshold is greater than 0%, then middle layer cloud cover (%) is estimated as follows

$$MC = HC + (RH_{thr})^2 \qquad (5.9)$$

5.2.4 Low Non-Convective Clouds

The estimate of low level non convective clouds is based upon two parameters: vertical velocity and potential temperature profile. From vertical velocity, the AGMg finds the layer with the largest negative vertical velocity and computes the critical relative humidity (Eq. 5.8) for the layer. If the vertical velocity is less than 0.1, the cloud cover is defined as

$$LC = (RH_{thr})^2 \qquad (5.10a)$$

else

$$LC = (RH_{thr})^2 * (-10 * VV) \quad (5.10b)$$

where

VV (m/s) is the vertical velocity for the layer.

The potential temperature is used only if there was no cloud cover from the vertical velocity method. Potential temperature lapse rates are computed for every sublayer between 1000mb and 700 mb as follows

$$_{lr} = -6.67 * \quad T / P \quad (5.11)$$

where

$_{lr}$ is the potential temperature lapse rate

T is the change in temperature over the layer
P is the change in pressure over the layer.

If the lapse rate is greater than zero, then the model tests upon relative humidity to compute the cloud cover. If the relative humidity is less than 60%, there is no cloud cover. If the relative humidity is greater than 60% and lower than the threshold relative humidity, then the cloud cover is

$$LC = {}_{lr} * (1 - (RH_{thr} - RH)/(1 - RH_{thr})) \quad (5.12a)$$

else

$$LC = {}_{lr} \quad (5.12b)$$

5.2.5 Global Statistics and Empirical Adjustments

We expect that clouds will be in the field-of-view (FOV) of LAWS 70-80% of the time. This estimate is based upon the recently reported analysis of two years of HIRS data (Menzel et al., 1992), the cirrus climatology derived from SAGE data by Woodbury and McCormick (1986) and the Nimbus-7 global cloud climatology (Stowe et al., 1989). Much of this cloud coverage is high cloud (above 400-500 mb) and is semi-transparent (~ 30-40%). Very thin or subvisual cirrus (<.07) is not detected by HIRS or Nimbus-7 but may be occasionally represented in the SAGE observations. Thus, we conclude that the occurrence of very thin cirrus is clearly underestimated in current climatologies.

Of particular interest to the LAWS program are the semi-transparent and optically thin clouds since they provide strong returns without full extinction. When one considers that the statistics given above are, in most cases, exclusive - i.e., they do not provide a good representation of coincident clouds at different altitudes, it is very likely that there are many occasions when there are multi-layers of thin clouds underlayed by opaque clouds.

The ECMWF T106 Nature Run provides accumulated convective precipitation (based upon Kuo's scheme) and the total cloud coverage. Unfortunately, layer by layer information on cloud cover is not provided. We have developed a means to use the basic concepts within the Slingo scheme to reproduce the layer by layer cloud statistics that we need for LAWS simulations (see Section 5.2).

The distribution of clouds (over a 1 x 1 area) based upon the ECMWF total cloud coverage as a function of latitude are shown in Figs. A.28 and A.29. While the total coverage is quite reasonable and compares well with the Nimbus-7 statistics, the amount of midlevel cloud forecast for the tropics is considerably

less than the 30-40% reported using the satellite data. Conversations with the National Meteorological Center's (NMC) personnel (Pan and Baker) suggest that this is an ongoing point for discussion and study with the modeling community suggesting that the interpretation of midlevel clouds in satellite imagery may be faulty.

For now, we do not attempt to force the models clouds to fit the climatology. The resulting global coverage maps for a single time period (0300 GMT) are shown in Figs. A.30-A.33.

5.2.6 Cloud Optical Properties

All cloud backscatter values are preset in the AGMg to be 1e-06 m^{-1} sr^{-1}. We believe this value is properly conservative, since recent mid-layer cloud backscatter, measured with a lidar in the Antarctic, range from 1e-06 to 1e-04 m^{-1} sr^{-1} (Del Guasta et al., 1993). Once a lidar shot gets a return from an opaque cloud, the shot is considered to be fully attenuated and is terminated. For cirrus cloud layers, we allow all the shots a finite probability to pass through to the next layer. The cirrus cloud attenuation model is a modified version of the analytical LOWTRAN cirrus algorithm (Kneizys, et al., 1983), where

$$= e^{-0.14 * L^2} \qquad (5.13)$$

is the cirrus transmittance
L   is the cirrus cloud thickness.

Since the AGMg is restricted to the coarse vertical resolution of the Nature Run, SWA uses the cirrus cloud percentage as a surrogate for cloud optical thickness.

The major assumption is that while the Slingo model derives a percent cirrus cloud coverage (i.e., 30%) from an average relative humidity within a grid volume, it is just as reasonable to interpret a thickness tendency from the same fields. Instead of using the percent coverage as literally meaning that 30% of the grid has cirrus cloud and 70% is totally cloud-free, the AGMg assumes that the whole grid area is covered by a cirrus cloud that has an optical thickness that scales to the percent coverage. The cirrus cloud attenuations is defined as

$$ci_{att} = 10 \ * (CLD_{\%} * 10)^2 \qquad (5.14)$$

where

ci_{att} is the cirrus cloud attenuation
is the LOWTRAN cirrus attenuation coefficient for a 1 km thick layer
$CLD_{\%}$ is the cirrus cloud percentage cover.

5.3 **Atmospheric Turbulence Model**

The LSMg has two options for estimating sub-grid scale variance. The first method uses rawinsonde uncertainties (300X300 km area) as a function of altitude as sampling scale variance in the gridded area. The second method attempts to better represent the sampling scale turbulence on the ECMWF grid scale and to provide "realistic" variability in the winds based upon the mean wind field. The model represents the uncertainties by scaling the mean ECMWF wind speed by 20%. Using one day of the ECMWF data, we computed the global distribution of variance for 12 atmospheric pressure levels. Table A.4 compares the total global averages with the NMC OI rawinsonde profiles and suggest that the simulated variances are not unreasonable. Figures A.34-A.39 show the global U and V variances for 500 mb and 250 mb levels.

5.4 **Terrain Model**

The terrain data base is a 1 x 1 direct access data file (OPTTOP) contains, in addition to the LOWTRAN index inputs, land elevation (M) and sea depths (M) for the globe.

Table A.4. Comparison of the NMC rawinsonde variances and the global average variances obtained by the LSM using a 24-hour sample of ECMWF data.

P (mb)	NMC Old	NMC New	LSMg Global Average
1000	1.8	1.4	1.44
850	1.8	2.2	1.65
700	2.4	2.4	1.81
500	3.8	2.8	2.42
400	4.7	3.4	3.00
300	5.9	3.4	3.61
250	5.9	3.2	3.87
200	5.9	3.0	4.00

150	5.5	2.7	3.83
100	4.9	2.5	3.12
70	4.9	2.5	2.80
50	3.9	2.7	2.80

6.0 Signal Processing Models

The LSMg has the option to use the narrow band signal to noise model or a consensus curve algorithm for signal processing of the lidar signal.

6.1 **Narrow Band Signal-to-Noise and Line-of-Sight Uncertainty Models**

There are several signal-to-noise (SNR) equations that have been suggested for use with LAWS. The narrow band SNR equation (along with default values) that we use in the LSMg is:

$$SNR_N = \frac{c\ _1\ _2\ _3\ _4\ J\ D^2\ \ ß\ e^{-2\ (r)dr}}{8\ h_v\ (R^2 + (.25\ D\ \ D/\ )^2)} \qquad (6.1)$$

where

c = speed of light (m/s)
$_1$ = heterodyne quantum efficiency
$_2$ = optical efficiency
$_3$ = beam shape factor
$_4$ = truncation factor
J = laser power (Joules)
D = mirror diameter (m)
 = pulse length (sec)
ß = backscatter ($m^{-1}\ sr^{-1}$)
e = 2 way attenuation
h_v = photon energy (J)
R = slant range (m)
 = laser wavelength (m)

As with the lidar SNR equation, there are several radial or LOS velocity error estimates, $_r$, that have been suggested for use with LAWS. While the Cramer-Rao Lower Bound may provide a limit to the extraction of a velocity estimate from a noisy signal, we have chosen the more conservative estimate based upon pulse pair autocorrelation processing of the Doppler signal. The following is derived from Eq. 6.22a in Doviak and Zrnic (1984).

$$_r = \frac{\ }{4}\left(\frac{f}{2t}\right)^{.5} (2\ ^{1.5}\ W + 16\ ^2\ W^2/SNR_w + 1/SNR_w^2)^{.5} \qquad (6.2)$$

where

= wavelength (m)

V_{max} = maximum velocity measured

f = sampling frequency = 2 V_{max}/

t = pulse duration (sec)

W = normalized frequency spread of return signal (m/s)

V_{bw} = uncertainty due to pulse bandwidth (m s^{-1})

V_{atm} = uncertainty due to turbulent eddies and wind shear within the pulse volume

SNR_w = 2 W SNR

6.2 Consensus Curves Model

Studies by Mike Hardesty and Barry Rye of NOAA have provided a general consensus algorithm used to simulate the processing of LAWS data.

The consensus algorithm computes wideband signal-to-noise (SNR_w) for each lidar shot along the slant path to a T106 nature run altitude as follows:

$$SNR_w = \frac{16\ h\ R^2\ V}{1.09 \times 10^{19}} * \ _m \ _o \ _{qe} * E_T D^2 * \ ^2 ß * \ ^2 \quad (6.3)$$

where

R = range

V = maximum wind window

$_m$ = mixing efficiency

$_o$ = optical transmission

$_{qe}$ = quantum efficiency

E_T = energy/pulse

D^2 = area of primary

^{2} = two-way transmission

ß = backscatter

^{2} = wavelength

The SNR_w is used to look up the probability of detection (POD) as shown in Fig. A.40a and the false alarm ratio (FAR) as shown in Fig. A.40b. The model uses the POD and the FAR to compute the probability of consensus as shown below

FARM = (FAR/100 * POD)/(1 - FAR/100)

(6.3)

CONS = POD + FARM

If the probability of consensus is greater than a random white noise value, the shot passes consensus. Once a shot passes consensus, the consensus algorithm test if the false alarm ratio is greater than a random white noise value. If true, then the shot is not a false alarm and the line-of-sight uncertainty is set to 0.5 m/s.

The measurement uncertainty for the observational error model is computed from Fig. A.40c (see Section 8.3).

7.0 Line-of-Sight Wind Products

If the lidar shot is not a false alarm, the LSMg computes the simulated line-of-sight wind velocity as

$$V_{los} = (U_{hwc} * \cos(ß) + V_{hwc} * \sin(ß)) * \cos(\ ) \quad (7.1)$$

$$+ W_{vwc} * \sin(\ ) + LOS_{unc}$$

where

V_{los}	is the line-of-sight wind velocity (m/s)
U_{hwc}	is the cross track wind velocity at the shot location (m/s)
V_{hwc}	is the along track wind velocity at the shot location (m/s)
W_{vwc}	is the vertical velocity at the shot location (m/s)
	is the elevation angle (radians)
ß	is the azimuth scanning angle (radians)
LOS_{unc}	is the line-of-sight uncertainty (m/s).

The along track and cross track winds at the shot location are computed from the input GADS horizontal wind components, orbit inclination and the sampling scale uncertainties that are obtained by using a random Gaussian distribution around the GADS turbulence profiles. The line-of-sight uncertainty comes from the consensus algorithm.

If the lidar shot is a false alarm, the LSMg compute the line-of-sight wind velocity using a random white noise value and the velocity maximum window as shown below

$$V_{los} = (RD - 0.5) * 2*V_{max} \quad (7.2)$$

where

V_{los}	is the line-of-sight wind velocity (m/s)
RD	is a random value (0-1)
V_{max}	is the velocity maximum window.

The LSMg LEVEL1 output products are listed in Table A.5.

Table A.5. Line-of-Sight Wind Products (LEVEL1)

Global grid box index for the LOS wind
Latitude of the LOS wind (deg)
Longitude of the LOS wind (deg)
Altitude of the LOS wind (km)
Time of the LOS wind (s)
Azimuth scan of the LOS wind (deg)
Sampling scale uncertainty in U of the LOS wind (m/s)
Sampling scale uncertainty in V of the LOS wind (m/s)
Sampling scale uncertainty in W of the LOS wind (m/s)
Line-of-sight uncertainty of the LOS wind (m/s)
Simulated line of sight wind velocity (m/s)

8.0 Horizontal Wind Component Models

The LSM uses either the Multi-Paired Algorithm (MPA) or the least squares techniques to compute the U and V horizontal wind components.

8.1 **Multi-Paired Algorithm (MPA)**

The MPA matches forward and aft lidar line-of-sight wind velocities to compute weighted horizontal wind components. The
MPA weights the winds by angular separation from orthogonality and by the distance between the lidar shots (Emmitt and Wood, 1988). The distance weight and angle weight are shown in Eqs. 8.1 and 8.2, respectively.

$$D_{wt} = 1 - (((\ _a * \ _c * COS(\ _a) - \ _f * \ _c * COS(\ _f))^2 + ((\ _a + 90) * \ _c - (\ _f + 90) * \ _c)^2)^{1/2})/D \qquad (8.1)$$

$$\ _{wt} = 1.0 - ((\ /2 - (ß_a - ß_f)) * (2/\ ))^4 \qquad (8.2)$$

where

$\ _{wt}$	angle shot pair weight
D_{wt}	distance shot pair weight
$\ _c$ (km/deg)	degrees to kilometers conversion factor at equator (111.11)
D (km)	diagonal distance across the grid area
$\ _a$ (deg)	longitude of the aft shot
$\ _f$ (deg)	longitude of the forward shot
$\ _a$ (deg)	latitude of the aft shot
$\ _f$ (deg)	latitude of the forward shot.

The U and V horizontal wind components are computed by Eqs. 8.3 and 8.4, respectively.

$$U = (VLOS_a / (COS(\ ) * COS(ß_a)) - (VLOS_f * TAN(ß_a)) /$$

(8.3) (COS( ) * SIN($ß_f$))/(1.0 - TAN($ß_a$)/TAN($ß_f$))

V = ($VLOS_a$ - U COS( ) COS($ß_a$))/(COS( ) SIN($ß_a$)) (8.4)

where

U (m/s)	U horizontal wind component
V (m/s)	V horizontal wind component
$VLOS_a$ (m/s)	line-of-sight lidar wind for aft shot
$VLOS_f$ (m/s)	line-of-sight lidar wind for forward shot
(deg)	elevation angle
$ß_a$ (deg)	scanner angle for aft shot
$ß_f$ (deg)	scanner angle for forward shot.

An analytical expression for the MPA Errors dependent upon shot separation was derived. The line-of-sight (LOS) velocities for two shots from different perspectives can be expressed as follows:

$VLOS_1 = (U_1 \cos{}_1 + V_1 \sin{}_1) \sin + W_1 \cos + N_1$ (8.5)

$VLOS_2 = (U_2 \cos{}_2 + V_2 \sin{}_2) \sin + W_2 \cos + N_2$ (8.6)

where:

$VLOS_i$	=	line-of-sight velocity for shot i
U_i	=	u component of the wind at location i
V_i	=	v component of the wind at location i
W_i	=	w component of the wind at location i
N_i	=	random noise
$_i$	=	azimuth for shot i from mathematic +x
	=	elevation angle from the nadir

Given the scarcity of shots, the following assumptions are made to solve for the horizontal wind components:

$u_1 = u_2$; $v_1 = v_2$; $w_1 = w_2 = 0.0$; $_2 = -{}_1$; $N_1 = N_2 = 0$

Thus:

$VLOS_1 = (U \cos{}_1 + V \sin{}_2) \sin$ (8.7)

$VLOS_2 = (U \cos{}_2 + V \sin{}_2) \sin$ (8.8)

To get two equations in two unknowns (u,v) we substitute $-{}_1$ for $_2$:

$VLOS_1 = (u \cos{}_1 + v \sin{}_1) \sin$

$VLOS_2 = (u \cos {}_1 - v \sin {}_1) \quad \sin$

Solving for u:

$VLOS_1 + VLOS_2 = 2 u \cos {}_1 \quad \sin$

(8.9)

$u = (VLOS_1 + VLOS_)/(2 \cos {}_1 \sin)$

Solving for v:

$VLOS_1 - VLOS_2 = 2 v \sin {}_1, \sin$

(8.10)

$v = (VLOS_1 - VLOS_2)/(2 \sin {}_1 \sin)$

However, if $U_2 \quad U_1$ and $V_2 \quad V_1$ then we make the following substitution into Eq. 8.8:

$U_2 = U_1 + u$

$V_2 = V_1 + v$

$VLOS_2 = (u \cos {}_2 + v \sin {}_2) \quad \sin + (u \cos {}_2 + v \sin {}_2) \qquad \sin$

(8.11)

$= (u \cos - v \sin) \quad \sin + (u \cos - v \sin) \sin$

Solving for u and v using Eqs. 8.7 and 8.11:

$u = (VLOS_1 + VLOS_2)/(2 \cos \sin) + \frac{u}{2} - \frac{v \tan}{2}$

$v = (VLOS_1 - VLOS_2)/(2 \sin \sin) - \frac{u}{2 \tan} + \frac{u}{2}$

The correct u and v values would be u + ½ u and v + ½ v. Therefore, the errors due to having different horizontal velocities at shot locations 1 and 2 are:

$U_{error} = - v \tan /2$ (8.12)

$V_{error} = - u/2 \tan$ (8.13)

If u and v are statistically non-zero (i.e., related to some coherent structure) then there will be a residual error regardless of the number of shots used in the velocity estimates. With Eqs. 8.12 and 8.13 we can compute the average u and v errors for a given shot pair spacing. For example ($_1$ = 301o):

$\underline{du} = 10^{-5} s^{-1}$

dx

$u = du/dx \quad x \qquad$ (x is $x(u_2) - x(u_1)$)

Let $x = 50$ km $u = .5$ m/s

$u_{error} = 0$

$v_{error} = .43 \text{ m s}^{-1}$

8.2 **Least Squares Model**

The least squares model performs a least squares fit to the lidar shots within a grid area to produce SNR weighted U and V horizontal wind components (Press et al., 1986).

8.3 **Observational Errors**

For a given number of lidar shots in a global grid area and at a one altitude, the LSMg computes the observational uncertainty in the global grid as follows

$$dn = 1 - \tan(ß_f)/\tan(ß_a) \qquad (8.14)$$

$$u_{12} = {}_{u}^{2} * (1/dn-.5)^2 \qquad (8.15)$$

$$u_{22} = {}_{v}^{2} * (\tan(ß_f)/dn)^2 \qquad (8.16)$$

$$u_{32} = {}_{w}^{2} * (\tan(\ )/\cos(ß_f)/dn)^2 \qquad (8.17)$$

$$u_{42} = 0 \qquad (8.18)$$

$$u_{52} = {}_{los}^{2} * (1/\cos(\ )/\sin(ß_a)*\tan(ß_f)/dn)^2 + \qquad (8.19)$$

$${}_{los}^{2} * (1/\cos(\ )/\cos(ß_f)/dn)^2$$

$${}_{u} = (u_{12} + u_{22} + u_{32} + u_{42} + u_{52})/(0.5 * nn) \qquad (8.20)$$

$${}_{u} = \text{sqrt}({}_{u}) \qquad (8.21)$$

where

nn	is the number of shots in the global grid,
	is the nadir scan angle (radians),
$ß_f$	is the forward scan angle (radians),
$ß_a$	is the aft scan angle (radians),
$_u$	is the cross track uncertainty (m/s),

v is the along track uncertainty (m/s),
w is the vertical velocity uncertainty (m/s),
los is the line-of-sight uncertainty (m/s),
u is the observational uncertainty (m/s).

9.0 Performance Display and Graphics Routines

Throughout the course of this contract, SWA has developed a series of graphics and display models for the LAWS simulations. The mapping algorithms allow the user to produce global (satellite view) or regional (mercator projection) graphics. The mapping models are the lidar shot location display, the lidar wind display model (plots input winds, output winds or both) and the T106/GADS color mapping model. A visual basic LAWS performance display model shown in Fig. A.41 was also developed which allows the user to plot line-of-sight and/or horizontal wind performance diagrams. The displays can be made for global performance or for a specific latitudinal band. The model has options for color or black and white.

APPENDIX B
1992 OBSERVING SYSTEM SIMULATION EXPERIMENTS

Under contract, SWA was tasked to conduct global OSSEs in co-operative efforts with GSFC and FSU to address orbit/design/configuration issues to assess the science related impacts of various shot management, signal processing and velocity computation algorithms. Five days of LAWS simulations were conducted using the following system configurations

20 joule/1.50 m/4.6 Hz system in a 525 km sunsynchronous orbit,
5 joule/0.75 m/4.6 Hz system in a 525 km sunsynchronous orbit,
0.1 joule/0.50 m/4.6 Hz system in a 200 km sunsynchronous orbit.

The simulated LAWS line-of-sight winds and horizontal wind components were delivered to NASA/GSFC (Atlas) and FSU (Krishnamurti) for assimilation.

One day global LAWS line-of-sight performance charts are shown in Figs. B.1-B.9 for the 20 J, 5 J and the 0.1 J laser system configurations for Nature Run days 16, 19 and 20 January, 1979. The charts give the percentage of time the lidar system can make useful measurements in terms of sufficient backscatter from aerosols, clouds and cirrus clouds. The charts also show the percentage of no returns due to clouds. All three system configurations indicate about 45-60% return from aerosols in the boundary layer. In the troposphere, where aerosols are represented as a background minimum, the 20 J system has 20% returns and the 5 J has less than 2%. The 0.1 J had no returns from aerosols. All three systems showed 0-10% returns from clouds.

One day of global LAWS U, V component (fixed area - 125 km X 125 km) performance charts are shown in Figs. B.10-B.20 for the 20 J, 5 J and the 0.1 J laser system configurations for Nature Run days 16, 19 and 20 January, 1979. The charts give the percentage of time that a lidar system can make good horizontal wind estimates for 5 categories ranging from 1 m/s accuracy to 5 m/s accuracy. Also shown is the percentage of time the system had an insufficient number of shots in the grid area to make a wind estimate. Figures B.10-B.12 give the LAWS performance as a function of the global grids that LAWS made an estimate in, whereas Figs. B.13-B.20 base the LAWS' performance upon the total number of global grids.

Figures B.10-B.12 demonstrate that all three systems can make good boundary layer wind estimates 15 to 20% of the time within 1 m/s accuracy and 30 to 50% of the time within 5 m/s accuracy. In the troposphere, the 20 J system's performance gives good wind estimates 15-25% of the time for 5 m/s or less. On the other hand, the 5 J and the 0.1 J performance is less than

10%. Figures B.13-B.20 illustrate the same relationship between the laser systems, but also considers the total area of the globe in the performance. The graphs show that LAWS could not make a wind estimate in 18% of the global grid areas due to lack of coverage.

In addition, we ran a LAWS simulation for a 0.2 joule/0.60 m/10 Hz system in a 300 km sunsynchronous orbit with a power conservation shot management algorithm (see Section 4.0, Appendix A). The power conservation shot management algorithm only takes lidar shots where groupings of orthogonal shots exist as shown in Fig. B.21. The line-of-sight and horizontal wind products were delivered to Atlas (NASA/GSFC) and to Krishnamurti (FSU) for assimilation.

Global LAWS line-of-sight performance charts for days 1 through 5 are shown in Figs. B.22-B.26. The charts give the percentage time the lidar system can make useful measurements in terms of sufficient backscatter from aerosols, clouds and cirrus clouds. The charts also show the percentage of no returns due to clouds.

Global LAWS U, V component (fixed area - 125 km X 125 km) performance charts for days 1 through 5 are shown in Figs. B.27-B.31. The charts give the percentage time that a lidar system can make good horizontal wind estimates for 5 categories ranging from 1 m/s accuracy to 5 m/s accuracy. Also shown is the percentage of time the system has an insufficient number of shots in the grid area to make a wind estimate. The coverage for the power conservation mode at a 300 km orbit is very small, but the quality of the wind products is extremely high. Figure B.32 shows the global coverage in 24 hours for the 0.2 J configuration.

An upgrade in the LAWS Simulation Model (LSM) - Global Version to include surface level LAWS winds added the option for a line-of-sight (LOS) performance profile with six atmospheric layers: surface, 1000 mb, low clouds, mid clouds, high clouds, and cloud free upper atmosphere. Figure B.33 shows the LOS observations for the 20 J LAWS configuration, where LAWS measurements due to high, middle and low clouds are seen 35%, 20% and 23% of the time, respectively. Conversations with M.J. Post and R. Menzies indicate enhanced backscatter levels due to volcanic eruptions. To bracket expected 20 J LAWS performance, we ran the LSM with aerosol backscatter increased by an order of magnitude higher than the background mode. Figure B.34 shows a potential increase in aerosol LOS winds by 70%, 40% and 50% in the cloud free upper atmosphere, high cloud layer and the mid cloud layer, respectively. Little or no impact is seen in the planetary boundary layer (PBL).

APPENDIX C
1991 OBSERVING SYSTEM SIMULATION EXPERIMENTS

Under contract, SWA was tasked to conduct global OSSEs in co-operative efforts with GSFC and FSU that are critical to the selection of orbit configurations for LAWS. We completed 5-day LAWS simulations with the following system configurations:

Perfect: 20 joule/1.5 m/4.6 Hz system in a 705 km-polar,
Perfect: 20 joule/1.5 m/4.6 Hz system in a 450 km-polar,
Perfect: 20 joule/1.5 m/4.6 Hz system in a 450 km-equatorial,
Perfect: 20 joule/1.5 m/4.6 Hz system in a 525 km-polar,
Perfect: 20 joule/1.5 m/4.6 Hz system in a 575 km-equatorial,
Realistic: 20 joule/1.5 m/4.6 Hz system in a 525 km-polar,
Realistic: 20 joule/1.5 m/4.6 Hz system in a 575 km-equatorial.

The simulated LAWS line-of-sight winds and horizontal wind components were delivered to NASA/GSFC (Atlas) and FSU (Krishnamurti) for assimilation.

Figures C.1-C.4 illustrate 12 hour coverage from the 450 km polar and equatorial perfect LAWS simulations. The perfect LAWS simulations considered cloud effects and terrain as the only mechanism to terminate a lidar shot.

Simulated LAWS winds and associated observational errors were computed for five ECMWF days and two satellite orbits (i.e., 575/55 and 525/98). These LSM runs considered the effects of clouds, cirrus clouds, aerosols, molecular attenuation, terrain, and sampling scale turbulence. Figures C.5-C.8 depict the 500 mb and the 1000 mb level Nature Run input winds for comparison with the LAWS winds shown in Figs. C-9-C.12. Cloud obscurations are noted by circled cross hairs. LAWS winds with observational errors greater than 25 m/s are not plotted.

APPENDIX D
GROUND-BASED DOPPLER LIDAR STUDIES

1.0 Signal Processing Tools and Data Display Models

Validation of the LMSg model has been a priority under the current contract. While GLOBE has improved the basis for modeled backscatter, the ground-based Doppler lidar at MSFC has provided some checks on assumed atmospheric variability parameterization and signal processing performance.

Many of the expectations for the performance of a space-based Doppler lidar are based upon the experienced gained from ground-based (and to a lesser extent airborne) systems. In particular, the LAWS concepts are derived from programs at NOAA and NASA - programs that have ground-based systems with pulse rates of 30-100 Hz and signal processing based upon the pulse-pair algorithm involving 3-50 shots per estimate. A space-based system will probably operate at < 20 Hz and each shot will be processed separately. SWA's experiments, therefore, have required that the single shot In-phase and Quadrature time series be recorded for post processing. Since the data volumes involved with single shot recording are much greater than the MSFC data system can handle, only 2 shots per record were recorded.

SWA has developed and refined a series of computer tools that are designed to process and analyze NASA/MSFC ground-based Doppler lidar polypulse-pair and single shot data. The models are the Pre-processor models, FFT Model, Sine Fit Model, PPI/UV and Range Display Models, Shot Pairing Model and the SNR/Consensus Model. Figure D.1 is a flow diagram that defines SWA's ground-based processing operations and conventions.

The NASA/MSFC ground-based lidar data pre-processor reads the raw NASA/MSFC data files providing data record counts to reformat the data into either VAD files (MSFC.VAD) for scanning angles or RANGE files (MSFC.RAN) for fixed scan angles. For single shot data, the models correct for the in-phase and quadrature data for offset and gain. For the scanning data sets, the models create line-of-sight wind velocity files (MSFC.PPI) for the PPI/UV Display Model and velocity/azimuth/signal (MSFC.VAS) files for the sine fit model.

After the polypulse-pair and single shot data sets are reformatted, the in-phase and quadrature single shot data is run through a FFT model to produce line-of-sight wind (LOS) velocities. Throughout this effort, we have evaluated several FFT models (MathLIB, MathCAD, numerical recipes, etc.) for Doppler lidar data processing. It became evident that a FORTRAN source code FFT model not only provides more flexibility (such as applying unity amplitude windowing) over a commercial off the shelf product, but is also faster. One VADS' worth of lidar

processing is approximately 7136 FFTs and takes 20 minutes on a 486 PC.

In an attempt to establish "ground truth" winds using poly-pulse data and in order to compute U/V horizontal wind components for the PPI/UV displays, SWA developed two sine wave fitting models that work with incomplete and extremely noisy data sets. The first model is a Fortran coded "brute force" model that tries to best fit the data by varying amplitude, phase and vertical offset for many combinations. Data filtering based on signal strength is used to remove wild data points. This model goes through many iterations to arrive at a best fit and is computationally expensive. Also, the sine fits have to be visually examined to evaluate the quality of the fits and to assess whether the applied filter was too weak or too strong. The second model was developed in Visual Basic and allows the user to quickly move the best sine fit around the data by changing amplitude, phase and offset. These sine fits are saved in MSFC.APO files. Figures D.2 and D.3 are examples of the Visual Basic Sine Fit Model. Figure D.2 shows a fit to unfiltered single shot data at range gate 9. The goodness of fit, 235.9, is not very good. Figure D.3 shows the same data with a signal filter used to remove weak pieces of information. The user was able to hone in the sine fit in seconds and found that the goodness of fit improved to 10.9.

The PPI/UV Display Model uses the MSFC.PPI line-of-sight velocity data file from the pre-processor model to produce a Plan Position Indicator plot (see Section 2.0). The model also uses the APO file from the Sine Fit Model to compute a vertical distribution of horizontal wind components. The RANGE Display Model uses the MSFC.RAN line-of-sight velocity data files from the pre-processor model to produce range verses elapsed time plots (see Section 2.0).

A space-based wind sounder will produce a pattern of shots that will provide generally bi-perspective samples within 100x100 km target areas. The angle between the two perspectives will vary depending upon the location of the target areas relative to the sub-satellite ground track. The wind estimate error will also vary with the geometry of the bi-perspective samples. The Ground-Based Shot Pairing Model processes the Doppler lidar data to yield shot combinations that are similar to those that will be achieved from space. Wind components are calculated for a range of bi-perspective shot pairs, which are compared to the true components "ground truth" derived from the sine fitted poly-pulsed pair VAD (see Section 2.0).

2.0 Ground-Based Studies

During the course of this effort, SWA cooperated with NASA/MSFC ground-based Doppler lidar engineers to produce a

series of lidar experiments to address issues such as system stability, amplitude averages, accuracy versus SNR, winds from cirrus decks, single shot pairing and associated errors. Table D.1 lists the NASA/MSFC data sets produced under this effort.

Table D.1. NASA/MSFC Ground-Based Doppler Lidar Data Sets

DATE	TYPE	AZM	ELV	#FILES	COMMENTS
07/10/92	SSHT	VAD	42	3	PARTLY CLOUDY
07/10/92	SSHT	VAD	74	1	PARTLY CLOUDY
07/10/92	PPP	VAD	42	1	PARTLY CLOUDY
07/10/92	PPP	FIX	3.8	1	MOUNTAIN
07/10/92	PPP	VAD	42	2	PARTLY CLOUDY

07/10/92	PPP	VAD	74	1	PARTLY CLOUDY
07/17/92	SSHT	FIX	3	1	MOUNTAIN
07/17/92	SSHT	FIX	3	2	SINE WAVE
07/17/92	SSHT	VAD	42	3	CLOUDY
07/17/92	SSHT	VAD	70	4	CLOUDY
07/17/92	PPP	FIX	3	1	MOUNTAIN
07/17/92	PPP	VAD	42	2	CLOUDY
07/17/92	PPP	VAD	70	2	CLOUDY
10/19/92	PPP	FIX	3	1	MOUNTAIN
10/19/92	PPP	FIX	90	1	CIRRUS
10/19/92	PPP	RHI	90	5	CIRRUS
02/27/93	SSHT	FIX	3.5	1	MOUNTAIN
02/27/93	SSHT	FIX	4.2	4	TOWER
02/27/93	SSHT	FIX	1	4	BLD 4200, PWR MODS
02/27/93	SSHT	FIX	3.5	1	MOUNTAIN
02/27/93	PPP	FIX	3.5	1	MOUNTAIN
02/27/93	PPP	FIX	4.2	4	TOWER
02/27/93	PPP	FIX	1	4	BLD 4200, PWR MODS
02/27/93	PPP	FIX	1	1	MOUNTAIN
03/01/93	SSHT	FIX	3.5	2	MOUNTAIN
03/01/93	SSHT	FIX	44	5	CLEAR, HIGH CLOUDS
03/01/93	SSHT	VAD	45	1	CLEAR, HIGH CLOUDS
03/01/93	PPP	FIX	3.5	3	MOUNTAIN
03/01/93	PPP	FIX	44	5	CLEAR, HIGH CLOUDS
03/01/93	PPP	VAD	45	1	CLEAR, HIGH CLOUDS
05/26/93	SSHT	FIX	4	1	MOUNTAIN
05/26/93	SSHT	FIX	1.8	1	BLG 4200
05/26/93	SSHT	FIX	1.8	2	NOISE
05/26/93	PPP	FIX	4	1	MOUNTAIN
05/26/93	PPP	FIX	1.8	1	BLG 4200
05/26/93	PPP	FIX	1.8	2	NOISE
05/27/93	SSHT	FIX	4	1	MOUNTAIN
05/27/93	SSHT	VAD	45	4	BKN CLOUDS, PWR MOD
05/27/93	PPP	FIX	4	1	MOUNTAIN
05/27/93	PPP	VAD	45	1	BKN CLOUDS, PWR MOD
05/28/93	SSHT	FIX	3.3	1	MOUNTAIN
05/28/93	SSHT	FIX	90	3	SUNNY, HIGH CLOUDS
05/28/93	PPP	FIX	3.3	1	MOUNTAIN
05/28/93	PPP	FIX	90	1	SUNNY, HIGH CLOUDS
06/03/93	SSHT	FIX	3.8	1	MOUNTAIN
06/03/93	SSHT	FIX	45	1	IN WIND DIRECTION
06/03/93	SSHT	FIX	45	1	IN CROSS WIND
06/03/93	PPP	FIX	3.8	1	MOUNTAIN
06/03/93	PPP	VAD	45	1	
06/03/93	PPP	FIX	45	1	IN WIND DIRECTION
06/03/93	PPP	FIX	45	1	IN CROSS WIND

07/07/93	SSHT	FIX	STEP	3	SUNNY
07/07/93	PPP	FIX	STEP	3	SUNNY

2.1 Frequency Stability

We examined three poly-pulse processed data sets to evaluate the frequency stability of the NASA/MSFC Doppler lidar wind returns from several stationary targets such as a mountain, building 4200 and a tower. Figure D.4 shows range versus time LOS wind velocity plots for returns off the building (~2.5 km). Figure D.5 shows the line-of-sight velocity, at range gate 8, as a function of time for returns off the same building. The LOS velocities varied with a of ~ .05 m s^{-1}.

2.2 SNR Studies

In processing the single shot data, VADS were used to identify a sequence of range gates where only noise existed. Figure D.6 shows the shot-by-shot average amplitudes of the signals in range gates 17 through 27. The average noise threshold is around 5.25 in arbitrary units. Figure D.7 is a velocity azimuth display (VAD) of the single shot data filtered along each shot by the average noise found in range gates 17-27. Using a VAD sine fitting model, vertical soundings were computed.

Single shot wind velocity estimates were analyzed per range gate foV_{sat} (km/s) satellite velocity,
V_{rot} (s^{-1}) scanner rotation,
ß' (deg) modified azimuth scan angle,
TIME (s) modified time.

The Atmosphere Generator Model (AGMg) creates Global Atmospheric Data Sets (GADS) that the LSMg use the accuracy of single shot estimates. The SNR value corresponding to all single shot estimates was also determined. From this information, the standard deviation of single shot estimates from poly-pulse pair "truth" was obtained for a range of SNR_N values. Figure D.8 shows the percentage of total single shot wind velocity estimates per SNR_N value which were less than the threshold values of 5 m/s, 4 m/s, 3 m/s, 2 m/s and 1 m/s. The trend in Fig. D.8 was as expected, i.e., better estimates were obtained at higher SNR_N. Additionally, Fig. D.8 shows that to obtain accurate single shot estimates to within 1 m/s 50% of the time an SNR_N value of about 2.5 dB is necessary. Likewise, to obtain accurate estimates to within 5 m/s 50% of the time an SNR_N value of approximately -2 dB is needed. This verified generally held expectations that an SNR (narrow band) of about 0 dB is the limit in obtaining "good" single shot estimates about 50% of the time. Finally, Fig. D.8 verifies another "rule of thumb" that confidence in single shot wind velocity estimates rapidly decreases for SNR_N values below 5 dB.

2.2 **Cirrus Studies**

We processed single shot and poly-pulse scanning data for July 10, 1992, 1623Z, which had a high thin cirrus cloud present. Each data set was noise filtered and least squares fit to a sine wave to establish "ground truth" winds. Figures D.__ and D.__ give examples of the sine fits, using the FORTRAN sine fitting model, for a series of range gates going from a strong SNR region to a weak SNR region for both unfiltered and filtered data. Figures D.__ and D.__ are selected poly-pulse and filtered single shot data for 1623 EST, VAD1, respectively. At range gate 5, the signal-to-noise (SNR) was very low for both data sets. At range gates 6-15 there were strong signals to make LOS estimates. However, at range gate 16, the single shot data showed loss of signal where the poly-pulse data (50 shot averages) was able to make a LOS measurement. Only noise was found from range gates 17-29. The cirrus cloud provided signal at range gates 30-32 as seen in Fig. D.__.

We selected the MSFC cirrus cloud data sets from July 17, 1992, 1623Z, to look at a time series of four sequential VADS, shown in Figs. D.__-D.__. The lidar wind information from the cirrus clouds appears to be very variable. A longer time series of lidar data is needed to examine this closer.

On May 28, 1993, the MSFC lidar was held in a fixed scan with an elevation angle at near zenith. At around 10 km, a thin cirrus deck was reported. Figures D.__-D.__ depict the poly-pulse lidar wind velocities and amplitudes for 400 seconds. As seen in Figs. D.__ and D.__, the cirrus return ends at around 370 seconds. Figure D.__ shows a time series at range gate 36 (cirrus deck) for the amplitude and wind velocity.

On October 10, 1992, the lidar was pointing straight up in a non-scanning mode at a thin cirrus deck. Figure D.__ is a range versus time wind velocity plot showing the cirrus deck at around the 41st range gate. Figure D.__ shows the vertical wind speeds as a function of range gate around the cirrus layer. For comparison, Fig. D.__ shows the vertical wind speeds for the mid-troposphere (strong SNR to weak SNR). The variance found in the mid-troposphere was approximately half that found in the cirrus layer.

2.3 **Shot Pairing**

As stated in Section 1.0 above, a space-based wind sounder will produce a pattern of shots that will provide bi-perspective samples within 100x100 km target areas. The ground-based data from July 10, 1992, 1623Z, was processed to yield shot combinations that are similar to those that will be achieved from space. Wind components were calculated for a range of bi-perspective shot pairs and compared to the true components derived from the sine fitted poly-pulse pair VAD. Figures D.__

and D.__ illustrate the distribution of errors for the U, V and total speed errors for unfiltered range gates 8-14 (high SNR) and for filtered range gates 30-31 (cirrus deck).

On May 27, 1993, the MSFC lidar ran a time series of scans (10 VADS) with an elevation angle of 45 degrees. Figure D.__ is a PPI diagram for the first VAD of poly-pulse data. We processed all ten VADS through the SWA ground-based Doppler lidar data processor. Each data set was noise filtered and least squares fit to a sine wave (see Section

__). Wind components were calculated for a range of bi-perspective shot pairs and compared to the true components derived from the sine fitted poly-pulsed pair VAD (see Section __). Figures D.__-D.__ illustrate the distribution of errors for the U, V and total speed errors for unfiltered range gates 10-15 (high SNR).

2.4 **Consensus Studies**

Of primary interest to the measurement of winds from space is the development of optimal signal processing algorithms. The Cramer-Rao lower bound is considered to be the ultimate target for algorithm performance. At this time, we expect to have an algorithm that is within 5-8 dB of that bound (Fig. 5 from John Anderson, University of Wisconsin). As part of our ground-based lidar research, we began assessing several processing algorithms. Since a consensus approach has been adopted for simulation studies, the ground-based data were processed to evaluate that technique. Some very preliminary results based upon 3 VADs of 1623 are shown in Fig. D.__. To note, is the much sharper fall-off of performance (percentage of estimates that meet consensus) than that expected from the curves in Fig. D.__.

On May 27, 1993, the MSFC lidar ran a time series of scans (10 vads) with an elevation angle of 45 degrees. In Section 2.2 above, we provide a discussion on how we processed the ten VADS of poly-pulsed data through the SWA ground-based Doppler lidar data processor to produce wind components and errors that are calculated for a range of bi-perspective shot pairs. We also used these data to populate our consensus algorithm to further evaluate the LAWS consensus signal processing technique.

(GDE input)

2.4 **Backscatter Inhomogeneity Feasibility Studies**

On July 7, 1993, the MSFC lidar performed a series of very low angle fixed scans varying from 2.5 degrees to 4 degrees to test if this configuration could be used to examine the backscatter and wind inhomogeneities in the boundary layer. Figures D.__-D.__ show the poly-pulsed wind velocities and amplitudes for elevations 2.5 degrees and 4 degrees, respectively. Strong lidar signals can be seen in the 7 to 10

range gate areas. Figures D.__-D.__ show a time series for both elevation angles at range gate 8 (high snr) and at the range gate where the signal hits the mountain.

APPENDIX E
ADDITIONAL STUDIES

1.0 Backscatter Comparison Between the Wavelengths 9.11 and 2.1 µm

The presentation, "Integration of Lowtran Into Global Circulation Models for Observing System Simulation Experiments" by S.A. Wood and G.D. Emmitt, compared the global relative backscatter performance of a 9.11 µm system to a 2.1 µm system.

A LAWS Simulation Model (LSM) has been developed to assess the potential impact of a space-based lidar wind sounder on global and regional features (see Appendix A). The LSM provides simulated winds for use in the Observing System Simulation Experiments (OSSEs) as described in Appendices B and C. The LSM contains "streamlined" versions of AFGL's LOWTRAN and FASCODE models that, coupled with the ECMWF model's atmospheric profile data, provides 1 X 1 estimates of global aerosol backscatter and molecular attenuation.

The LSM uses a unique 1 X 1 global LOWTRAN input data base that, given a location on the earth, provides a latitudinal location profile, haze model, coastal influence parameter, stratospheric profile and upper atmosphere profile. These LOWTRAN inputs along with global GCM inputs such as winds, relative humidity, temperature, pressure and clouds are used to compute aerosol backscatter and molecular attenuation.

Given the baseline LAWS wideband Signal-to-Noise (SNR) equation (described in Appendix A, Eq. 6.3), we can rewrite the SNR_w equation as follows

$$SNR_w = \frac{K}{R^2} * \;^{2} * ß(\;) * \;\; 2 \;\; (\;)\, dr$$

where

K = system variables
 = wavelength
ß ( ) = aerosol backscatter
 ( ) = attenuation
R = range

and grouping all system variables in the constant K, we simplify the equation to atmospheric variables and wavelength. For our discussion, we multiplied all optical properties by ^{2}, thus our performance comparison is in terms of the optical properties and wavelength and any additional enhancement of the SNR must come from the system (K) itself.

Figures E.1 and E.2 show the global unattenuated aerosol backscatter for 9.11 µm and 2.1 µm, respectively. Since our aerosol backscatter model is largely driven by relative humidity, Fig. E.3 gives the corresponding global relative humidity field to highlight the high relative humidity regions (i.e., high backscatter). This study shows that at high relative humidity regions, the 2.1 µm had distinct advantages over the 9.11 µm system. However, no advantage was evident at low relative humidities and in some dry regions the 2.1 µm system actually performed worse than the 9.11 µm. Figures E.4 and E.5 show the 9.11 µm and 2.1 µm attenuated aerosol backscatter for a given LAWS baseline configuration, respectively. Likewise, the performance of the 9.11 µm system is much better than the 2.1 µm system in most of the northern hemisphere. Whereas, in most of the "wetter" southern hemisphere, the 2.1 µm is better than 9.11 µm.

Since LSM cloud fields are also driven by the moisture fields, we looked at high ß regions where total cloud cover did not exist. Figure E.6 shows all 9.11 µm attenuated backscatter greater than 2.5 e^{-6} m^{-1} sr^{-1} for integrated cloud cover less than 90%. Figure E.7 shows all integrated cloud cover to the surface. Large areas of high backscatter regions are found in the northern hemisphere even though there is a large amount of cloudiness present. For most of the southern hemisphere, either the backscatter was weak or cloud cover was dominate.

In conclusion, any significant advantages of a 9.11 µm system over a 2 µm system, or visa versa, over the obvious wavelength considerations must come from lidar system parameters such as power, telescope diameter or system constants.

APPENDIX F
HDF FORMAT FOR GLOBE DATA

The GLOBE database project involves the development of a system for archiving, distributing and retrieving aerosol data obtained from seven different sites. Over the past several months, we have worked on completing the archival strategy for one site (GSFC) with the goal of developing a storage format that will be useful for data from all of the sites. We focussed on several factors when creating the format for storing data from GSFC. First, data storage needed to be both efficient and consistent. Our goal was to maintain consistency in the format so that it could be used for all sites, while making the best possible use of disk space. Second, we developed a storage format from which data could easily be retrieved, both by internal project scientists and by external users. This involved using reliable and freely available database management tools and maintaining information about the data as part of the storage format.

The GLOBE database project scientists at MSFC had already begun to develop a system of data storage using the database management system known as the Hierarchical

Data Format (HDF). HDF is distributed by the National Center for Supercomputing Applications (NCSA) and is freely available to users. We chose to continue using HDF to format the data because it met the criteria outlined above, and it is also the data format distributed by the DAACs.

HDF consists of a set of routines that are callable either from C or Fortran. These routines allow for the storage of many data blocks in one file. They also allow the user to store meta-data referring to either individual data blocks or the file as a whole. Thus, a wide range of meta-data can be saved, from information about variables themselves (e.g., their units of measure) to information about the creation of the dataset as a whole (e.g., the type of equipment used for data collection). Although the original computer programming for the storage of GLOBE data was done in Fortran, we agreed to re-write these programs in C as part of the project, so that all programs would be written in one language. C is the language of choice when using HDF because it allows for more flexibility in formatting and because some new HDF functions are not callable from Fortran.

The first phases of work on the GLOBE database project began in early March after we received and configured a Silicon Graphics workstation. HDF (version 3.2R3) was downloaded from NCSA and compiled on our computer. It was tested using examples provided by NCSA and with small sample datasets created at SWA. After this initial testing period, we downloaded two sample datasets which were provided by Dean Cutten of MSFC. These datasets contained similar information but were from two different sites (GSFC and JPL). They were representative of the

different formats in which the data are received. We focussed on the GSFC data as our test case for developing a complete system for formatting the data.

We continued our initial testing process by creating a small sample HDF data file from the GSFC data provided by MSFC. The practicality of creatingV_{sat} (km/s) satellite velocity,

V_{rot} (s^{-1}) scanner rotation,

ß' (deg) modified azimuth scan angle,

TIME (s) modified time.

5.0 Atmosphere Generator Model for the Global Atmospheric Data Set

The Atmosphere Generator Model (AGMg) creates Global Atmospheric Data Sets (GADS) that the LSMg usgrams provided by Dean Cutten for formatting the GSFC data. Our goal was to create new C programs that would incorporate all of the information saved in the old programs while modifying the database structure and adding some information that was not previously saved. Toward this end, we began with a sample dataset that contained only one block of data collected at GSFC. This data block was used to test all aspects of data storage, and the program was expanded to store many data blocks only after the test block was successfully formatted in an HDF data file.

The HDF format that was developed for storing each data block consists of several components. First, the data are stored in an unmodified form as a matrix of data points. Second, information about the data block is stored as a separate matrix associated with the data matrix. This matrix contains the numeric header information, and it consists of the following variables:

Number of layers in the data block
Lidar pointing up (1) or down (2)
Start Year (indicating the beginning of data collection)
Start Month
Start Gregorian Day
Start Julian Day
Start Hour
Start Minute
Start Second
End Year (indicating the end of data collection)
End Month
End Gregorian Day
End Julian Day
End Hour
End Minute
End Second
Start Latitude
Start Longitude

Start Altitude
End Latitude
End Longitude
End Altitude
Maximum Latitude
Maximum Longitude
Maximum Altitude
Minimum Latitude
Minimum Longitude
Minimum Altitude

Our program obtained some of this information directly from the raw data files provided by GSFC. For instance, the direction in which the lidar is pointing is available in the original data file. Other variables are created or modified by the program before being stored in HDF. For example, our program provides routines for converting Julian days to Gregorian days and vice versa. Also, the position information stored as part of the numeric header is interpolated by the program using data obtained from DADS data files provided by Dean Cutten at MSFC. This was done to maintain consistency between sites. Although position information is provided by GSFC in their data files, other sites do not provide this information. It was felt that the use of one source for latitude, longitude and altitude information would be more consistent than relying on multiple sources depending on the original site of data collection.

In addition to the numeric header, two other pieces of information are stored with each block of data in an HDF file. First, a short block label is assigned to each block. This label provides information about the start year, Julian day, hour, minute and second of data collection (formatted as yydddhhmmss). Although these data are already available in the numeric header matrix, the short label provides easy access for a user who might wish to retrieve a data block collected at a particular time. Second, a block descriptor is also associated with each block of data. This consists of numeric information that is specific to GSFC (that is, it does not generalize across sites) and thus was not included as part of the numeric header for the data block.

Several categories of meta-data stored in each HDF data file did not pertain to individual data blocks but were instead associated with all blocks of data stored in the file. These types of information were stored only once in each data file. They fall into four different classifications. First, every HDF data file contains a file label. This is a short text string that consists of the project name (GLOBE I or GLOBE II), the flight number, the institution from which the data were received, and the instrument that was used for data collection. Second, a longer file descriptor is saved with each HDF file. This contains information about whether the flight was transit or local, the start and end locations for the flight (saved as city names rather than latitude and longitude coordinates), and any special comments regarding the flight (such as its purpose). The information stored in the file descriptor is obtained by user input during the archival procedure.

The third type of data stored at the file level consists of a number of text header fields that contain information about the flight and the scientists involved in the flight. The following fields are saved as part of the text header:

Task Type
Experiment Name
Experiment Sub-Name
Platform
Instrument Type
Instrument Name
Processing Level
PI Name
PI Phone
PI E-Mail Address
PI Fax
Other Researcher's Name
Other Researcher's Phone
Other Researcher's E-Mail Address

Other Researcher's Fax
Institution

These fields represent information that might be of use during data retrieval, when a user might want to search data files to find particular instruments or experiments or to contact the PI with questions about the data. All of the fields are obtained from user input during the archival process with the exception of the institution. This field refers to the site from which data are received. It is specific to the data being processed, and it is assigned by the program.

Finally, each HDF file contains documentation stored on the file level that describes the process by which the file was created so that the history of each dataset is preserved. This documentation consists of the following variables:

HDF File Creation Date
HDF File Creation Time
Program Name and Version used for HDF File Creation
Processing Steps
Remarks

Information about 'processing steps' include details about when and in what form the data were received from the original source, the types of transformations that were performed on the data, and the process by which the data were stored in an HDF format. This differs from the 'remarks' section in which information about the

experiment itself is stored (e.g., details about instrument performance, weather problems, etc). This type of documentation was saved for two purposes. First, it provides a user with useful information about the condition and storage of the data. Second, it documents every transformation that has occurred in the data so that earlier versions of the data can be recreated if necessary.

The HDF format outlined above represents a departure from the format previously used by the GLOBE database project in several ways. First, this format allows all information (data and meta-data) to be stored in one HDF file. The previous format used two separate files to store these types of information. Second, the current format stores more meta-data than
DeTd$ea"
"#IJ L

HT B- D.HDJ42*@BT
!P PT Ha -QQ~Q"AEDB Z`V-Z0"F42 @42J0"
5AF @X