PHAiRS

Public Health Applications in Remote Sensing

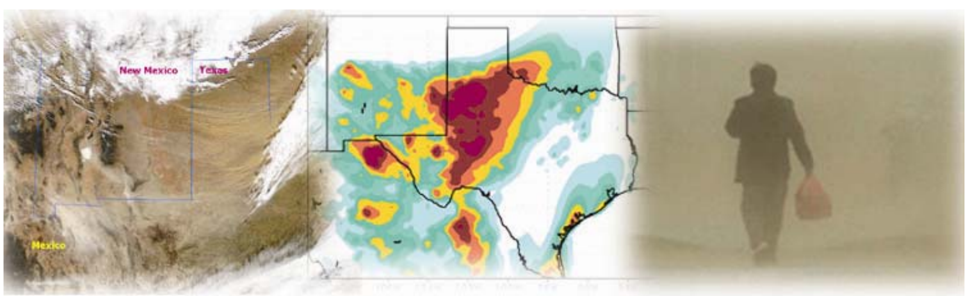
Verification and Validation Report

Agreement NNSO4AA19A

Report submitted to:

John Haynes, Program Manager, NASA Headquarters
and
Terry McPherson, Stennis Space Center

By
Dr. Stanley A. Morain, Principal Investigator, University of New Mexico
and
Dr. William A. Sprigg, Co-Principal Investigator, University of Arizona



September 30, 2007

Contributing Project Members:

Dr. Karl Benedict, Amelia Budge, Thomas Budge, and Dr. William Hudspeth – University of New Mexico

Dr. Brian Barbaris, Dr. Dazhong Yin, and Patrick Shaw – University of Arizona

Executive Summary

Cardiovascular and respiratory diseases increase in populations exposed to airborne mineral dust. Sand and dust storms that entrain and carry particles to unsuspecting populations are also a hazard to air and ground transportation, spread bacteria and toxic materials mixed with the soil, and affect weather and climate through radiation and condensation processes.

NASA Earth Science support for “Public Health Applications in Remote Sensing” demonstrates that using MODIS products greatly improves sand and dust storm simulations and forecasts. The current project also helps narrow the field of potential remote sensing products that could map the three-dimensional characteristics of dust clouds over the landscape. Current capabilities developed through PHAiRS in simulating and forecasting airborne particulates have led public health and air quality offices in the states of New Mexico and Arizona to reevaluate their assessment of the emerging technology. Today, these offices participate actively in PHAiRS, testing and designing products and providing a beacon for future research.

The tools developed herein are mindful of the needs by public health services. Every component is based on current obligations of government. The simulation and forecast models are driven by the operational weather forecast models of the U.S. National Weather Service. Weather data and analyses are provided in real time through the national and international operational services of the World Weather Watch, the European Centre for Medium-Range Weather Forecasting, and the U.S. National Weather Service. Remote sensing and data products from the A-Train are provided by NASA and partners. State agencies responsible for monitoring the region’s airshed or for public health surveillance and warnings, and other departments of environment, air quality, and health services, provide the ultimate test of new applications for remote sensing products assimilated into air quality models.

PHAiRS products are being validated against several measures of airborne particulate matter obtained from both satellite sensors and ground-based monitors and from *in-situ* samplers of PM_{10} and $PM_{2.5.}$ Dust model outputs are evaluated also against observed and modeled weather variables. These verification and validation (V&V) studies show marked improvement of PM_{10} and $PM_{2.5}$ concentration simulations and forecasts over the first three-and-a-half years of the project, principally due to improvement of source identification using advanced MODIS products. These studies also point out highly desirable future directions for verification (more speciation of particles and use of ground- and space-based LIDAR) and of study and product development (higher space and time resolution of surface landscape characteristics and model product outputs).

Table of Contents

1.0 Introduction

1.1 Importance of dust in human health

Everyone's health is impacted by exposures to microscopic minerals, chemical particulates, by organisms bonded to air particles, and by toxic gases (Figure 1). Many of these constituents are infectious and can become contagious; others are patient-specific. For example, ozone (O_3) is a molecule that can lead to chronic asthma; viruses lead to influenza; bacteria lead to intestinal problems; and pollen leads to hay fever. Dust and smoke particles and industrial emissions in the PM_{10} to $PM_{2.5}$ micron (µm) range contribute directly to respiratory health responses and serve as carriers for respirable viruses and bacteria (Kuehn, 2006).

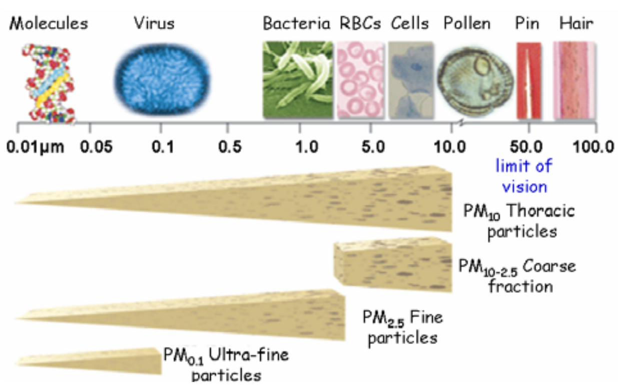


Figure 1. Particulate size distribution and related biophysical impacts. Source: Kaiser, 2005 Science 307, p. 1859.

One of the challenges for integrating satellite Earth observations into human health practice is to demonstrate that these data improve model predictions of dust levels that could trigger respiratory responses. The body of medical and epidemiological knowledge linking dust and smoke to health responses is growing rapidly (Pope, 1989, 2004; Schwartz and Dockery, 1992; Dockery et al., 1993; Pope et al., 1995; Griffin, 2007; National Research Council and Institute of Medicine, 2007). Through these linkages, it is increasingly clear to science and government that satellite observations can play a prominent role in forecasting short term weather episodes, and longer-term environmental changes that cycle over several human generations. Earth system scientists are modeling complex biological, chemical, and physical processes at the ecosystem level, and finding quantitative measures for tracking ecosystem health over regional domains.

Another challenge is for medical science to extract from this knowledge the consequent flow of pathogens and chemicals through airborne mechanisms, and to translate findings into actionable human health interventions for populations at risk. This challenge implies adding health-care professionals into efforts that merge environmental surveillance with human health surveillance. Effective public health surveillance requires an appreciation of natural processes that impact environments and that could impose secondary impacts on exposed populations. Rewards will be realized when health care providers and health authorities are included in collaborative efforts with Earth scientists.

While the medical community recognizes the adverse effects of PM_{10} and $PM_{2.5}$ in patients with respiratory conditions (Pope, 2004), they lack reliable information for forecasting dust storms so that public alerts can be issued. Respiratory diseases and syndromes typically are monitored by surveillance systems consisting of electronic databases into which data are entered and accessed by doctors and clinicians. But these systems do not provide enough information to issue public warnings in advance of a dust event. This technological gap is accented by the high numbers of deaths that could be exacerbated by dust and aerosols. Table 1 lists the world's five leading causes of death. Three of these include some deaths that are actually caused by declining air quality and/or dust episodes, but for which there is no clinical confirmation of dust involvement.

Table 1. Leading causes of death, worldwide (estimated), 2002. Deaths reported in cardiovascular, infectious and parasitic, and chronic lung categories include those actually induced by dust. Source: Centers for Disease Control and Prevention, 2005.

Cause of Death	*Est. # (%) of Deaths*
Cardiovascular	16.73M (29%)
Infectious & Parasitic	14.86M (26%)
Malignant neoplasms	7.121M (12%)
Violence/injuries/ accidents/suicides	5.168M (9%)
Chronic lung	3.02M (6%)

Box 1 summarizes a few of the social and economic costs related to asthma and myocardial infarction (MI) in the United States.

Box 1: Health Costs and Respirable Particulates

Economically, there is ample evidence that respirable particulates result in costly health effects. Asthma and MI are among these. Asthma is a progressive disease that afflicted 20M Americans in 2003 (American Lung Association, 2005). It is a chronic disease, especially in arid and semi-arid areas of the U.S. (ca. 25% of the domestic land area). Between 1980 and 1994, the prevalence of asthma in the U.S. increased 75%; in children under 5, it increased 160%. In 2003 there were 12.7M physician office visits and 1.2M outpatient department visits related to asthma (CDC, 1998). Direct health care costs currently exceed $11.5B annually, including $5B in prescription drugs. Indirect costs (lost productivity) add another $4.6B (Myers, 2006). Annual treatment costs in 2003 were over $4,900 per asthmatic. From the health and health care cost perspectives, there is a strong argument for forecasting outdoor dust and ozone environments based on time series Earth observations of dust episodes.

Medically, the epidemiology connecting declining air quality and respiratory diseases in desert regions is poorly understood; but, patterns of rising health care costs are agreed in the health communities-of-practice to be associated with rising levels of atmospheric contaminants. Economic studies in the environment and health sector provide adequate stimulus for investing in quantitative environmental measurements that reduce medical care costs and improve air quality that someday will reduce chronic diseases (cf, Ackerman, 2002; Landrigan et al., 2002; Pear, 2003; Massey and Ackerman, 2003; Jerrett et al., 2003; and, Davies, 2005).

1.2 PHAiRS integrated system solution

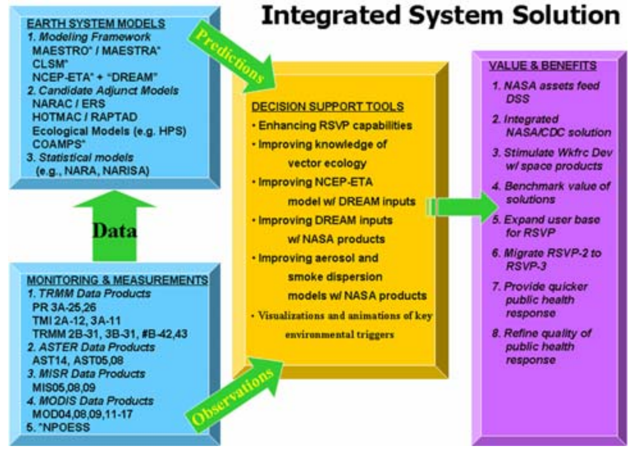


Figure 2. Plan for inserting public health into NASA's integrated system solution (ISS).This is the solution proposed by the PHAiRS project team for verification and validation.

PHAiRS has three parallel thrusts. The ***first*** assimilates satellite observations from MODIS Terra and other sources into the Dust Regional Atmospheric Model (DREAM). DREAM is nested within the National Centers for Environmental Prediction (NCEP/eta) weather forecasting model, which is a numerical model driven by both satellite and *in-situ* atmospheric measurements. The aim of PHAiRS is to: (a) verify that advanced satellite image data from NASA sensors can replace terrain parameters from traditional non-satellite sources, or from earlier (coarser resolution) satellite sources; and, (b) validate that parameter replacements lead to more reliable model forecasts of dust episodes.

The ***second*** is to optimize DREAM model outputs by iterating model inputs with a variety of satellite products and assessing incremental improvements. The questions of greatest interest are: (a) how well, and with what degree of sensitivity, can NCEP/eta combined with DREAM forecast dust lifted from a landscape? (b) how well can this combined model simulate the speed and direction of moving dust clouds?; (c) can medically sound evidence be generated that couples dust episodes to documented respiratory health responses at the population level?; and (d), can areas affected by dust clouds be forecasted in a timely fashion to alert health officials and populations at risk?

The ***third*** thrust is to establish collaborative relations with public health authorities to test whether there are statistically valid relationships between dust episodes and increased respiratory complaints. This is a difficult task in the United States because public health authorities are distributed throughout all levels of government, and because standardized record keeping is not mandatory within or between these levels. Furthermore, patient confidentiality makes it impossible to know the geospatial coordinates behind any given record.

Ultimately, the goal of PHAiRS is to improve public health decision support systems that can evolve toward operational status for the next generation of space-based sensing. The National Polar-orbiting Environmental Satellite System (NPOESS) is scheduled for launch in the 2010 timeframe. It will consist of several platforms carrying operational versions of NASA's current experimental sensors. The PHAiRS project is helping to build the scientific and technological underpinnings of these near-future capabilities, and testing them with appropriate public health user communities.

1.3 Activities leading to V&V report

Table 2. Milestones in PHAiRS tasks.

Activity	***Calendar Quarters Beginning March 2003***															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Baseline pre-PHAiRS DREAM performance	✶	✶	✶	▲ (red)												
Assimilate EO products into DREAM				✶	✶	✶	✶	▲ (red)								
V&V outputs against dust observations			✶	▲ (red)	✶	✶	✶	▲ (red)	✶	✶	✶	▲ (red)	✶	✶ (yellow)	✶ (yellow)	▲ (yellow)
Iterate model runs w/variety of EO data; Initial Benchmark					✶	✶	✶	▲ (red)	✶	▲ (orange)	✶	✶				
Develop Web client/Start AIR-Now, etc.					✶	✶	✶	✶	✶	✶	✶	▲ (red)	✶	✶ (yellow)	✶ (yellow)	▲ (yellow)
Statistical analyses & forecasts								✶	✶	✶	✶	▲ (red)	✶	✶ (yellow)	✶ (yellow)	▲ (yellow)
Engage user communities	✶	✶	✶	▲ (red)	✶	✶	✶	▲ (red)	✶	✶	✶	▲ (red)	✶	✶ (yellow)	✶ (yellow)	▲ (yellow)
User training workshops													✶	✶ (yellow)	✶ (yellow)	▲ (yellow)
Submit V&V Report													✶	▲ (orange)		
Final Benchmark Report	……	……	……	……	……	……	……	……	……	……	……	……	✶ (yellow)	✶ (yellow)	✶ (yellow)	▲ (yellow)

▲ (red) =Annual Reviews; ▲ (orange) =Initial Benchmark Report and V&V Report; ✶ (yellow) Projected;

▲ (yellow) = 4th Projected Annual Review; Final Review in 2008 not shown but scheduled for March 2008.

1.4 Initial benchmark results from 2005

The initial benchmark report to NASA (Morain and Sprigg, 2005) showed the fundamental socioeconomic and political importance that dust storms play in human health, and how Earth observations will play significant roles in future public health services.

When the operational National Weather Service forecast model was modified to accommodate a dust entrainment process (i.e., DREAM), meteorological variables were well simulated and predicted, especially *after* experimental satellite observations were assimilated.

Model comparisons showed that modeled meteorological fields, both surface and 500 millibar levels, were in agreement with measured observations. The modeled vertical profiles of wind speed, wind direction, temperature, and specific humidity matched the observed profiles. Statistical evaluation of the modeled and observed surface winds and temperatures showed the model performed reasonably well in reproducing the measured values.

The MOD12 barren ground assimilation had a much larger influence on modeled dust concentrations than on meteorological fields. The peak hour correlation was least affected by the change. However, major gains were made in modeling the magnitude and duration of near-surface high dust concentrations. The enhanced model predicted accurately the order of magnitude of the dust storm event at almost all locations in the model domain.

2.0 V&V Methods and Results

In the integrated system solution, verification and validation are tied to uncertainties in the data and product inputs, assumptions in the processing steps, and interpretation of the outputs. Some of these are specific and measurable; others are known but not measurable; and still others are suspected and require further definition. Aerosol optical depth/thickness (AOD/AOT) is an example of the latter. It does not appear that these data will be useful in model verification other than being another qualitative measure for comparison. However, if work at The University of Arizona with speciated $PM_{2.5}$ is successful, these data may be more correctly compared to model output.

2.1 Pre-PHAiRS DREAM performance

The pre-PHAiRS baseline DREAM/eta model has been verified and validated over the Mediterranean and North Africa by Nickovic et al. (2001); Nickovic et al. (2004); and Perez et al. (2006). The Nickovic, et al. papers made qualitative comparisons of the horizontal plume of a Saharan dust event. Perez et al. compared observations of a 17-day Saharan dust event that affected the western Mediterranean in June 2002. Intensive LIDAR observations at Barcelona (Spain) and sun-photometer data from two stations located along the dust plume (El Arenosillo, Spain; Avignon, France) were used to examine *vertical structure* and optical properties, and to evaluate DREAM performance. Evaluations were performed also to show the dust *horizontal spread* and *vertical structure* simulated by DREAM, as observed by SeaWiFS and as measured by LIDAR and sun photometers in the region. Figures 3 through 7 show the horizontal spread and vertical structure of dust plumes originating in the Sahara. In Figure 3 images from SeaWiFS were compared visually to DREAM dust loading maps. Agreement between the two patterns is encouraging. There is also encouraging agreement between the modeled vertical structure and the observed vertical profiles over Barcelona (Figure 4).

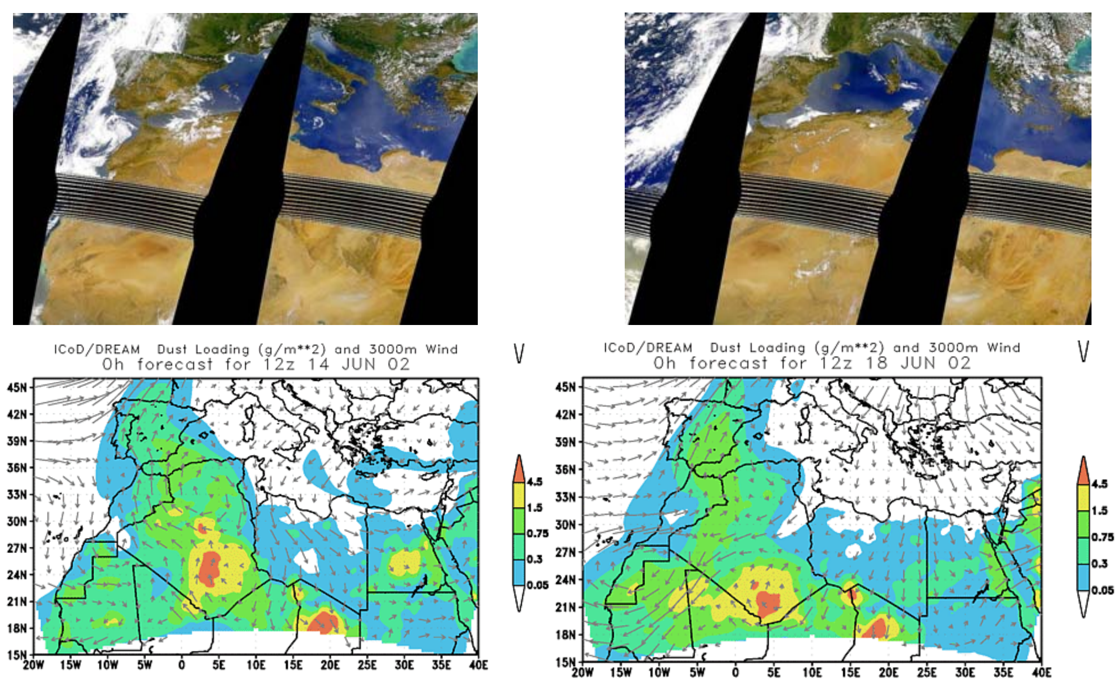


Figure 3. Horizontal spread of a Saharan dust storm. SeaWiFS Images and modeled dust loading and winds at 3000m. (Left, top). SeaWiFS image; (left, bottom) DREAM output for 14 June, 2002. (Right, top) SeaWiFS image; (right, bottom) DREAM output for 18 June, 2002 (Perez et al., 2006).

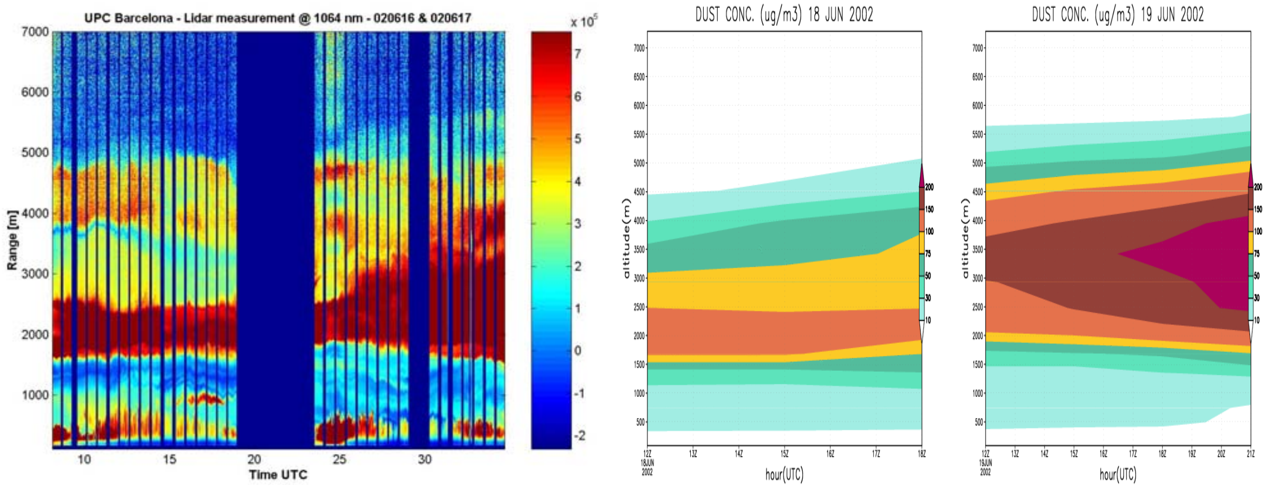


Figure 4. Vertical structure of a Saharan dust storm over Barcelona, Spain, 18-19 June, 2002. (Left): LIDAR measurements. Dark blue columns indicate no measurements on the range corrected 1064nm signal (arbitrary units; temporal resolution is 60 sec.); (Right) DREAM modelled vertical dust concentration (Perez et al., 2006).

Figure 5 compares modeled and observed Aerosol Optical Depth (AOD) over Arenosillo, Spain. Figure 6 shows LIDAR vertical profiles of measured extinction coefficients at 1064nm and 532nm compared to modeled results over Barcelona, Spain. In general, the modeled profiles from three parameterizations designated G8 and D8 are in good agreement with observations, but show a tendency to over-predict in the upper levels of the dust plume. Note that LIDAR profiles may contain error-bars of 30 percent due to the assumption of a constant LIDAR ratio in the profile.

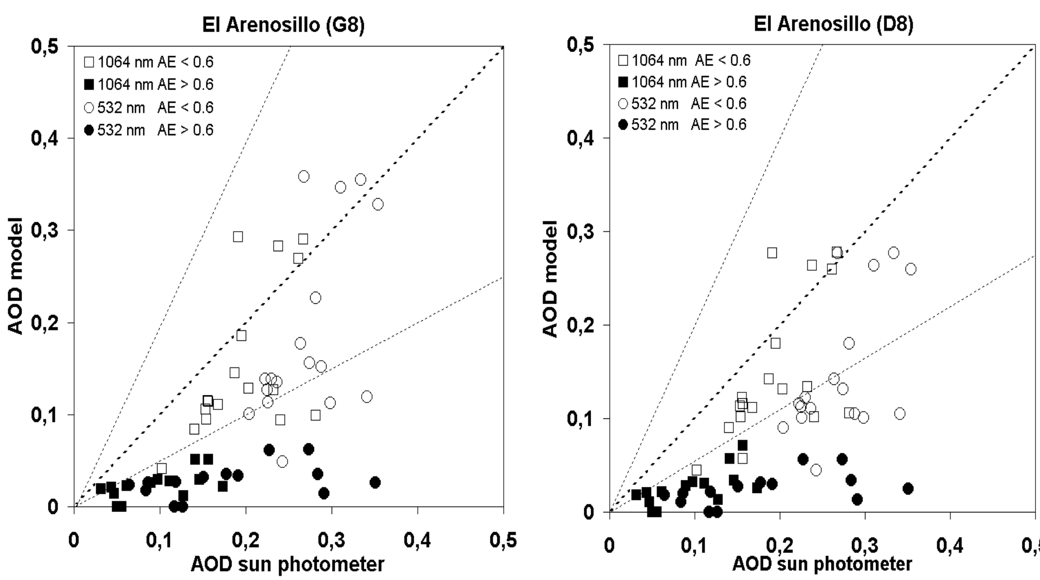


Figure 5. Modeled vs observed Aerosol Optical Depth (AOD) for G8 (left) and D8 (right) at El Arenosillo, Spain (Perez et al., 2006).

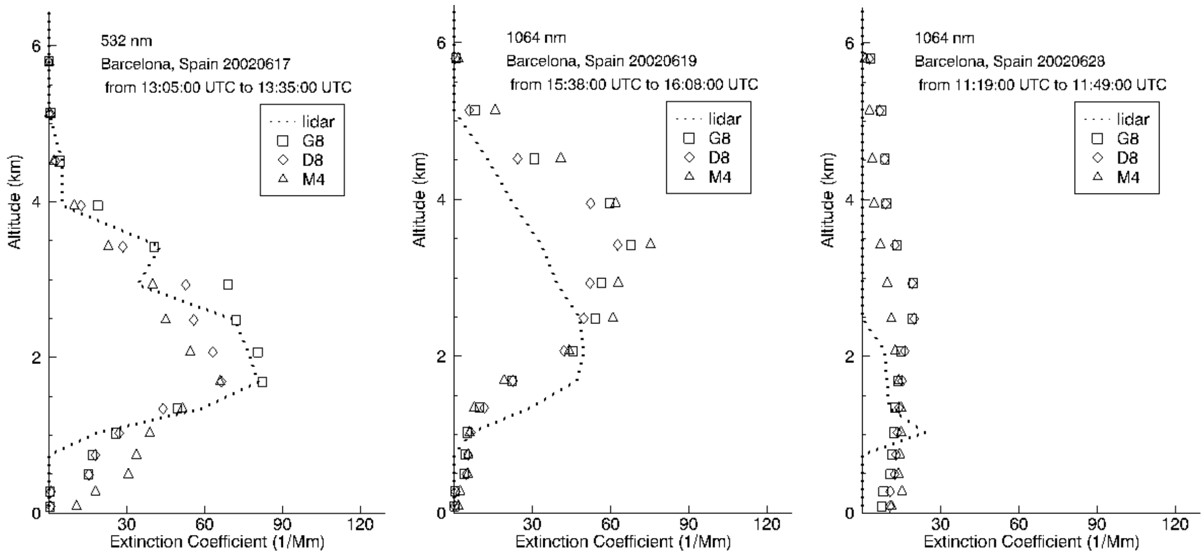


Figure 6. Comparisons of modeled and observed vertical profiles of the extinction coefficient for M4, G8 and D8 over Barcelona, Spain. (left) 532nm on June 17, 2002 between 13:00 & 13:35 hrs; (center) 1064nm on June 19, 2002 between 15:38 & 16:08 hrs; (right) 1064nm on June 28, 2002 between 11:19 & 11:49 hrs UTC (Perez et al., 2006).

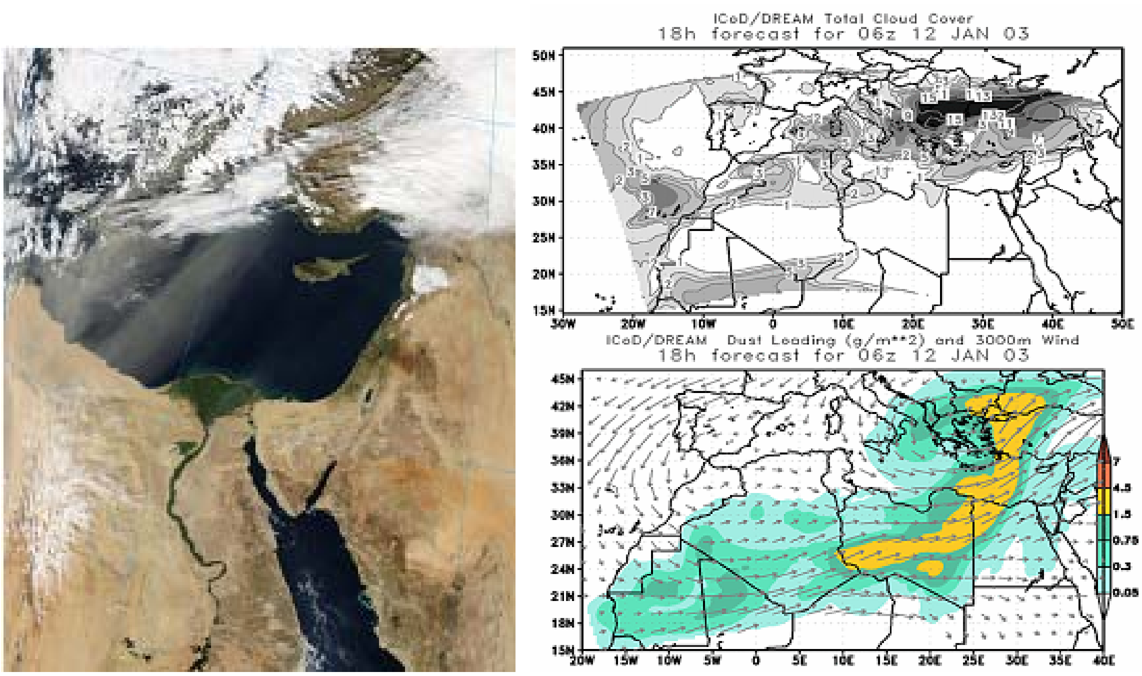


Figure 7. Mediterranean dust transport on 12 January 2003. (Left), Satellite image; (right), DREAM output (Nickovic et al., 2004).

Passive satellite sensors only show horizontal 2-dimensional features of dust plumes that often remain undetected over continents because sensors cannot distinguish easily between the color of atmospheric aerosols and surface background reflectance especially in arid and semi-arid environments. Sun-photometers deliver column integrated results with no distinction of layered aerosols or particulates. On the other hand, deposition or surface concentration data involve close-to-ground characteristics of the dust process. Ground based LIDAR complements other measurements and depicts dust structure to allow vertical model validation, but deposition or air sampling that yields quantitative measurements of the airborne dust is the most difficult test for model validation. This type of validation has been performed by the PHAiRS team for this V&V report.

2.2 Performance-PHAiRS domain

Verification and validation of DREAM performance over the PHAiRS domain are based on statistical measures that compare surface level dust (PM_{10} and $PM_{2.5}$) concentrations recorded at particulate air quality sampling sites. Few have attempted such rigorous tests.

DREAM was adapted for the PHAiRS model domain centered at (109ºW, 35ºN). It covers the Southwest US and its surrounding areas (Morain and Sprigg, 2005). Table 3 lists the data sets and general properties upon which the model is based.

Table 3. Data sets employed to V&V baseline PHAIRS DREAM/eta domain performance. See Nickovic et al., 2001; Zobler, 1986; and Cosby et al., 1986 for details.

Data Set	Purpose/Properties
ECWMF medium-range weather forecast	Initial & boundary conditions; Res. = 1°
NCAR monthly SST	Sea surface temp.; Res. = 1°
USGS terrain data	Res. = 1km
Olsen Wld Ecosysts[1]	Land cover; Res. = 10min.; Dust categories = 8, 50, 51, 52
FAO Wld Soil Map[2]	Res. = 2min.; 134 categories reduced to Zobler/Cosby categories for soil texture

As an initial test of the model's performance, modeled meteorological fields were evaluated against measurements and analyses obtained from surface synoptic, surface Meteorological Aerodrome Report (METAR), and upper-air radiosonde reports. The modeled dust field patterns and dust concentrations were compared with satellite images, measured visibility distributions, and the surface $PM_{2.5}$ and PM_{10} observations from Texas Commission on Environmental Quality (TCEQ) and the US Environmental Protection Agency (EPA) Air Quality System (AQS). Graphical measures, such as pattern comparison, site against site time series vertical profile comparison, and statistical metrics, were used (Yin et al., 2005).

Table 4 lists the performance statistics for modeled surface wind and temperature for a December 15-17, 2003 dust storm over New Mexico and Texas. The modeled wind, temperature, and humidity profiles were verified against sounding data. The modeled meteorological fields were compared against analyses using observational data. These comparisons showed that DREAM performed well in forecasting meteorological parameters.

Table 4. Performance statistics of modeled surface wind and temperature.

Metrics	Wind Sp (N=31967)	Wind Dir (N=31968)	Temp (N=37094)
Mean Obs	5.53 (m/s)	229.96°	276.30 K
Mean modeled	4.60 (m/s)	227.83°	275.21 K
Mean bias	-0.93 (m/s)	-2.14	-1.11 K
Mean error	1.98 (m/s)	50. 52°	4.02 K
Norm. mean bias (%)	-16.82	-9.93	-0.40
Norm. mean error (%)	35.75	21.97	1.46
Fract. bias (%)	-0.13	-0.01	-0.004
Frac. error (%)	0.41	0.30	0.02
Agrmt index	0.73	0.74	0.73

Figure 8a compares DREAM dust concentrations at 20Z with a GOES satellite image (Fig. 8b) acquired at 20:26Z on Dec 15, 2003 and a visibility analysis interpolation from ground observations taken at 20z (Fig. 8c). The comparison shows that themodeled dust concentration pattern is similar to the observed pattern, but there are significant differences with respect to details.

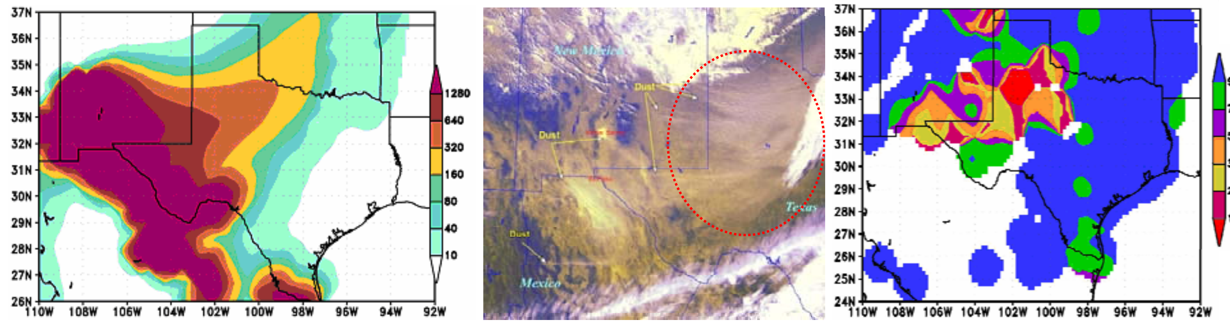


Figure 8. (Left), DREAM modeled dust concentration; (center), GOES 12 satellite image; (right) Measured visibility in miles (no observational data in the white areas).

Table 5 lists the performance statistics of modeled $PM_{2.5}$. They were calculated using modeled and observed $PM_{2.5}$ at 40 sites in the dust affected areas. The average modeled $PM_{2.5}$ concentrations at these sites are significantly (more than 3 times) higher than the measured average, possibly because DREAM outputs include dust at altitudes above the *in-situ* monitors. The mean bias and mean error are quite high. The agreement index of 0.12 is low. These metrics suggest there is considerable room for improvement if NASA experimental data sets were assimilated into the model to replace baseline parameters.

Table 5. Performance statistics of modeled surface $PM_{2.5}$ concentrations.

Metrics	***$PM_{2.5}$***
Mean observed	8.66 (μg/m^3)
Mean modeled	26.33 (μg/m^3)
Mean bias	17.67 (μg/m^3)
Mean error	26.51 (μg/m^3)
Agreement index	0.12

2.3 Assimilated EO products

The strategy of PHAiRS is to replace baseline DREAM parameters with satellite observations of the same, or closely allied, parameters that characterize the land surface and drive the dust entrainment module. Table 6 lists the baseline parameters and those selected for assimilation. These are: (1) dust source regions; (2) digital elevation; (3) aerodynamic surface roughness length; and (4), soil moisture.

Uncertainties in satellite observations exist in sensor design, the algorithms defining sensor products, and assimilation requirements. These limitations are governed largely by the laws of physics and chemistry that can be more or less engineered and characterized. All NASA data sets used by PHAiRS have validation programs.

Table 6. Baseline DREAM parameters and candidate assimilation parameters.

Baseline Parameters	***Assimilated Parameters***
Land Cover: Olson World Ecosystem 10-min. (19km) Res	MOD-12 1km resolution
Elevation: USGS 1km terrain data	SRTM-3 arcsec (90m) terrain data* resampled to 30 arcsec (1km)
Aerodynamic roughness length: predicted using 12 SSiB land cover types	Look-up table linked to MOD-12 land cover
Dust source areas	FPAR "Fill" class 254-255
Soil Moisture: simulated using a land surface model	AMSR-E

Dust Sources

Up-to-date patterns of land cover are important for DREAM to identify dust source areas and evolving dust storm episodes. PHAiRS assimilated the Moderate Resolution Imaging Spectroradiometer (MODIS) Land Cover product (MOD12Q1). It is based on 17 classes of land cover defined by the International Geosphere-Biosphere Programme (IGBP). The product supplies an assessment of the quality or confidence placed in the classification. Five of the 17 categories are relevant to dust entrainment: open shrublands; grasslands; croplands; urban and built-up; and, barren, or sparsely vegetated land.

Land cover categories for MOD12Q1 were produced by the MODIS Science Team using a supervised approach. Training sites were developed by analyzing high resolution imagery in conjunction with ancillary data. The classification used a decision tree algorithm (C4.5) in conjunction with a technique for improving classification accuracies known as “boosting.” Boosting improves classification accuracies by estimating classifiers iteratively using a base learning algorithm (e.g., a decision tree) while systematically varying the training sample. The training sample is modified through iteration to focus the algorithm on examples that are difficult to classify correctly. This modification is performed by providing a weight for each training example. The importance of misclassified training samples is increased and the classification algorithm focuses on learning these samples. The boosted classifier's prediction is then based upon an accuracy-weighted vote across the estimated classifiers. The boosting algorithm used for creating MOD12Q1 is Adaboost-M1, which is the simplest multi-class boosting method.

Recently, boosting has proven to be a form of additive logistic regression. As a result, probabilities of class membership are obtained from boosting. These probabilities are a means for assessing the confidence of the classification results, as well as a means for incorporating ancillary information in the form of prior probabilities to improve discrimination of cover types that are difficult to separate in the spectral domain.

In addition to the classifications and assessments, MOD12Q1 also provides mandatory quality information on whether each pixel has been newly classified or is dependant on a persistent value, and an embedded land/water mask as bit flags in the 8 bit Land_Cover_Type_QC parameter[3]. Overall, the product assimilated easily into DREAM. Its major drawback is that it is not being updated on an annual or seasonal basis for robust health applications.

To help identify dust source areas, the team examined Leaf Area Index (LAI), Enhanced Vegetation Index (EVI), and Fraction of Photosynthetically Active Radiation (FPAR). The MOD15 FPAR product holds the greatest promise for DREAM assimilation since it has a class (value 253) labeled "barren, desert, or very sparsely vegetated." In the FPAR algorithm, this value, among others for water, urban, and permanent snow and ice, is known as a "fill" class. Since the FPAR algorithm requires MOD12 as an input, it was thought to use class 253 to seasonally update MOD12 for DREAM model runs. The idea was tested over the White Sands National Monument in New Mexico by substituting MOD12 pixel values with FPAR class 253 values. Results were modest, indicating that the relationship is complex. It is questionable also whether FPAR fill values are updated seasonally along with non-fill classes.

Topography

Another data set for assimilation into DREAM is digital elevation. This parameter gives the model a realistic representation of the air/land interface. Terrain induced systems include land/sea breezes, mountain/valley winds, and forced airflow over and around rough terrain. Data from the Shuttle Radar Topography Mission (SRTM) were assimilated. The most recent version of this data set (released in May 2006) is called SRTM30, the global 30 arcsec [1km] product.

The SRTM Mission obtained elevation data on a near-global scale to generate a nearly complete high-resolution digital topographic database of Earth[4]. SRTM consisted of a specially modified radar system that flew onboard the Space Shuttle Endeavour during an 11-day mission in February 2000. To acquire digital elevation data, the SRTM payload was outfitted with two radar antennas. One was located in the shuttle's payload bay, the other on the end of a 60-meter mast that extended from the payload bay once the Shuttle was in space.

Before SRTM level-1 data could be assimilated, they had to be contiguous with no spikes, wells, or large voids. Voids are caused by geometric artifacts, specular reflection off water, phase unwrapping artifacts, and, complex dielectric constant (Dowding et al., 2004). SRTM data for the PHAiRS domain showed that the primary concern focused on two types of voids: small “salt and pepper” voids consisting of pixels having no SRTM response; and larger voids representing areas of contiguous pixels. The “salt and pepper” voids were replaced by interpolated values using a neighborhood filter. The larger voids were filled using ancillary data (Sanchez, 2007). Figure 9 illustrates the raw and filtered SRTM data for a small part of the DREAM domain.

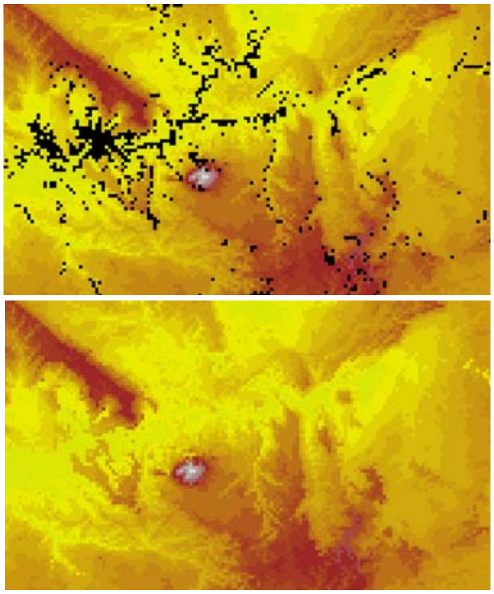


Figure 9. (Top) “salt and pepper” voids in SRTM data; (bottom) appearance after voids were removed using a neighborhood filter.

Aerodynamic Surface Roughness Length

SRTM data were used along with MOD12 to create a surface roughness layer for the DREAM model. Aerodynamic surface roughness length (z_0) is defined as the height above the ground at which wind speed is zero under neutral atmospheric stability (that is, where air temperature is isothermal and equal to that of the surface)[5].

For z_0, the less water held in a soil, the more prone it is to wind erosion and dust entrainment. Retention of soil water consists of two factors: (a) molecular adsorption on the surface of the soil grain; and (b), inter-particle capillary forces. The latter of these determines whether dust will be lifted from a surface at a given wind speed. As soil moisture is increased, the threshold wind velocity is also increased, thus reducing the amount of dust injected into the atmosphere (van Deursen et al., 1993; Nickovic et al., 2001).

To estimate z_0, one must measure surface momentum, soil temperature, and water vapor, among other surface properties. Conceptually, it is possible to measure these properties using sensors from different satellites, but the technology for creating a z_0 data set from different sensors into a form that can be assimilated into DREAM does not exist. To overcome this hurdle, the PHAiRS team transformed the MOD12 land cover data into a simulated z_0 product for assimilation. Table 7 is considered a “best practice” z_0 defined by standard physiognomic cover types extractable from MOD12 satellite observations by Stennis Space Center. It did not replace any Pre-PHAiRS DREAM parameter and therefore represents a novel derivation of data in the model. Model outputs showed modest improvement.

Table 7. Look-up values for surface roughness length (z_0). Source: Stennis Space Center.

DN	Land Cover Category	Z_0 Range (m)	Default z_0
8	Woody Savanna	0.10-0.20	0.15
9	Savanna	0.03-0.10	0.06
10	Grassland	0.03-0.07	0.05
12	Cropland	0.04-0.18	0.11
14	Crops/Natural Mosaic	0.10-0.30	0.20
16	Barren/Sparse	0.00-0.01	0.01
253	Fill	0.00	0.00

Soil Moisture

The Advanced Microwave Scanning Radiometer (AMSR-E) is a multi-frequency, dual-polarized sensor that detects emissions from the Earth's surface and atmosphere[6]. Passive microwave emissions can be used to estimate soil moisture in the surface centimeters (NSIDC, 2000). However, there are several challenges to assimilating data into DREAM: (a) the effective data footprint is almost 70km, while the model outputs are aiming toward 1km resolution; (b) the data are formatted to an Equal-Area Scalable Earth Grid or EASE-Grid, which is not readily compatible with PHAiRS remote sensing software; (c) there are serious data voids in areas of dense vegetation (high Leaf Area Index) and under snow cover; and (e), there are measurement errors associated with sampling depth and vegetation density

Despite formatting and resolution issues, soil moisture data from AMSR-E were assimilated into PHAiRS DREAM. Outputs showed little improvement in the model’s performance.

Aerosol Optical Depth

From its inception, PHAiRS expected that aerosol patterns would help define areas of elevated dust concentrations near reported dust events. The MODIS Aerosol Product (MOD04) monitors ambient AOT globally over the oceans, and over a portion of the continents. Aerosol *size* distribution is derived over the oceans, and aerosol *type* is derived over the continents. Level-2 data are produced daily at a horizontal resolution (at nadir) of 10×10km. Aerosols are one of the greatest sources of uncertainty in climate modeling. Concentrations and distributions vary in time and space and can lead to variations in cloud microphysics that in turn impact cloud radiative properties and climate. The PHAiRS team tested whether MOD04 AOT patterns would reveal dust events (Mahler, 2006) and discovered that they are not. While there are some cases where AOT data are reported for dust events, in most cases the data seem to be incoherent for the desert southwest, perhaps because they are interspersed with many pixels of no data.

2.4 Post assimilation model outputs

Data sets described in Section 2.3 were used to replace parameters in the model domain after the baseline model runs. MOD12Q1 replaced the out-dated OWE data set based on 1970s/80s data sources. The MOD12Q1 has a higher spatial resolution than OWE, 30-second versus 10-minute and it represents land cover conditions in 2001.

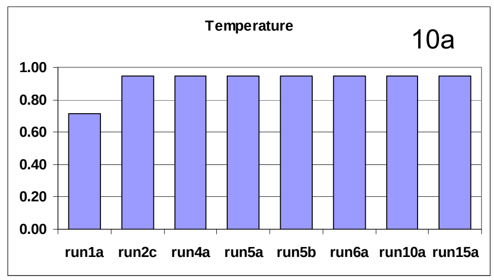
Following the land cover data, SRTM terrain data, z_0, FPAR data, and AMSR-E soil moisture data were assimilated into the DREAM system. FPAR data were used to locate areas covered with vegetation and, therefore to locate barren surfaces indirectly that are susceptible to generating airborne dust. Because of temporal resolution and incomplete coverage of the AMSR-E data, a composite of AMSR-E soil moisture data was assembled for initial assimilation into DREAM.

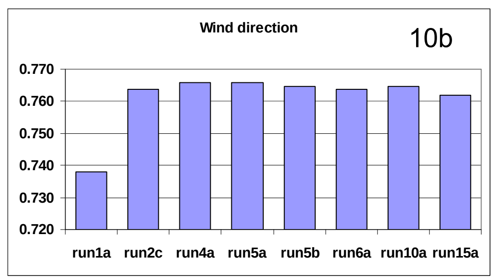
Table 7 lists key model runs and the NASA data sets that were assimilated (marked with Y). Run1a is the baseline (pre-assimilation) run. Mod12 (barren class) was a standard replacement set in all the model runs and was the only parameter replacement in run 2a. Run 4 assimilated the barren class and digital elevation from SRTM; runs 5a and 5b added z_0; run 6 was a test of Mod12 with Mod15 FPAR substituted for part of the domain; run 15 assimilated barren ground with soil moisture without digital elevation; and run 10a assimilated barren ground, SRTM and soil moisture.

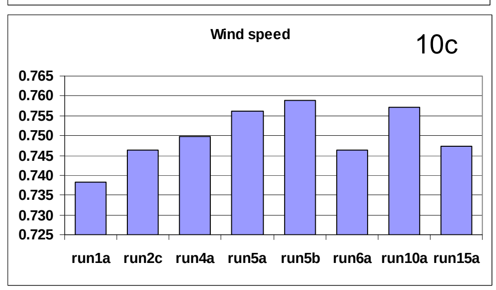
Table 7. Model runs using Earth observation data.

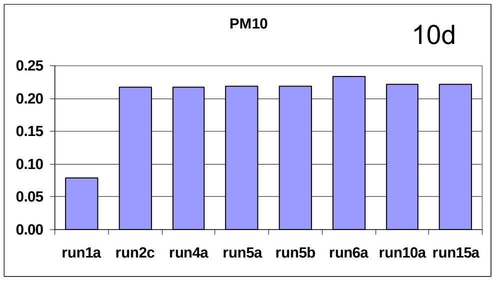
Run #	***MOD12***	***SRTM***	***Surface roughness length***	***FPAR***	***AMSR-E***
Run 1a					
Run 2c	Y				
Run 4a	Y	Y			
Run 5a	Y	Y	Y		
Run 5b	Y	Y	Y		
Run 6a	Y			Y	
Run 15a	Y				Y
Run 10a	Y	Y	Y		Y

Figures 10a-e show the agreement indices of modeled surface wind, temperature, $PM_{2.5}$ and PM_{10} concentrations compared to observations. Although impacts vary with different model runs, in general, the assimilation of NASA Earth observations data improved DREAM's performance measurably.









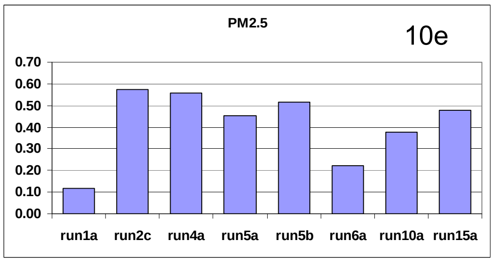


Figure 10. Agreement indices of modeled surface wind, temperature, $PM_{2.5}$ and PM_{10} concentrations compared to observations.

Figure 11 illustrates the pattern of modeled surface dust concentrations when MOD12Q1 replaced OWE land cover data. This pattern shows a much closer agreement with observed patterns than those patterns shown in Figure 8.

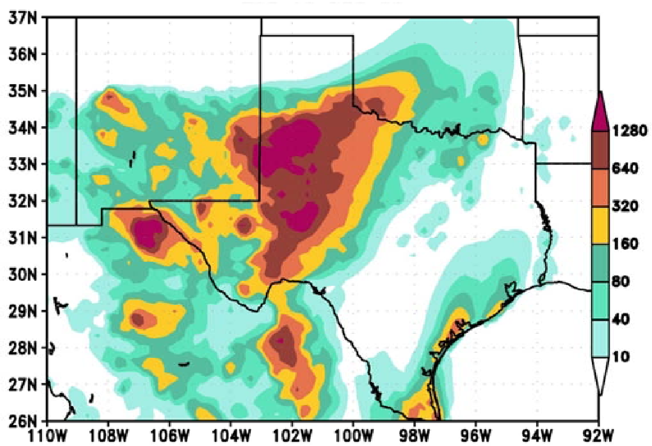


Figure 10. Modeled DREAM output using MOD12 data. Compare patterns shown in Figure 8 (p.8).

2.5 The V&V system

Verification and validation of DREAM outputs has been done by making qualitative and quantitative (statistical) comparisons of model outputs with *in-situ* dust concentrations reported by EPA's AIRNow network. To this end, development efforts at EDAC focused on three tasks.

Model Output Archive

The first task was to generate an archive of DREAM dust concentration data. This includes a daily DREAM model run for the 48-hour forecast beginning at 00:00:00 hours of the previous day. It also includes a *twice-a-day* DREAM model run for all days for 2006. The archiving system is designed to execute three model runs per day (two historical model runs for 2006, and one ongoing 72-hour forecast for the current day). The configuration of the model prevents concurrent execution of runs, so they are scheduled to minimize the potential for conflict. A single model run executes in approximately 5 hours, so a two hour buffer has been built into the execution schedule.

As of September 13, 2007, there were 335 separate 48-hour forecast datasets in the archive for the period January 1, 2006 to September 11, 2007. These are stored on the PHAiRS data server. Most model runs reach completion, but there are days when they do not. On 32 occasions (ca. 10%) the runs did not complete and had missing data. The team is exploring this situation and will report its findings in its final benchmark report.

Data Management and Web Services

The second task was to develop web services that permit system developers and health-care users to search for, access, and download dust concentration data generated by the DREAM model, as well as data collected *in-situ* by EPA's AIRNow network. Both the historical and daily forecasts are integrated into the PHAiRS data management system for delivery to public health decision support systems through simple object access protocols (SOAP) and web mapping service (WMS) interfaces published by the project.

The PHAiRS web service architecture allows users to search for and download both EPA AIRNow $PM_{2.5}$ and PM_{10} particulate data, as well as DREAM model output values for specific locations. Users can download $PM_{2.5}$ or PM_{10} AIRNow data for a defined date range, or for a single day. Similarly, SOAP service functions allow one to download both EPA and DREAM dust concentration values for a single station, or for all stations within the modeling domain, or to download data for a specific day, a 48-hour period corresponding to a DREAM model run, or a date range specified by the user. Note that, at present, the EPA AIRNow data values are not segregated into species. The downloadable *in-situ* values thus represent a composite measure of both geologically-derived and anthropogenically-produced particles.

Statistical Measures

The third task was to create web services that allow developers to generate statistical measures and indices. One of these, the *DREAM Data Access and Statistical Wizard*, allows one to extract modeled dust values for specified X-Y coordinates at specified times, and combine them with AIRNow values to generate statistics. In order to verify and validate the performance of consecutive versions of the model, web services have been designed to calculate *measures of central tendency* and *measures of variability* for both observed and modeled dust concentration values. These measures include the mean and standard deviation. Another set of statistics provides *measures of association* between these two variables. These include: mean observed value at each site; mean bias (0 if perfect); mean error (0 if perfect); normalized mean bias (0% if perfect); normalized mean error (0% if perfect); fractional bias (0% if perfect); fractional error (0% if perfect); and index of agreement (1 if perfect); the correlation coefficient (R); and the centered root mean square (RMS). These statistics can be obtained for a single station for a 48-hour DREAM run, or for a date range specified by the user.

2.6 In situ *V&V data streams*

AIRNow Reporting Stations

Historical AIRNow data (hourly $PM_{2.5}$ and PM_{10}) for the entire model execution period (2006-Present) are available through the DataFed's *AIRNow Web Coverage Service (WCS)*[7]. These data are acquired daily as a CSV file for all EPA stations within the DREAM domain for the previous 60 days. The daily reacquisition for the previous 60 days corrects data for stations that experienced delays in submitting values either to EPA's network or to DataFed's data ingest system.

During the development phase of the PHAiRS V&V system, questions arose regarding the timestamps encoded into the CSV files. Initially it was thought there was an undocumented offset to UTC, but subsequent discussions with DataFed revealed that timestamps encoded in the AIRNow data files vary by day and station, and that these timestamps are not consistently converted to UTC. This led DataFed to reconfigure its services to provide AIRNow data in UTC, regardless of the offset in the original data. This standard UTC format now provides unambiguous alignment of DREAM model outputs with well-defined ground observation times.

The *Data Access and Statistical Wizard* provides hourly DREAM output from 2006 to present and *in-situ* PM_{10} and $PM_{2.5}$ data from DataFed. The current web interface has 94 $PM_{2.5}$ and 41 PM_{10} sites for which modeled and observed data are collocated for side-by-side comparisons (Figures 12 and 13). Many sites have missing data for lengthy periods, especially for days of known dust events. It is suspected that *in-situ* sensors fail under extreme conditions and/or reporting of these events is delayed. It is unclear how many sites within the *in-situ* network have this problem, but often it happens that dust events of interest have missing data at many sites. It is sometimes possible to obtain data from the AIRNow website itself rather than through the DREAM web interface. Also, there is an obvious gap in station coverage for PM_{10} in central Texas, a region known to experience widespread dust events. Most AIRNow sites are located in cities, making validation over rural areas difficult. It has been shown also that the MOD12Q1 data for Mexico in the modeling domain improve validation statistics at US stations (Yin et al., 2007), yet to date there are no *in-situ* measurements from Mexico.

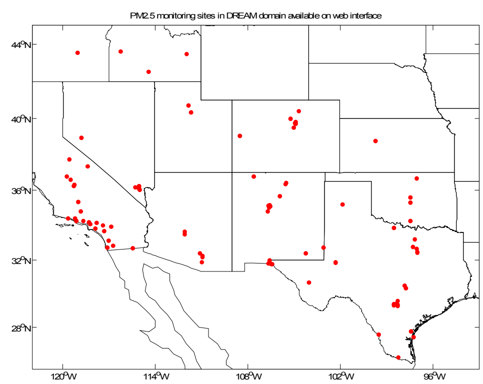


Figure 12. $PM_{2.5}$ monitoring sites in the DREAM

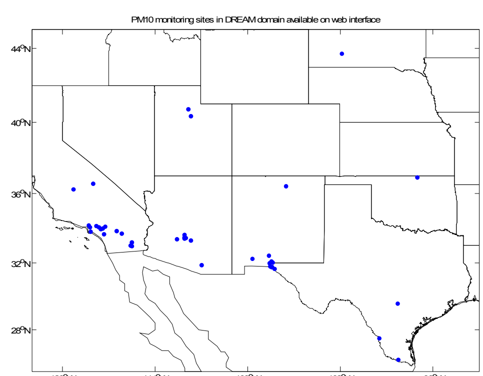


Figure 13. PM_{10} monitoring sites in the DREAM

Speciation in PM_{10} and $PM_{2.5}$ Dust

There are drawbacks to comparing model outputs with AIRNow data for PM_{10} and $PM_{2.5}$ because each fraction contains materials that are not generated by natural atmospheric processes. A more robust approach for health applications is to verify and validate these fractions continuously on the basis of individual species' concentrations.

PM_{10}, being larger in diameter and mass than $PM_{2.5}$, requires more momentum and higher wind speeds to be entrained. After lifting, this fraction also settles out of the atmosphere quicker. Because DREAM is strictly wind driven, and PM_{10} is almost always mechanically entrained, the coarse fraction is a better indicator of atmospheric dust events than $PM_{2.5}$. However, *in-situ* PM_{10} may be present in arid environments even in the absence of wind, and in such cases would not be predicted by DREAM. Anthropogenic concentrations often are present when DREAM predicts none. Fugitive dust from off-road vehicles, agricultural and construction dust clouds and emissions of larger pollutants from automobiles and factories add biases to PM_{10}. During non-windy conditions, it is still possible to observe other sources of PM_{10} that DREAM has no way of simulating. Due to its relatively large size, PM_{10} deposits in the upper thoracic region of the human respiratory system, and is often a concern for silicosis (Policard et al., 1952; Bar-Ziv and Goldberg, 1974; Norboo et al., 1991).

$PM_{2.5}$, on the other hand, may be present before and linger after weather-driven events. It penetrates deeper into the lungs and is a serious concern for chronic asthma, MI, and other respiratory conditions. Furthermore, its smaller size, makes validation more difficult. There are many more types of particles in the fine fraction. These finer particles include organic carbon as smoke from fires, soot from automobile emissions, and photochemical products. Other gases react photochemically forming ammonium sulfates and ammonium nitrates in this size range. Trace metals are produced via industrial emissions. Finally, natural aerosols are created mechanically as sea salt or windblown mineral dust. PHAiRS modelers have been concerned primarily with the mineral dust component, but these other components of $PM_{2.5}$ material complicate measurement of particulate concentrations, and therefore model performance. Total $PM_{2.5}$, as referred to here is the net concentration of all species in the air for that size range. DREAM has no anthropogenic emission module, so the other species and the anthropogenic signal in total $PM_{2.5}$ have been ignored.

The importance of speciation is evident in analyses of urban areas. El Paso, TX for example, experiences both desert dust storms and anthropogenic pollution episodes. DREAM can only model the former, so distinguishing the two using speciation is extremely beneficial for V&V. It is evident that during days of dust storms, the soil component comprises a much larger fraction of the total $PM_{2.5}$, while on non-windy days the other species dominate. While this is promising for V&V purposes, more frequent *in-situ* data are needed. Presently, only daily averages taken every 3rd day are used for speciation, so DREAM can be validated discretely only at this frequency. Continuous hourly data are ideal, but are probably not feasible due to cost and time restraints. The representation of cities in EPA's Speciation Trends Network (STN)[8] data is limited mainly to large metropolitan areas that monitor anthropogenic species. The soil component will usually be small in proportion to other species at these sites (Figure 14), but it is assumed to be larger in rural areas that are routinely exposed to desert dust and relative absence of a large human influence. Speciation at these sites may support the claim that a soil component is needed to validate the windblown dust model, and attempts to find such data are underway. One likely source is the Interagency Monitoring of Protected Visual Environments (IMPROVE)[9], a program designed to measure air quality in rural National Parks. Speciation and/or visual range data from this program could be used in future PHAiRS V&V efforts.

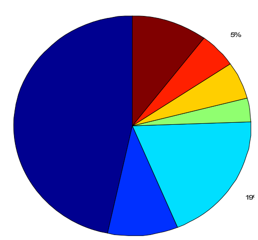


Figure 14. Mean dust conditions for El Paso, 2006: deep red= sulfate (10%), red = nitrate (5%), yellow = metals (5%), green = ions (4%), lt. blue = soil (19%), med. Blue = carbon emissions (10%), and deep blue = organic carbon (47%).

2.7 Model runs and statistics

For statistical V&V purposes, we regularly compare DREAM model runs to observational PM_{10} data during dust events that occur in the model domain, particularly in Texas and southern California. An example of these occurred on January 5^{th}, 2007. A severe wind and dust storm near Barstow caused a crash in which a minivan hit a tour bus, killing two and leaving others with severe injuries. Across the Southland, residents woke up to stacks of palm fronds on the ground, downed trees and other debris. The wind hobbled the morning commute, as freeways were jammed because of wind and several big-rigs toppled or jackknifed on freeways across the region (*High Winds Aren't Over Yet*, L.A. Times, January 6, 2007, p. A1). High winds occurred across the desert southwest, including parts of Texas. Using DREAM model hind-casting, the team investigated this dust event using seven PM_{10} AIRNow monitoring stations, four located in Southern California (Burbank, Riverside, Palm Springs, Indio) and three others in Texas (El Paso, Mission, Selma).

Figure 15 shows a 72-hour plot for each station (January 4-6, 2007) and illustrates the dust event that occurred around 23:00 UTC on January 5^{th} at most stations. The stations are plotted geographically west (l) to east (r). Southern California was affected most by this event. Both the observed and modeled data show a strong dust gradient from east to west, with the exception of Riverside, where no significant dust event was recorded in the observed data. Model improvements between an earlier version (15a) and the most recent version (20a) are evident in the decrease in magnitude of the latest model outputs. This improvement was accomplished in May 2007 with a correction to the bin size algorithm. Previous versions were 'grabbing' too much of the bin to represent PM_{10} values.

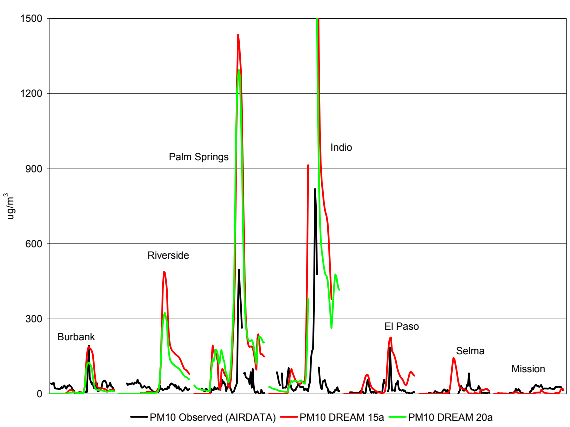


Figure 15. PM_{10} concentrations, modeled and observed, at seven AIRNow stations across the southwest for January 4-6, 2007.

Figure 16 represents the correlation between modeled and observed data for the January 4–6, 2007 test case. Derivation of the performance statistics is described in Yin et al. (2005). A total of 443 hourly values were used to compare modeled forecasts to the observed AIRNow data. Correlation lines are skewed toward the modeled data axis, illustrating the models tendency to over-predict dust events. Model improvements are indicated in the improved correlation from version 15a to 20a.

A statistical analyses that included the seven sites using the latest version of the model (20A) is shown in Table 8. These statistical parameters will be used to validate future versions of the DREAM model and to demonstrate improvements over previous versions.

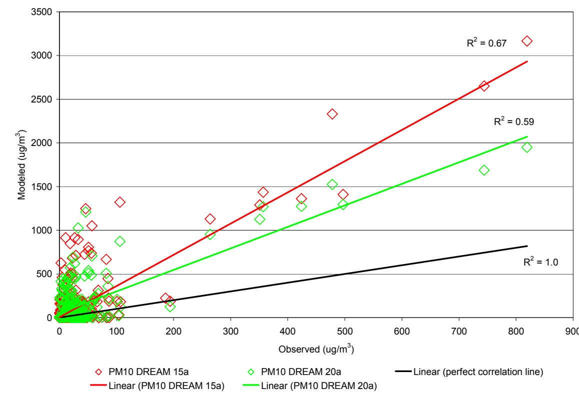


Figure 16. Magnitude correlation (seven sites, N = 443) during the Jan 4-6, 2007 dust episode.

Table 8. Statistical analysis of seven test sites.

N (seven sites)	***443 obs / 443 mod***
Mean	29.2 obs / 26.3 mod
Mean bias	2.8
Meas error	26.0
Normalized mean bias	10.8
Normalized mean error	76.2
Fractional bias	12.1
Fractional error	88.1
Index of agreement	0.63

The timing correlations for two test cases are shown in Figure 17. A 72-hour event clock is represented on the axes, and the modeled vs. observed peak hour is plotted. Several sites had more than one peak hour during the three-day event. A plot of daily peak hour for each site, would yield 21 data points. Occasionally, however, no peak hour was evident particularly on January 4th. These results (R^2 = 0.95) for model version 20A show an improvement over previous versions of the model published in earlier work (R^2 = 0.76, Yin et al, 2005).

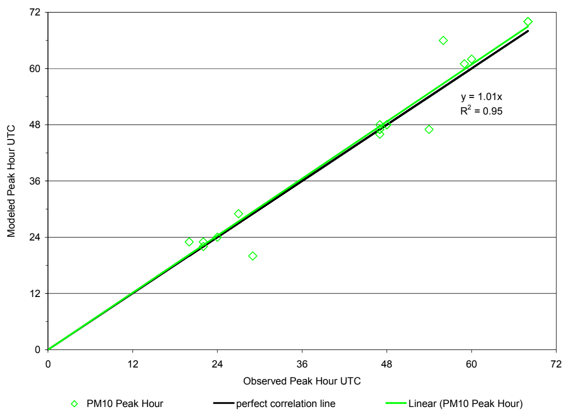


Figure 17. Timing Correlation (N=18 peak hours, seven sites) for the Jan 4-6, 2007 dust event.

Another test case was run for a high wind event during the last week of February 2007. Very strong and gusty westerly winds caused blowing dust over a large area of eastern New Mexico and northwest Texas on the afternoon and early evening of February 24th (Figure 18). A huge dust cloud was blown eastward across much of the eastern half of the state on the 25th and then stagnated over parts of Central, Southeast, and South Texas on the 26th and 27th. PM_{10} levels in parts of the southern Panhandle were hazardous on EPA's Air Quality Index (AQI) scale.

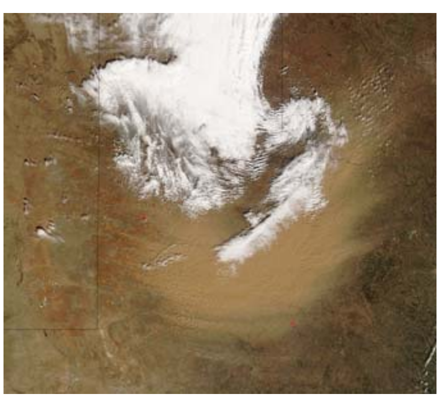


Figure 18. Dust storm in Texas on February 24, 2007 . Image is a cropped 500m resolution MODIS Terra image from Rapidfire.

Figure 19 shows the 72-hour plot for each station (February 23-25, 2007) and illustrates the dust event that occurred around 00:00 UTC on February 24th at most stations. The stations are plotted geographically west (l) to east (r). Two versions of DREAM are plotted (15a and 20a). The DREAM model under-predicted the event at Palm Springs, over-predicted the event at Indio, but performed well at the Texas sites, particularly at El Paso. Observed data from Selma and Mission, TX indicated a minor event and the DREAM model outputs were in fairly good agreement for these sites.

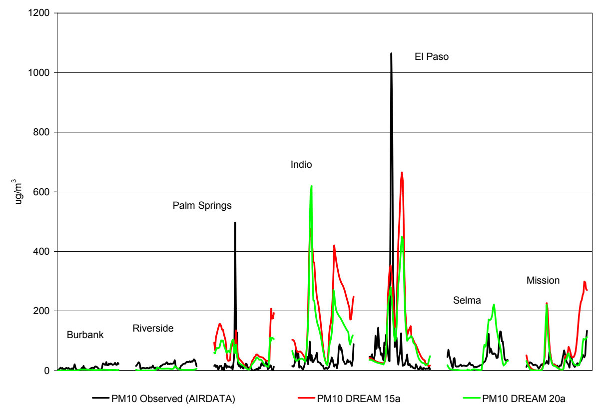


Figure 19. The February 23-25, 2007 dust episode, seven sites located in the model domain.

Figure 20 illustrates the magnitude correlation between modeled and observed data for the February 23-25, 2007 test case. Correlations for both model versions were poor for this test case (R^2 ~ 0.1), due primarily to the Palm Springs and Indio data discrepancy. In spite of this, the timing correlations (Figure 21) again show excellent agreement between observed and modeled peak hour.

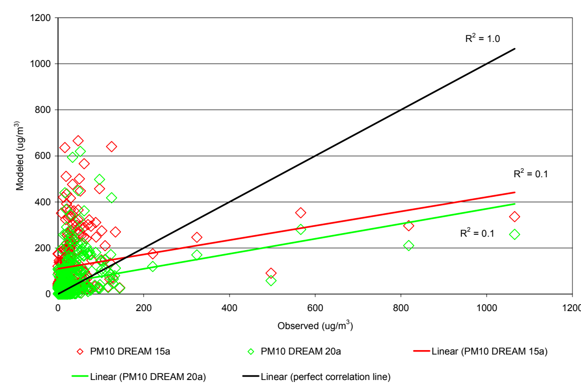


Figure 20. Magnitude correlation between observed and modeled data, February 23-25, 2007 test case.

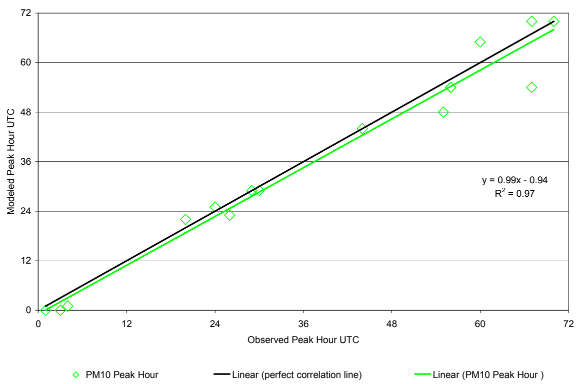


Figure 21. Timing correlation, February 23-25, 2007 test case (N=16 peak hours).

The same statistical analyses that included seven sites using the latest version of the model (20a) is shown in Table 9 for the February test case. The statistics indicate that the model had a negative bias, or under-predicted the February event. The January test case had a positive bias and a much better index of agreement (0.63 vs. 0.42).

Table 9. Statistical analysis of seven test sites, Feb 23-25, 2007.

N (seven sites)	**346 obs/346 mod**
Mean	*34.1 obs/59.3 mod*
Mean bias	*-25.0*
Mean error	*56.0*
Norm. mean bias	*-42.4*
Norm. mean error	*67.7*
Fractional bias	*9.7*
Fractional error	*122*
Index of agreement	*0.42*

In summary, these results will be used to assess the improvements made in future model versions. The two test cases illustrated here indicate that the model can accurately predict the timing of the dust events, but the prediction of the magnitude of events are a mixed result.

3.0 Health Data

The integrated system solution executed in PHAiRS has focused on inputs and outputs (i.e., Missions and Models) that might be used by the health community to formulate decision support systems. A NASA directive from the Earth Science Applications Division strongly discouraged use of funds to "develop a DSS". Nevertheless, a requirement of the project was to identify a candidate DSS into which project outputs could be inserted and tested. The Rapid Syndrome Validation Program (RSVP) was PHAiRS' proposed decision support system. However, between 2003 and 2005, RSVP morphed into a commercial system called the Syndromic Reporting Information System (SYRIS). This system represents a sophisticated convergence of modeled geostatistical and biostatistical processes. RSVP was beta-tested over a 25,000 square mile area surrounding Lubbock, TX (Morain and Sprigg, 2005) and was subsequently deployed as SYRIS over the Texas Department of State Health Service, Public Health Region 1, covering 41 counties surrounding Lubbock (Lindley, 2006).

Efforts to relate PHAiRS modeled results with hospitalizations, school nurse records, and emergency room admissions have been made, but initial results are too few for verification and validations purposes. We expect baseline biostatistics will include Poisson regression, zero-inflated Poisson regression (ZIP), generalized additive models of daily visit counts, and logistic regression of daily proportion of respiratory visits diagnosed as asthma, MI, or other respiratory conditions. In the meantime, the PHAiRS system has been designed to facilitate these statistical analyses, and to make dust forecasts and compliant aggregated health data available through web-based services to qualified health authorities for statistical analyses.

The following Sections are intended to provide the reader with information about the data and analytical processes.

3.1 Types of data and uncertainties

Typically primary sources of health outcome data are derived from statewide hospital data queries of emergency department and hospital inpatient discharge records for asthma and MI. Other common health records are kept by Vital Records and Health Statistics, Indian Health Service, Medicaid, and a variety of surveys such as behavioral and risk factors surveys.

Uncertainties in public health data far surpass those for environmental measurements and modeling. Uncertainties in health data begin with individual genetics, and magnify at each step in the reporting chain from the onset of symptoms or syndrome (e.g., knowing the exact location of the individual at the time of exposure, what that individual was doing at the time of exposure, the duration of the exposure, and post-exposure activities). More than likely the patient can only describe the answers to these questions in general terms, which leads first responders like school nurses, ER personnel, physicians, and others to treat the case along prescribed best practices aided by patient history.

Longitudinal studies of chronic respiratory diseases like asthma are in their infancy (Gauderman, et al. 2004). In the absence of mandatory autopsies, very few MI cases are reported as "caused by dust."

Exacerbating uncertainties inherent in health records is the health care system itself. This system is balanced between being a commercial enterprise and a social/humanitarian requirement. Hospitals earn revenue from inpatient care (i.e. number of beds occupied). Comparatively little revenue is realized from emergency room (outpatient) care. Increasingly, ER patients are diagnosed and released rather than admitted as inpatients. Because ER operations are financial "loss leaders," episodic increases in outpatient arrivals are diagnosed quickly and reported through a coding system that validates reimbursements to hospitals, but because of time constraints, frequently results in partial or misdiagnosis. Respiratory diseases are often assumed to be infectious, resulting in patients being given antibiotics for an asthma condition that is chronic but exacerbated by atmospheric contaminants. The loss of inpatient admissions has led many hospitals to reduce the number of beds and the accompanying requirement to have permanent, full-time personnel to service those beds.

3.2 Geo- and Biostatistical Processes

Data mining and clinician-based syndromic surveillance strategies are both being explored by CDC. Situational awareness is essential for early detection of infectious diseases and bioterrorism threats, but most public health compliance reporting focuses on notifiable diseases. There is a critical time lag of several weeks between situational awareness and notifiable reporting, when what is needed is rapid syndromic surveillance that provides actionable information within hours.

Initial work integrating the geostatistical capabilities of the PHAiRS system with biostatistical analyses has resulted in statistical routines that summarize the hourly DREAM model outputs and AIRNow measurements for Lubbock and the Midland/Odessa Texas areas for the first two months of 2006. These summary data were generated using the R statistical programming language, and are based upon data retrieved from the PHAiRS HTTP interface to the data extraction SOAP services. Requests for comma-separated-value (CSV) data may be submitted to the PHAiRS web server. These requests are converted by the web server interface into SOAP service calls to the PHAiRS analytical services that extract pixel values from a series of DREAM model outputs and query the database for corresponding AIRNow measurements for the same location. The resulting data are formatted as CSV files and delivered to the requesting system in a format suitable for data ingest and processing. Since R can use a network-accessible resource (such as the above described HTTP-based system) as a data source in an analysis, the product generated by R consists of a new CSV file containing the daily summary data for both the DREAM model and AIRNow, and a set of URL web addresses where the hourly data from which the daily summaries are derived may be obtained. Such a CSV file has been used to integrate biostatistical analysis for correlation between $PM_{2.5}$ concentration and emergency room admissions for respiratory problems in the Lubbock area.

The issue of catchment modeling in the biostatistical analyses has also been considered. Specifically, in order to better represent the particulate concentrations to which a population has been exposed, the geographic area of that population must be defined. That geographic area then is used to extract and process air quality data. While not yet implemented, it appears that this will be a necessary next step in developing a reasonable model for the capture and presentation of air quality and health data in a consistent and statistically valid manner. An initial capability for the summarization of DREAM model outputs for irregular polygons (counties thus far) has already been developed as part of the PHAiRS SOA, so a capability for providing daily summaries for regional model outputs (as opposed to single model cell/pixel) has already been developed. Further development of this capability will contribute directly to this catchment-based analytical approach.

4.0 Summary and Conclusions

Remote sensing of the environment is critical in advanced systems to warn of imminent, life-threatening sand and dust storms and to reduce risk of exposure to mineral dust concentrations that contribute to cardiovascular and respiratory disease. MODIS data improve identification of active mineral dust sources, and thus, numerical model simulations and forecasts of dust generation, entrainment, and downwind dispersal and deposition.

An advanced numerical dynamical model of dust generation and entrainment (DREAM), driven by operational, validated, weather forecast models of the U.S. National Weather Service (eta), initialized with MODIS landscape information, can forecast the timing of an advancing dust storm verifiably to meet the needs of many users. While the dust forecast system developed under PHAiRS simulates and predicts the three-dimensional size-concentration characteristics of the dust cloud, verification of model output is problematic.

For V&V of airborne particulate concentration, PHAiRS relies mainly on a regionally sparse network of *in-situ* particulate sampling stations for statistical comparison DREAM-generated PM_{10} and $PM_{2.5}$ concentrations. Furthermore, the sampling networks are concentrated in densely-populated, large urban areas, subject to PM_{10} and $PM_{2.5}$ sources generated by human activity, as in construction and combustion. And, too few speciated particle sampling sites are available to identify natural vs. man-made sources. The PHAiRS comparisons of optical depth in the NASA/ AERONET network of photometers, and airport networks measuring visibility, have provided other quantitative measures against which to compare model output. A-Train's CALIPSO and GLORY offer near-term opportunities to test satellite-based measurements of aerosol profiles for future V&V, as would greater access to ground-based LIDAR sensors, which have been used to validate dust model performance in the Mediterranean region.

Specific uncertainties exist in each data-set/product. For example, the MOD12Q1 product offers only the one class for "barren." This class includes not only barren ground, but rock surfaces and un-vegetated urban pixels. Typically, seasonally active agricultural dust sources are not distinguished. Even though use of the MOD12Q1 product improved the DREAM output, it is not certain why this product made a difference. Was it only the spatial resolution of the assimilated data vis a vis the surface data used in baseline DREAM? We intuit that soil moisture is important; but when AMSR-E soil moisture data were assimilated in the model runs, no significant improvement in model performance occurred.

Products specifically designed with the end user in mind are being evaluated in key state offices with operational health and air quality responsibilities. These products will be modified as needed, and further V&V will play a large role in adapting/adopting the new technology developed under PHAiRS for public health services.

5.0 References

Ackerman, F. 2002. Pricing the Priceless: Cost-Benefit Analysis of Environmental Protection. Georgetown University.

American Lung Association, 2005. Trends in Asthma Morbidity and Mortality. Epidemiology and Statistics Unit, Research and Program Services. May

Bar-Ziv, J. and G.M. Goldberg. 1974. "Simple Siliceous Pneumoconiosis in Negev Bedouins," *Arch. Environ. Health*, 29: 121.

CDC. 1998. Surveillance for Asthma—United States, 1960-1995. MMWR. 47(SS-1).

Centers for Disease Control and Prevention. 2005. *Syndromic Surveillance: Reports from a National Conference*, 2004. MMWR 2005:54 (Suppl). 208 pages.

Cosby, B.J., G.M. Hornberger, R.B. Clapp, T.R. Ginn, 1984. A Statistical Exploration of the Relationships of Soil Moisture Characteristics to the Physical Properties of Soils, *Water Resources Research*, 20: 682-690.

Davies, K. 2005. How Much Do Environmental Diseases and Disabilities Cost? *Northwest Public Health*. Fall/Winter.

de La Beaujardiere, J. 2004. Web Map Service, Version 1.3. Open Geospatial Consortium. OGC 04-024.

Dockery, DW. C.A. Pope III, X. Xu, et al. 1993. An Association Between Air Pollution and Mortality in Six US Cities. *N. Engl. J. Med*. 329:1753-1759.

Dowding, S., Kuuskivi, T., & Li X., 2004. Void fill of SRTM elevation data – principles, processes and performance, *Images to Decisions: Remote Sensing Foundations for GIS Applications,* ASPRS Fall Conference, September 12-16, Kansas City, MO, USA.

Evans, J.D. 2003. Web Coverage Service (WCS), Version 1.0.0. Open Geospatial Consortium. OGC 03-065r6.

Gauderman, W.J., E. Avol, F. Gilliland, H. Vora, D. Thomas, K. Berhave, R. McConnell, N. Kuenzli, F. Lurmann, E. Rappaport, H. Margolis,

D. Bates, and J. Peters. 2004. The Effect of Air Pollution on Lung Development from 10-18 Years of Age. *N. Engl. J. Med*. 351(11):1057-1067.

Griffin, D.W. 2007. Atmospheric Movement of Microorganisms in Clouds of Desert Dust and Implications for Human Health. *Clinical Microbiology Reviews*, 20(3):459-477.

Jerrett, M., J. Eyles, C. Dufournaud, and S. Birch. 2003. Environmental Influences on Healthcare Expenditures: An Exploratory Analysis from Ontario, Canada. *J. Epidemiol. Comm. Health*. 57:334-338.

Kaiser, J. 2005. Mounting Evidence Indicts Fine Particle Pollution. *Science*. 307(1717): 1858-1861.

Kuehn, B.M. 2006. Desertification called Global Health Threat. *JAMA*. 295(21):2463-2464.

Landrigan, P.J., C.B. Schechter, J.M. Lipton, M.C. Fahs, and J. Schwartz. 2002. Environmental Pollutants and Disease in American Children: Estimates of Morbidity, Mortality, and Costs for lead Poisoning, Asthma, Cancer, and Developmental Disabilities. *Environmental Health Perspectives*. Vol. 110, # 7, July.

Lindley, C. and T. Ward. 2006. Experience with Clinician-based Syndromic Surveillance in West Texas. Annual Conference, International Society for Disease Surveillance. Baltimore MD. October 19-20. CD-ROM

Mahler, A-B., K. Thome, D. Yin, and W.A. Sprigg. 2006. Dust Transport Model Validation Using Satellite- and Ground-based Methods in the Southwestern United States. In: A. Chu, J. Szykman, and S. Kondragunta (eds.). *Remote Sensing of Aerosol and Chemical Gases, Model Simulation/Assimilation, and Applications to Air Quality*, Proc. SPIE Vol. 6299, 62990. 12 pgs.

Massey, R. and F. Ackerman, 2003. Costs of Preventable Childhood Diseases: The Price we Pay for Pollution. Tufts University. September.

Morain, S.A. and W.A. Sprigg. 2005. Initial Benchmark Report for Public Health (February 2004-September 2005). NASA Agreement NNSO4AA19A.

Myers, N. 2006. A Networker Index on Environmental Health Costs. *The Networker*. Vol. 11(3). May.

Norboo, T., P.T. Angchuk, M. Yahya, S.R. Kamat, F.D. Pooley, B. Corrin, I.H. Kerr, N. Bruce, and K.P. Ball. 1991. Silicosis in a Himalayan Village Population: Role of Environmental Dust. *Thorax*, 46: 341-343.

NSIDC. 2000. Advanced Microwave Scanning Radiometer For EOS (AMSR-E) Science Data Validation Plan Version 2. National Snow and Ice Data Center, University of Colorado, Boulder, CO July, 76 p.

National Research Council and Institute of Medicine. 2007. *Earth Materials and Health: Research Priorities for Earth Science and Public Health*. National Academies Press. 176 pgs.

Nickovic, S., A. Papadopoulos, O. Kakaliagou, and G.Kallos. 2001. Model for Prediction of Desert Dust Cycle in the Atmosphere. *J. Geophys. Res*. 106:18113-18129.

Nickovic, S., W. Sprigg, R. Clark, and A. Micallef. 2004. Environmental Modelling Programme of the World Laboratory, Malta Centre. In: Annual Meeting, World Federation of Scientists. Geneva, Switzerland.

Perez, C., S. Nickovic, M. Baldasano, F. Sicard, F. Rocadenbosch, and V.E. Cachorro. 2006. Saharan Dust Over the Western Mediterranean: LIDAR, Sun Photometer Observations and Regional Dust Modeling. *JGR*, Vol 111, 38 pgs.

Pear, R. 2003. Spending on Health Care Increased Sharply in 2001. *New York Times*. Jan. 8.

Policard, A. and A. Collet. 1952. Deposition of Silicosis Dust in the Lungs of the Inhabitants of the Saharan Region, *Arch. Indust. Hyg. Occupat. Med.,* 5: 527-534

Pope, C.A. III. 1989. Respiratory Disease Associated with Community Air Pollution and a Steel Mill, Utah Valley. *Am. J. Public Health*. 79:623-628.

Pope, C.A. III. 2004. Air Pollution and Health. *New England Journal of Medicine*. 351(11): 1132-1133.

Pope, C.A. III, M.J. Thun, M.M. Namboodiri. 1995. Particulate Air Pollution as a Predictor of Mortality in a Prospective Study of US Adults. *Am. J. Respir. Crit. Care Med*. 151:669-674.

Sanchez, G.M., 2007. The Application and Assimilation of Shuttle Radar Topography Mission Version 1 Data for High Resolution Dust Modeling. MA Thesis, Univ. New Mexico, Department of Geog. 51 pgs.

Schwartz, J. and DW Dockery. 1992. Increased Mortality in Philadelphia Associated with Daily Air Pollution Concentrations. *Am. Rev. Respir. Dis*. 145:600-604.

van Deursen, W.P.A., G.W. Heil, and A.M.J.V. van Boxtel. 1993. Using Remote Sensing Data to Compile Roughness Length Maps for Atmospheric Deposition Models. In: J.L. van Genderen, R.A. van Zuidam and C. Pohl (eds.) *Operationalization of Remote Sensing in the Netherlands*. Vol 3, pgs 79-90. ITC Enschede, The Netherlands.

Yin, D., S. Nickovic, B. Barbaris, B. Chandy, and W.A. Sprigg. (2005). Modeling Wind-blown Desert Dust in the Southwestern United States for Public Health Warning: A case study. *Atmospheric Environment*. 39(33):6243-6254.

Yin, D., S. Nickovic, and W.A. Sprigg. (2007). The Impact of Using Different Land Cover Data on Wind-blown Desert Dust Modeling Results in the Southwestern United States. *Atmospheric Environment*. 41(10):2214-2224.

Zobler, L., 1986. "A World Soil File for Global Climate Modeling," *NASA Technical Memorandum 87802*. NASA Goddard Institute for Space Studies, New York, New York, USA.

Acronyms

AE – Angström Exponent

AMSR-E – Advanced Microwave Scanning Radiometer for EOS

AOD – Aerosol Optical Depth

AOT – Aerosol Optical Thickness

AQI – Air Quality Index

AQS – Air Quality System

BGC – BioGeochemical Cycles

BLM – Bureau of Land Management

CDC – Centers for Disease Control and Prevention

CGI – Common Gateway Interface

CSV – Comma-separated-value

DAAC – Distributed Active Archive Center

DREAM – Dust Regional Atmospheric Model

DSS – Decision Support System

EASE-Grid – Equal-Area Scalable Earth Grid

ECMWF – European Center for Medium-Range Weather Forecast

EO – Earth Observation

EOS – Earth Observation System

EPA – Environmental Protection Agency

ER – Emergency Room

ESMF – Earth System Modeling Framework

ESR – Earth Science Results

FAO – Food and Agriculture Organization

FPAR – Fraction of Photosynthetically Active Radiation

GEO – Group on Earth Observations

GIS – Geographic Information System

GLAS – Geoscience Laser Altimeter System

HTTP – Hypertext Transfer Protocol

ICESAT – Ice, Cloud, and land Elevation Satellite

IGBP – International Geosphere-Biosphere Programme

IMPROVE – Interagency Monitoring of Protected Visual Environments

ISDS – International Society for Disease Surveillance

ISS – Integrated System Solution

LAI – Leaf Area Index

LiDAR – Light Detection and Ranging

METAR – Meteorological Aerodrome Report

MI – Myocardial Infarction

MODIS – Moderate Resolution Imaging Spectroradiometer

MRLC – Multi-Resolution Land Characteristics Consortium

NASA – National Aeronautics and Space Administration

NCAR – National Center for Atmospheric Research

NCEP – National Centers for Environmental Prediction

NGA – National Geospatial Intelligence Agency

NLCD – National Land-Cover Database

NMB – Normalized Mean Bias

NME – Normalized Mean Error

NOAA – National Oceanic and Atmospheric Administration

NOMADS – National Operational Model Archive and Distribution System

NPOESS – National Polar-Orbiting Environmental Satellite System

NPS – National Park System

NRCS – Natural Resources Conservation Service

NSIDC – National Snow and Ice Data Center

O_3 – ozone

OGC – Open Geospatial Consortium

OWE – Olson World Ecosystems

PHAiRS – Public Health Applications in Remote Sensing

POI – Plan of Implementation

PM2.5 – Particulate matter at 2.5 μm

PM10 – Particulate matter at 10 μm

QA/QC – Quality Assurance / Quality Control

REASoN – Research, Education, and Applications Solution Network

REGAP – Regional Gap Analysis Project

RMSE – Root Mean Square Error

RSVP – Rapid Syndrome Validation Project

SeaWiFS – Sea-viewing Wide Field-of-view Sensor

SOA – Service Oriented Architecture

SOAP – Simple Object Access Protocol

SRTM – Shuttle Radar Topography Mission

SST – Sea Surface Temperature

STN – Speciation Trends Network

SYRIS – Syndrome Reporting Information System

TCEQ – Texas Commission on Environmental Quality

TM – Thematic Mapper

UMD – University of Maryland

UN – United Nations

UNESCO – United Nations Education, Scientific and Cultural Organization

URL – Uniform Resource Locator

USFS – United States Forest Service

USFWS – United States Fish and Wildlife Service

USGS – United States Geological Survey

UTC – Coordinated Universal Time

V&V – Verification and Validation

WCS – Web Coverage Services

WHO – World Health Organization

WMS – Web Mapping Services

WSSD – World Summit on Sustainable Development

ZIP – Zero-inflated Poisson regression

---

[1] http://biodi.sdsc.edu/Doc/Why/Where/meta/owe14d.txt

[2] www.fao.org/ag/agl/agll/wrb/wrbmaps/htm/soilres.htm

[3] www.modis.bu.edu/landcover/userguidelc/consistent.htm

[4] www.2.jpl.nasa.gov/srtm/SRTM_paper.pdf

[5] www.Ieeexplore.ieee.org/iel5/10226/32599/01526482.pdf

[6] http://sharaku.eorc.jaxa.jp/AMSR/ov_amsr/index.htm

[7] http://datafed.net/

[8] http://www.epa.gov/ttn/airs/airsaqs/detaildata/

[9] http://vista.cira.colostate.edu/improve