PHAiRS Public Health Applications in Remote Sensing

Final Benchmark Report

(February 2004-September 2008)

Agreement NNSO4AA19A

Report submitted to:

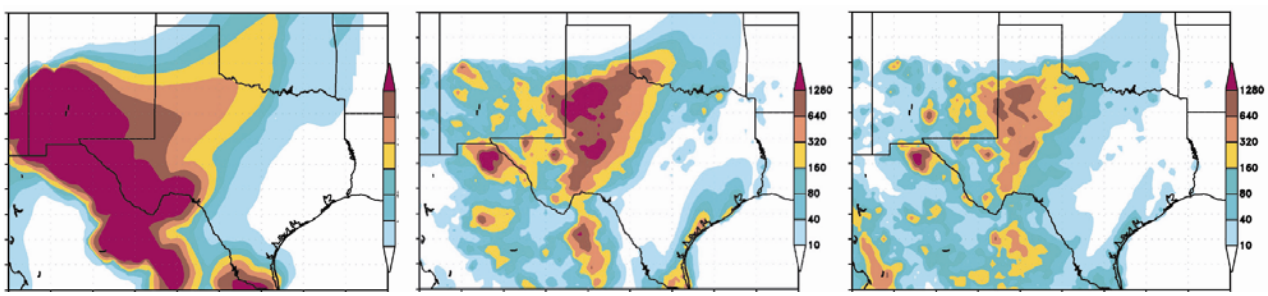
John Haynes, Program Manager, NASA Headquarters

by

Stanley A. Morain, Principal Investigator, University of New Mexico

and

William A. Sprigg, Co-Principal Investigator, University of Arizona



September 30, 2008

Contributing Project Members (alphabetically by last name and by Institution):

University of NM: Karl Benedict, Amelia Budge, Thomas Budge, William Hudspeth, and Gary Sanchez

University of AZ: Brian Barbaris, Christopher Catrall, Beena Chandy, Anna-Britt Mahler, Patrick Shaw, Kurt Thome, Slobodan Nickovic, and Dazhong Yin

Stennis Space Center: Donald Holland, SSAI

Texas Tech Health Science Center: James Spear

University of NM Health Science Center: Gary Simpson, Alan Zelicoff

Cover Photo

The triptych represents three generations of DREAM model outputs. At left, DREAM/SW shows the pattern of dust concentrations for a dust storm across New Mexico and Texas on 15-16 December, 2003 as modeled by DREAM/SW before NASA Earth science data sets were assimilated to replace the original design parameters of DREAM/SW; at center is the output from *enhanced* DREAM/SW (*e*D/SW) after MOD12Q1 and SRTM topography were assimilated; at right is the same *e*D/SW configuration nested within the higher-resolution NCEP/NMM (Non-hydrostatic Mesoscale Model). The left image was produced in 2005, the center in 2006, and the right in 2007.

Executive Summary

Many challenges in Earth system science require integrating complex physical processes into system models, and coupling environmental biogeochemical and chemical phenomena. Among the most important of these challenges are the coupled processes that affect human health. Future generations of healthcare providers and scientists will need to form teams representing both the biogeophysical realm and the medical and health realms to properly understand disease outbreaks that could have devastating epidemic outcomes. The *Public Health Applications in Remote Sensing (PHAiRS)* project is concerned with pulmonary and cardiovascular diseases that are exacerbated by poor air quality, primarily in this report by particles that are entrained and transported in dust storms across the American Southwest. These storms affect the rates and severity of chronic lung diseases, particularly in the very young and aging populations.

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. Coupling the environmental processes that lift dust into the atmosphere along with associated airborne pathogens would allow epidemiologists to better understand the medical consequences of dust and pathogen transport across the Southwest. Through Internet and Intranet surveillance and reporting systems, medical professionals might better diagnose individual patient symptoms in ways that enable them to alert those at higher risk during severe dust episodes. Health officials, on the other hand, seek early warning of dust episodes that could lead to disease outbreaks, harmful working conditions, or other hazards so they can plan interventions that reduce exposures throughout their constituencies. The role dust and airborne pathogens play in human health is an important part of Earth system science, with consequences that are rooted deeply into social and economic stability.

Asthma is a primary concern for the American College of Allergy and Immunology. It is one of the most common chronic diseases in the United States and is the most prevalent chronic disease in children. For poorly understood reasons, the rate of asthma among children in the northern mid-latitudes has more than doubled in the last 20 years. Nationwide, more than 9 million children struggle to breathe. According to the National Center for Health Statistics, asthma causes more missed school days than any other chronic condition, and is the leading cause of hospitalization for children under 15. Asthma is the most common reason that children younger than five go to the emergency room. Based on outpatient visits, the prevalence of asthma has increased by 50 percent over the last decade. The increase in morbidity and mortality has not been uniform throughout the country (Moorman et al., 2007). Urban settings, and minority groups within them, have experienced higher rates of increase. However, much about the disease is unknown, particularly in regard to incidence and prevalence in rural areas where access to care is limited and exposure to organic and inorganic dusts is inescapable.

The Earth Science Applications Division in NASA's Science Mission Directorate supports a public health benefit area to demonstrate that Earth science sensor products can improve dust storm simulations and forecasts. The PHAiRS project has defined a set of products that can map the three-dimensional characteristics of dust clouds, and has verified and validated the improvements in a model specifically designed to incorporate these products. The model is named *enhanced* DREAM/SW (*e*D/SW).

Ultimately, the goal of NASA's investments in PHAiRS is to insert a space-based component into a public health decision making system that can continue to evolve in-step with the next generation of space-based sensor systems. The National Polar-orbiting Environmental Satellite System (NPOESS) will consist of platforms carrying operational versions of NASA's current experimental sensors. As a precursor to such systems, this report documents the scientific and technological underpinnings of inserting

Earth observations data into models and verifies and validates model outputs for use by appropriate public health user communities.

The Project

PHAiRS is a step toward linking atmospheric dust events to human respiratory health outcomes. The *e*D/SW forecasts dust patterns and concentrations by being nested within, and driven by, an operational numerical weather forecast model of the U.S. National Weather Service called the National Centers for Environmental Prediction, eta version (NCEP/eta). 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. State agencies like public health departments, environment departments and air quality offices are responsible for monitoring the Southwest's air quality for public health conditions, and provide the ultimate test of *e*D/SW's new capabilities. The project has therefore worked closely with these communities to describe the capabilities and to test the output products.

The project had three goals. The first focused on assimilating satellite data from NASA's Terra and other platforms into a baseline model developed originally for use in the Mediterranean region, called the Dust Regional Atmospheric Model (aka here DREAM/MED) This model was adapted for use in the Southwest as DREAM/SW. Both are driven in their atmospheric parameters by NCEP/eta. The aim was to: (a) verify that advanced satellite image data from research sensors could replace DREAM/SW parameters to evolve *e*D/SW; and (b), validate that parameter replacements could lead to more refined model forecasts of dust episodes. NCEP/eta provided the meteorological parameters for traditional weather forecasts, while *e*D/SW provided the terrain parameters needed for simulating dust entrainment.

The second goal focused on optimizing model outputs by iterating inputs with a variety of satellite products and assessing incremental improvements to a health system called the Syndromic Reporting Information System (SYRIS™). The questions of greatest interest were: (a) could various iterations of the model system simulate dust entrainment over the Southwest? (b) could these models predict the speed and direction of moving dust clouds? (c) could reliable and repetitive products be generated for health authorities to use for public health alerts, or for healthcare practitioners to use in day-to-day syndromic reporting? and (d), could dust storm movement be forecast in a timely fashion to alert populations at risk?

The third goal involved establishing collaborative relations with public health communities to develop statistically valid relationships between dust episodes and increased respiratory complaints. This is difficult in the United States because health authorities are distributed throughout all levels of government, and because standardized record keeping is not mandatory within or between these levels. Developers of SYRIS designed its system to encourage public health officials, air quality monitoring offices, doctors, and clinicians to report their information electronically, and in appropriate ways to protect patient confidentiality. The system also allows group attributes to emerge in geospatially explicit ways that populations at risk can be forewarned.

Results

Results showed that meteorological fields (both surface and 500 hPa geopotential height) modeled by *e*D/SW were in agreement with measured observations. The modeled vertical profiles of wind speed, wind direction, temperature, and specific humidity also matched the observed profiles. Statistical evaluation of the modeled and observed surface winds and temperatures showed the model performed well in reproducing the measured values.

Comparing model runs before and after NASA data assimilation (that is, DREAM/SW compared to *e*D/SW) showed that sea level pressure, 500 hPa geopotential height, and temperature patterns matched well with traditional weather observations. Differences between the before and after model runs occurred in sea level pressure fields, but did not affect the overall pattern. Most importantly, the upper-air fields were *not* affected by assimilating NASA data as replacement parameters into the model. Had this happened, the utility of the NCEP/eta model as a global weather simulator would have been compromised. Figure 1 compares the dust patterns and concentrations of DREAM/SW (before) and *e*D/SW (after) assimilating NASA Earth science data. The *e*D/SW pattern matched more accurately the observed pattern in the visibility analysis shown at right.

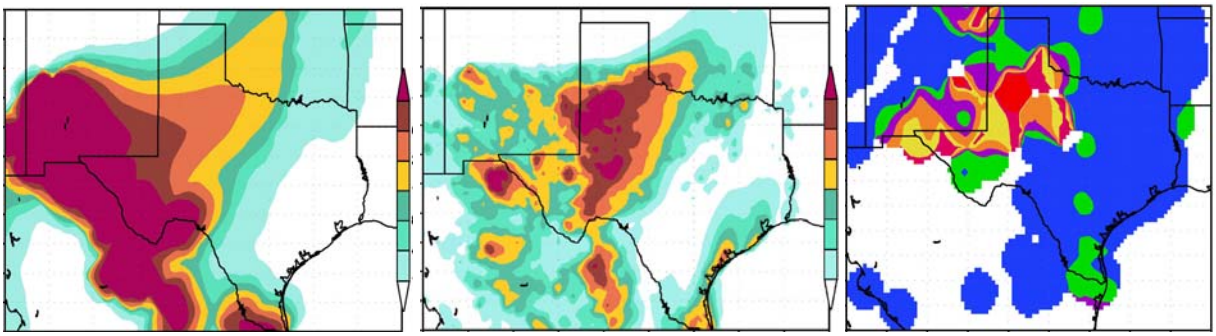


Figure 1. DREAM/SW (left) vs. eD/SW (center) dust patterns from model run 4a for the dust storm of 15-17 January, 2003. At right is a visibility map prepared from ground-based data using a Cressman analysis.

For model performance, improvements were realized by assimilating MOD12Q1 data into DREAM/SW. The peak hour correlation was least affected by this parameter replacement. However major gains were made in modeling the magnitude and duration of near-surface high dust concentrations. The *e*D/SW model forecasted the timing of two pilot dust storm events very well at almost all locations in the model domain, but had variable success in forecasting dust concentrations. This is encouraging for the future use of *e*D/SW simulations in public health alerts, at least for dust storm warnings in the Southwest.

What remains to be demonstrated is that there is a pattern of progressive improvements in *e*D/SW performance with each incremental parameter replacement and that there are statistically valid correlations between dust episodes and reported health outcomes. For the former, preliminary results suggest there is a performance plateau beyond which further replacements yield little measurable benefit. For the latter, collaborative statistical analyses of joint health and dust storm data are needed for etiological and epidemiological research; and/or public health and medical workers need to adopt a system like SYRIS™ for contextual information about dust episodes and syndromic reports.

Model performance, verification and validation took several forms: simple correlations using observed and hourly data, calculation of agreement indices, and application of tools to calculate model skill and threat scores. For model performance, simple correlations (peak hour and concentration) were used between hourly observed and modeled outputs to assess how well *e*D/SW predicted dust events. Results suggest that there are lags in model timing and concentration averaging, which help improve verification of model performance. Twenty-four-hour averages were compared to test the model's ability to predict exceedances of the EPA health standards for $PM_{2.5}$ and PM_{10}. The *enhanced* model performed better than DREAM/SW in reporting fewer EPA exceedances, and it had fewer false alarms.

Figure 2 is a plot for each station 24 hours before, during, and after a dust event on 4-6 January, 2007. The stations are plotted geographically from west on the left to east on the right. Southern California was hardest hit. Both the observed and modeled data show a spike in the dust gradient with the exception of Riverside, where no significant event was recorded by the AIRNow station. In Figure 3, a total of 443 hourly values were used to calculate the correlation of dust magnitudes between modeled and observed data. Correlation lines are skewed toward the modeled data axis, illustrating the model's tendency to over-predict dust magnitudes. Some improvement is indicated in the higher correlation from model run 15a to 20a.

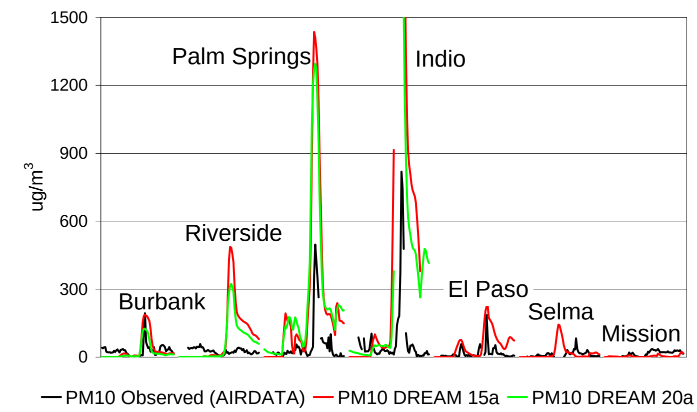


Figure 2. Modeled and observed PM_{10} magnitudes at seven AIRNow stations for 4-6 January, 2007 for model runs 15a (MOD12Q1 and AMSR-E) and 20a (REGAP).

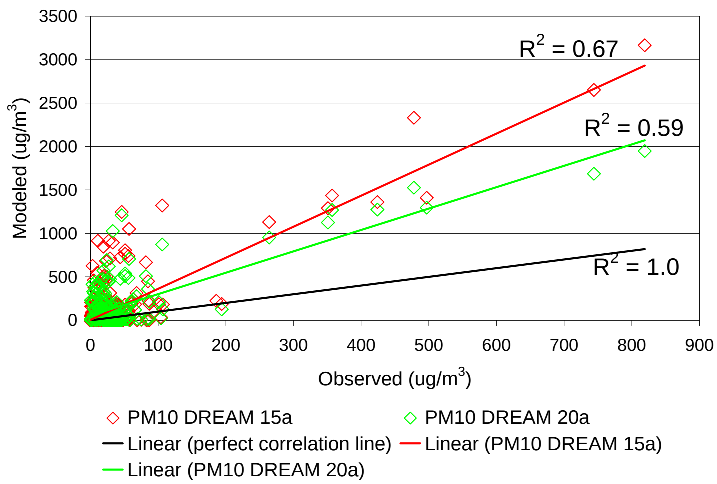


Figure 3. Magnitude correlation (R^2) for seven sites, (N = 443) during the 4-6 January, 2007 episode.

Metrics were defined to assess which parameter replacements led to improvements in *e*D/SW. These metrics relate to surface meteorological parameters important in dust entrainment and resulted in agreement indices between observed and modeled data sets. Table 1 lists the performance statistics. The biggest differences between DREAM/SW and *e*D/SW were for temperature at two meter height above ground where dust entrainment occurs. The agreement index after NASA data assimilation was 0.95, compared to 0.71 using the original DREAM/SW parameters. This is the single most significant improvement that led to the dust patterns shown in Figure 1 (above). Results suggest that there are lags in model timing and concentration averaging, which help improve verification of model performance. Twenty-four hour averages were compared to test the model's ability to predict exceedances of the EPA health standards for $PM_{2.5}$ and PM_{10}.

Table 1. Comparative performance statistics for DREAM/SW and eD/SW surface wind and temperature for Case 2. DREAM/SW values are in normal font; eD/SW values in bold font

Metrics	Wind Speed	Wind Dir. (degree)	Temp (K)	Definition M=modeled O=observed

Mean Observed	5.53	231.40	276.74	$\frac{1}{N}\sum_{i=1}^{N} O_i$
Mean Modeled	4.65 4.37	226.60 230.38	275.56 277.48	$\frac{1}{N}\sum_{i=1}^{N} M_i$
Mean Bias	-0.88 -1.16	-4.80 -1.02	-1.20 0.72	$\frac{1}{N}\sum_{i=1}^{N}(M_i - O_i)$
Mean Error	1.97 2.03	51.76 47.85	4.09 2.67	$\frac{1}{N}\sum_{i=1}^{N}\lvert M_i - O_i\rvert$
Agreement Index	0.74 0.75	0.74 0.76	0.71 0.95	$1-\frac{\sum_{i=1}^{N}(M_i-O_i)^2}{\sum_{i=1}^{N}\left(\lvert M_i-\overline{O}\rvert+\lvert O_i-\overline{O}\rvert\right)^2}$

Finally, a weather forecaster's approach was used to model verification using the WRF Model Evaluation Tools. Skill and threat scores were calculated in much the same way meteorologists predict rainfall patterns. These show great promise as *e*D/SW verification tools. Performance stats were calculated for two 2007 test cases using the Point-Stat tool. As indicated in Table 2, only ten 'hits', or exceedances, of the 'dust threshold' occurred over the entire model domain during these events.

Table 2. Modeled vs. observed hourly dust forecasts of dust concentrations over the entire eD/SW domain during the CA, AZ, NM, TX events of 4–6 January, and 23–25 February 2007.

Case	Fraction	N	Hits	Misses	False alarms	Non-events
Jan '07	pm10	27	2	0	3	22
Feb '07	pm2.5	267	4	0	15	248
Feb '07	pm10	52	4	4	2	42
Combined	pm2.5	267	4	0	15	248
Combined	pm10	79	6	4	5	64
Combined	Combined	346	10	4	20	312

The approach was modified using the Phoenix metro area. Using the combined values in Table 2 yielded the performance statistics shown in Table 3. These indicate that *e*D/SW successfully forecasted 71 percent of the hourly averages and 29 percent of the hours exceeding the dust threshold. These results indicate that *e*D/SW can be used in the same way weather forecasters predict the possibility of rainfall over metropolitan areas. The *e*D/SW had a threat score of 65 percent from January–April 2007. In an operational sense this suggests that the odds of correctly forecasting a dust event in the Phoenix metro area during that time period were two-out-of-three.

Table 3. Point Stat Evaluation – Phoenix Metro Area. POD = probability of detection; POFD = probability of false detection; TS = threat score; SS = skill score.

Accuracy	What portion of forecasts were either hits or non-events?	93%
POD	What fraction of events was correctly forecasted?	71%
POFD	What fraction of forecasted events did not occur?	6%
TS	What fraction of events was successfully modeled? (ignores non-events)	29%
SS	How well does the model discriminate between events and non-events?	65%

Benchmark

Given the promising results from *e*D/SW model runs, the project team is well satisfied that data replacements improve dust episode forecasting. The team and its public health partners are encouraged that these improvements will lead to more timely forecasts that will enable public health officials to issue early warning alerts and implement health interventions for populations at risk. The paragraphs below give the basis for this optimism.

SYRIS™

Key to mitigating disease epidemics is situational awareness, both before an outbreak occurs and during an outbreak. Physicians, veterinarians and their assistants see many of the first cases. A large number of professionals who, though not usually considered to be part of the "clinical" community, also see ill people or animals and collect related data, which is of great value to public health officials (PHOs). Emergency medical technicians, school nurses, animal control personnel, laboratory technicians, and medical investigators fit into this "other" clinical group.

The Syndrome Reporting Information System™ (SYRIS) is a JAVA-based, platform-independent system that runs on most PCs and laptops, and does not require a Web browser. SYRIS™ supports two-way disease information reporting and data sharing for these medical professionals. It provides a fast, reliable, portable method for reporting suspicious or novel symptoms that may be part of a known disease or disease-complex. Reporting is based on symptom complexes known as syndromes. These can be defined with a high degree of specificity (e.g., flu-like syndromes) or can be made more general, reflecting common medical care parlance. SYRIS includes the PHAiRS modeling and information system in its development client. This group of qualified, practicing health officials, doctors, nurses, and others will provide another form of validation to *e*D/SW outputs.

PHAiRS data system

From the outset of PHAiRS, the intention was to generate a modeling system that incorporated NASA Earth science data and that would enhance an existing public health decision support system. Three tasks were designed to generate an archive of *e*D/SW forecasts. These are true forecasts in the sense that they are not based on comparative data from AIRNow or other observation sites. Comparative data are not available for the forecast period because, by definition, they are not available before the event. They are used only for testing model performance *after-the-fact*, for verifying and validating the outputs, and as a supplemental data source for historical trend analyses. Figure 3 shows the data management and web services system.

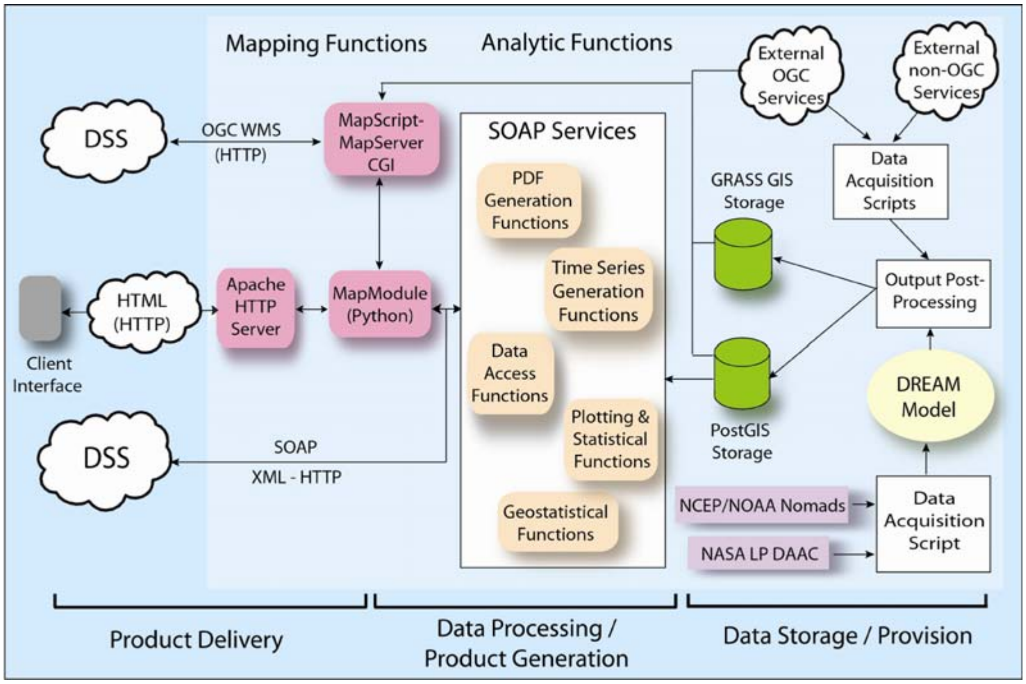


Figure 4. PHAiRS system architecture for providing, storing, processing, and delivering *e*D/SW dust event forecasts and animations; and for retrieving archival data for health research.

Public health users

Arizona Department of Health Services (Lea Trujillo):

As a syndromic surveillance epidemiologist, I am always searching for useful sources of data to track syndrome illnesses that I can add to my program. One of the problems with disease surveillance in general is that we do not know when and where events are going to take place and therefore we are reactive, not proactive. Another problem specific to syndromic surveillance is that with the non-traditional data sources commonly used in syndromic surveillance, there is no common user interface. We must use many different programs and softwares to visualize and analyze the data. Based on the demo we at Arizona Department of Health Services were shown, DREAM has the potential to add to existing data sources for syndromic surveillance. First, the dust storm model can help predict when and where respiratory illnesses are potentially going to increase, which is a much needed addition to disease surveillance tools. Being forewarned about the possibility of dust storm-related illnesses will help health officials better cope with the resulting illnesses. Second, the model seems simplistic enough to integrate into existing programs instead of requiring its own user interface and program. I understand that it will be possible to format this model to be added as an extra button/tab built into existing visualization systems. This aspect alone will increase the utility of the model for sydromic surveillance. If the PHAiRS program can help us prepare for events and be integrated into current program operations with such ease, it will be a very welcome and useful tool.

Pima County AZ (Beth Gorman):

The visualization of the data was an exciting way to see the numbers on a page come to life. It was especially intriguing to watch changes in the dust plume over time and from different perspectives. Our department is looking forward to continued coordination with PHAiRS and others to develop a method of forecasting airborne dust events to protect the many individuals who are at risk in our community.

The DREAM model visualization was quite interesting. It provided a virtual look at the formation of a dust event with indications of the originating area. I believe with some modifications it might prove useful in pinpointing sources of dust events which could prove useful in remediation.

City of Lubbock Health Department (Tigi Ward):

RSVP (Zelicoff et al., 2001) (now SYRIS) was the only active Syndrome-based Decision Support Systems (SBDSS) available for comparison to the passive systems. RSVP and SYRIS both define six common syndromes worded in the daily parlance of medicine and public health, and further provided an electronic interface that operated on virtually any computer connected to the Internet. It also provided primitive, but useful geographic mapping tools.

Their experience with RSVP was generally positive. Physician compliance was high (contrary to the popular, but incorrect belief that physicians will not take time to enter cases) because the number of cases of seriously ill patients who fit into one of the syndrome categories [is], on average, a case per month per physician (except during large epidemics). Further, RSVP provided information of immediate clinical importance to physicians thus increasing their cost-effectiveness in practice. Finally, on rare occasions, RSVP enabled public health officials to contact doctors within minutes of a case report when the data suggested unusually worrisome symptoms that might require immediate contact investigation. Thus, RSVP cut down the time from initiation of contact investigation from days to mere minutes.

Table of Contents

Cover Photo 2
Executive Summary i
The Project ii
Results ii
Benchmark vi
SYRIS™ vi
PHAiRS data system vi
Public health users vii
Table of Contents ix
Chapter I: Dust and Health 1
Atmospheric dust and health effects 1
Challenges for health surveillance 2
Linking dust with health 3
America's public health system 4
Need for decision support systems 5
Approaches to syndromic information systems 5
Chapter II: DREAM Models and Parameter Replacement 7
Terminology 7
PHAiRS project goals 7
DREAM/MED 8
Design 8
Performance 9
DREAM/SW 12
Parameter replacements 13
Dust sources 13
Topography 17
Aerodynamic surface roughness 17
Soil moisture 18
Aerosol optical depth 18
Chapter III: DREAM/SW Performance 21
Surface patterns 21
Upper air patterns 22
Profiles through the atmosphere 23

Surface wind and temperature 24
Modeled & storm-generated dust cloud 24
Point by point comparison 25
Peak hour and peak concentration ($PM_{2.5}$): 26
Dust episode duration: 26
Chapter IV *Enhanced* DREAM/SW (*e*D/SW) performance 29
ESR data assimilation 29
Model runs and comparative agreement indices 29
*e*D/SW Performance 30
Surface patterns 31
Upper air patterns 31
Profiles through the atmosphere 32
Surface wind and temperature 32
Modeled & storm-generated dust cloud 33
Point-by-point comparison 33
Peak hour and peak concentration (PM2.5) 33
Dust episode duration 34
Southwest REGAP experiment 34
Chapter V. Verification and & Validation 37
Approach 37
AIRNow data 37
Dust storm cases (2007) 38
Model statistics 38
The Point-Stat tool 43
Chapter VI: Data and Information Systems 49
SYRIS™ 49
PHAiRS system 50
*e*D/SW output archive 50
Data management and web services 50
Statistical measures 51
Systems integration 52
Chapter VII: Final Benchmark 55
Improvements to *e*D/SW 55
Improvements to SYRIS™ 56
Other health system improvements 57

NM Department of Health (EPHTS) ... 57
Arizona Department of Health Services ... 58
Pima County AZ ... 58
City of Lubbock Health Department ... 59
Model verification issues ... 59
New technology development ... 60
Data uncertainties ... 60
Replacement data sets ... 60
Dust speciation ... 61
Health data ... 62
Chapter VIII: Outreach & Transition ... 63
Engaging stakeholders ... 63
Collaborations at the national level ... 64
Testing the PHAiRS client ... 65
Training workshops ... 65
Lubbock test ... 65
Future opportunities ... 66
Opportunities in China ... 66
World Meteorological Organization ... 67
GEO/GEOSS ... 67
Media interest in PHAiRS ... 68
Publications and presentations ... 69
Chapter IX: Summary and Conclusions ... 71
V&V summary ... 71
References ... 75
Appendices ... 81
Appendix 1: Zelicoff / Forslund emails ... 81
Zelicoff – ARES Corporation, Albuquerque ... 81
Forslund – Los Alamos National Labs ... 81
Appendix 2: Lubbock beta test of RSVP ... 83
Introduction ... 83
Past Experience with SBDSS in Lubbock ... 84
Current Experience ... 85
Summary ... 86
Appendix 3: Publications ... 87

Appendix 4: Oral presentations and posters 89
Appendix 5: Terminology 93
Appendix 6: Acronyms 95

Chapter I: Dust and Health

Atmospheric dust and health effects

Observations from Earth-orbiting platforms cannot reveal a population's health, but interpreting environmental factors that impact health is routine. Once airborne, mineral dust and pollutants affect human health in many ways; most contribute to chronic and costly respiratory conditions while others are lethal. The first etiological evidence of a relationship between air quality and lung development in 10-18 year olds was reported by Gauderman et al. in 2004. It is one of a very few longitudinal studies connecting air quality and health. However, numerous studies have confirmed that satellite acquired data can detect moderate to severe dust events, and that dust can be traced in the atmosphere across continents and oceans (e.g. Lee, 1989; King et al., 1999; Prospero, 1999; Kaufman et al., 2000; Chu et al., 2003; Grousset et al., 2003; Gu et al., 2003; Miller, 2003; Kaya et al., 2004; Stefanov et al., 2003). Likewise, weather forecasting models, augmented with regional dust forecasting capabilities, show promise for better predicting the onset and tracking of dust events. Lastly, it is well known that naturally dusty environments exacerbate irreversible lung diseases (Policard and Collet, 1952; Bar-Ziv and Goldberg, 1974; Norboo et al., 1991; Goudie and Middleton, 2001; Xu et al., 1993; Mathur and Choudhary, 1997; Wiggs et al., 2003; Wright, 2005). Whereas lifetime exposures to atmospheric dust may result in silicosis or pneumoconiosis in high altitude or desert-dwelling populations, asthma is a global chronic respiratory disease triggered by numerous indoor and outdoor attributes. What remains to be demonstrated is that there is a direct coupling of specific dust episodes with health response statistics; and that such events can be forecasted and tracked in time to issue effective alerts.

Everyone's health is impacted by exposures to microscopic minerals, chemical particulates, by organisms bonded to dust particles, and by toxic gases (Figure 1). Airborne thoracic particles range in size from 10ųm to 0.01ųm, a size range that includes pollen, bacteria, viruses, and molecules. Only the coarse particle fraction ($PM_{2.5}$-PM_{10}) is being addressed in this report, but these particles serve as vehicles for potentially lethal concentrations of finer biological material. Many of these organisms are infectious and can become contagious throughout whole populations; others are patient-specific (Kuehn, 2006; Griffin, 2007).

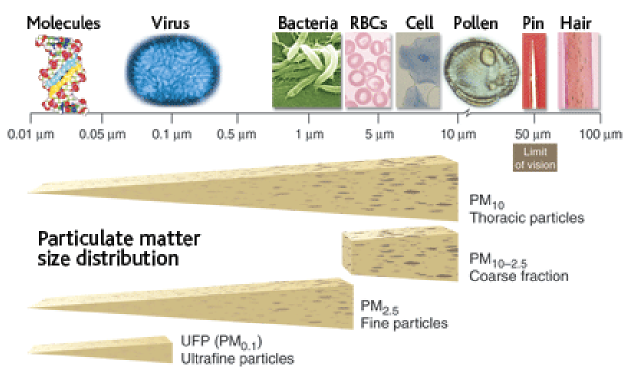


Figure 1. Medically relevant size distribution of atmospheric particulates. Source: *Kaiser, 2005*.

Asthma is a primary concern for the American College of Allergy and Immunology. It is one of the most common chronic diseases in the United States and is the most prevalent chronic disease in children. For poorly understood reasons, the rate of asthma among children in the northern mid-latitudes has more than doubled in the last 20 years. Asthma and acute myocardial infarction (MI) are among these. Asthma is a progressive disease that afflicted 20M Americans in 2003 (American Lung Association, 2005) and was a direct cause of death for an estimated 13-21M Americans between 1980 and 2004

(Moorman et al., 2007). Between 1980 and 1994, the prevalence of asthma in the U.S. increased 75%; while 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).

Nationwide, more than 9 million children struggle to breathe. According to the National Center for Health Statistics, asthma causes more missed school days than any other chronic condition, and is the leading cause of hospitalization for children under 15. Asthma is the most common reason that children younger than five go to the emergency room. Based on outpatient visits, the prevalence of asthma has increased by 50 percent over the last decade. During this time asthma fatalities have increased more than 80 percent. This increase in morbidity and mortality has not been uniform throughout the country (Moorman et al., 2007). Certain urban settings and minority groups have experienced higher rates of increase, but much about the disease is unknown, particularly in regard to incidence and prevalence in rural areas where access to care is limited and exposure to organic and inorganic dusts is inescapable. Table 1 illustrates a few of the more alarming rates for asthma visits and deaths over the last 2-3 decades.

Table 1. National Asthma Surveillance Data (Source: Moorman et al., 2007). Note: all data are based on asthma as the first-listed diagnosis.

Statistic	Reporting Period	Rate
Hospital discharges	1980-2004	15-20/10^4 patients
Physician office visits	1980-2004	270-470/10^4 patients
Outpatients visits	1992-2004	25-53/10^4 patients
Emergency visits	1992-2004	57-71/10^4 patients
Asthma deaths	1980-2004	13-21/10^6 patients

Economically, there is ample evidence that respirable particulates result in costly health effects. 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. These data alone give ample reason for forecasting outdoor dust events based on time series Earth observations.

Medically, the etiology of declining air quality and respiratory diseases in desert regions is poorly understood; but, patterns of rising health care costs are agreed in the health community 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 should reduce chronic diseases (cf, Ackerman, 2002; Landrigan et al., 2002; Pear, 2003; Massey and Ackerman, 2003; Jerrett et al., 2003; and, Davies, 2005).

Challenges for health surveillance

Health bulletins are posted every day on the Internet. Most concerns focus on containment through quarantines of animals or people, the fear of rapid global epidemics if containment fails, or the ability of public health authorities to treat sudden onset of epidemic situations. The challenge for this and following generations will be to develop highly reliable surveillance and monitoring systems that ensure detection and recording of individual cases even in remote parts of the world. Among the contributing challenges are: (a) developing technologies that permit quantitative assessment of population health status; and (b), developing technologies that allow assessment of individuals for multiple conditions or pathogens at points-of-care (Varmus et al., 2003). The stimulus for these developments is compelling because emerging infectious diseases (EIDs) in the 21st Century are overwhelmingly centered on the densely populated temperate zones of the Northern Hemisphere. In earlier centuries, EIDs were associated with the equatorial zone, especially tropical humid regions (Binder et al., 1999; Fauci et al., 2005). The fact that EIDs today can spread at the speed of international airlines underscores the need for coordinated, global surveillance and reporting systems.

Among the major challenges for integrating Earth science results (ESR) into human health practice is to demonstrate that these data improve model performance for forecasting dust episodes 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 this literature, it is increasingly clear to science and government that satellite observations should play a prominent role in forecasting short term weather episodes and longer-term environmental changes that cycle over several human generations. The opposite challenge is for medical scientists to extract from these capabilities the consequent flow of pathogens and chemicals through airborne mechanisms, and to translate findings into actionable human health interventions. This challenge implies adding health-care professionals into efforts that merge environmental surveillance with human health surveillance. Happily, the medical community already recognizes the adverse effects of PM_{10} and $PM_{2.5}$ in patients with respiratory conditions (Pope, 2004). What they lack are reliable forecasts of dust episodes that stimulate health alerts. Table 2 lists the world's five leading causes of death. Three of these (cardiovascular, infectious, and chronic lung) include a proportion of deaths actually caused by atmospheric dust, but for which there is no etiology for the role dust might played in these data.

Table 2. 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%)

Linking dust with health

In 2002, NASA implemented a plan to create integrated system solutions (ISS) for its application program elements. Its 2007-2011 plan for the public health program element is arguably the most aggressive attempt to link dust with health outcomes. Its 2002-2008 roadmap aims to use ESR in weather models to integrate "public health surveillance [that can] track weather-climate-environmental factors to improve [detection of] disease outbreak[s], predictions, and increased warning times (NASA, 2006)." This plan is predicated on ESR from any of a suite of satellite sensors that provide observations or products that support decision support tools; or, that could be integrated with Earth system models to enhance their performance in routine applications. The Public Health Applications in Remote Sensing (PHAiRS) solution was designed to integrate ESR products into an atmospheric modeling system that would forecast dust events and magnitudes, and that would improve respiratory health decisions (Figure 2).

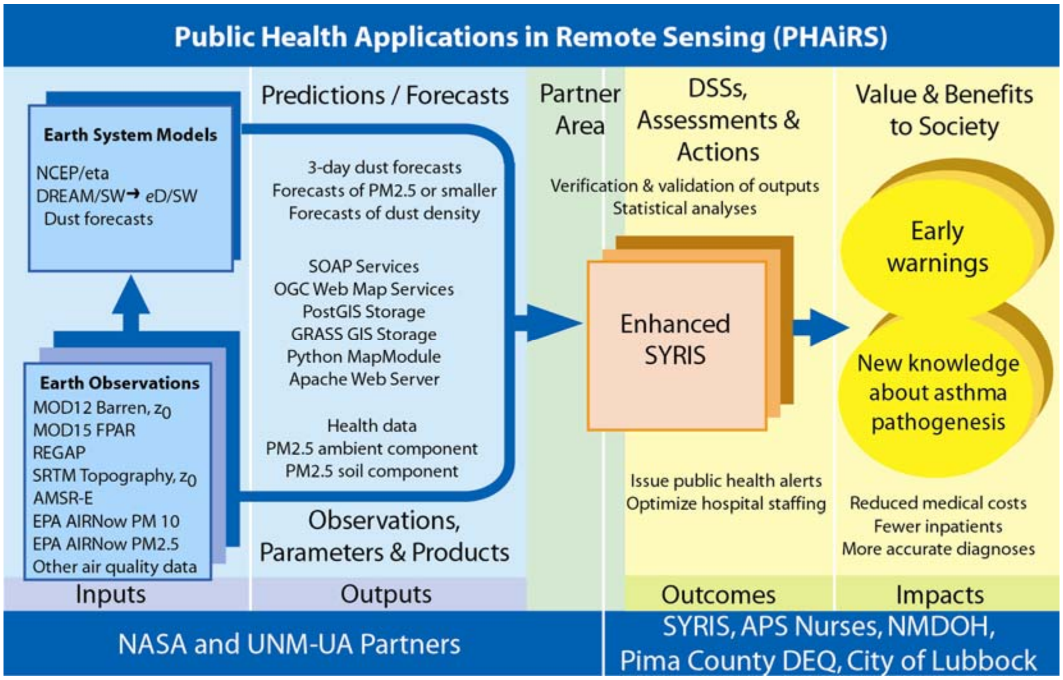


Figure 2. PHAiRS approach to an Integrated System Solution (ISS) for respiratory health.

In PHAiRS, the framework for coupling atmospheric dust processes with respiratory health responses began with experimental NASA satellite data products and modified them for assimilation into the *Dust Regional Atmospheric Model* modified from its home domain in the Mediterranean region (DREAM/MED) to a new domain over the American Southwest (DREAM/SW). The output from the enhanced DREAM/SW model (*e*D/SW) became input to the PHAiRS web-based product and data management system. This system is accessible to both medical and health providers who desire additional corroborating information about similar cases reported by their local or regional colleagues. The ultimate goal was to serve geospectral and geospatial data and information to medical and health professions for decision making and for epidemiological research.

America's public health system

"The heart of the public health system [in the United States] is comprised of over 3,000 Local Health Departments. Almost 96 percent are in small cities, towns, and rural areas that serve fewer than 25,000 people. It is here that public health decisions are most likely to affect the public's health (Parsons, 1997)." These circumstances represent a powerful motivation for health departments to transition to better, faster, and cheaper ways of making decisions. At the local level, which is where all public health decisions are made, departments are always under-staffed and under-funded. Nevertheless they deliver essential public health services through surveillance, health education, and prevention (Anon., 1997). Furthermore, electronic syndrome reporting systems must serve many purposes because some health issues are contagious or communicable, and others are case specific. There is a need to develop reporting systems that start with a few well-known syndromes observed at local and regional levels, and progress to those that are national and global. Systems should have rigorous geospatial content for tracking the frequency and distribution of cases, and be supported by interoperable health data through statistical packages and products that facilitate decisions.

Need for decision support systems

An article in *Vaccine* (Oxford et al., 2005) hypothesized that the pandemic flu of 1918-19, which killed more than 50 million people worldwide, began at the end of WW-I in an overcrowded British base camp surrounded by livestock markets for pigs, horses, geese, ducks, and chickens. The flu virus probably jumped rapidly from avian forms to pigs and/or horses and thence to humans, and spread over the following 18 months as millions of men returned home from the war. A far less devastating event (SARS) occurred in the chicken markets of Hong Kong in 2003-4. Detecting and reporting individual cases is essential to monitor and contain such viruses in today's world of rapid transit. Clearly, there must be a rapid reporting system based on modern computing technology. Science and technology have always supported the practice of medicine; but, appropriate technology for reporting health syndromes is exacerbated by paper-based rather than digitally-based information systems that would expedite alerts (Lohr, 2008). According to Zelicoff and Bellomo (2005), current public health reporting systems are "exquisitely designed for failure" because they rely on doctors filling out tedious paper forms and/or relying on the instant availability of public health officials by telephone. Neither of these means of communication helps a doctor to learn whether there are similar syndromes being reported nearby, or whether his/her patient is in a life-threatening situation. Electronic reporting systems have been prototyped only in the past few years (Centers for Disease Control and Prevention, 2005).

Approaches to syndromic information systems

At least two approaches to health information systems are being developed. One is based on medical reporting by doctors, clinicians, nurse practitioners, school nurses, first responders and others; the other is based on electronic information gathering through data mining and statistical analyses from historical reports and current medical record databases. Both approaches rely on "syndromic surveillance," the ability to detect outbreaks of illnesses earlier than disease-specific reporting systems, and with sensitivities to detect outbreaks that might otherwise be missed (Hadler et al., 2005). As expressed in the exchanges between Drs. Zelicoff and Forslund, each approach has its advantages and disadvantages (Appendix 1). The PHAiRS project focused its development efforts on the Syndrome Reporting Information System (SYRIS), a clinician-driven approach. It was selected because its predecessor, the Rapid Syndrome Validation Project had already been tested by a public health coordinator in Texas, and because it addresses human resource issues, timeliness and accuracy, and cost effectiveness (Appendix 2).

Rapid Syndrome Validation Project (RSVP):

In 2002, RSVP was a novel prototype for decision support systems aimed at modernizing health care reporting. It was an Internet-based syndromic surveillance system designed to facilitate rapid communication between epidemiologists (public health officials in local jurisdictions) and health care providers (physicians, physician assistants, and nurse practitioners). It was a reporting and discovery system for primary care physicians and clinicians who wanted to determine if their patient's syndrome had been reported by others in their jurisdiction or surrounding area. It provided medical and environmental information in a geospatially explicit architecture in three modules: (a) a syndrome information collection module whereby doctors could submit an inquiry, (b) a communication module whereby a public health official could respond to an inquiry; and (c), a data visualization module that permitted both parties to review collective inputs in the medical and geographic domains.

The prototype system was successfully beta-tested for six syndromes in several states in the U.S. and internationally. Beta testers expressed a universal desire for more visualization tools, especially those of a geospatially explicit kind. Based on this experience, the PHAiRS team partnered with developers of

RSVP to insert an imagery and geospatial module into which outputs could be placed and made available via the Internet.

Syndrome Reporting Information System (SYRIS):

Much of the SYRIS^{TM} material throughout this report has been extracted from the SYRIS *Manual*, and is used with permission. Use of SYRIS^{TM} for this purpose does not constitute endorsement.

SYRIS is an advanced, fully developed version of RSVP designed for commercial applications. RSVP was well accepted by physicians and public health officials where it was tested, but suffered because it lacked multi-community surveillance to be widely implemented. SYRIS represents an operational, highly-secure system for syndrome-based disease reporting that offers high-fidelity information with minimal false positives and at least two dramatic successes in ruling out a bioterrorism threat and in early detection of influenza. It captures the clinical and professional judgments of physicians, veterinarians, nurses (especially school nurses), coroners and medical investigators, emergency medical response teams and ambulance services, animal control, environmental health, clinical laboratory chemists, microbiologists, immunologists, and wildlife rehabilitators.

When clinicians see a seriously ill patient with presumed infectious disease, it takes less than 20 seconds to report that case via SYRIS. National studies suggest that this is less than 0.1 percent of all clinical encounters in human medicine (perhaps slightly higher in veterinary medicine), but these cases are precisely the ones needing to be identified in near real-time to avoid possible epidemics. SYRIS contains summaries and analyses from local public health officials that focus on diseases of importance in a particular geographic area that can be accessed with a single click of the mouse.

Experience with properly designed active, clinician-driven surveillance systems demonstrates that these health professionals will report cases of suspected infectious disease, if the system is fast (less than 15 – 30 seconds), provides immediate feedback to clinicians on local infectious disease outbreaks, permits selective interaction between public health officials and clinicians on a real-time basis as warranted, and which is inexpensive. SYRIS meets all of these criteria. Moreover, unlike the “passive” or “data-mining” approaches, SYRIS has a low false-positive rate (thus mitigating the investigation of a large number of false alarms) while at the same time facilitating enhanced relationships between local public health officials and all health care providers.

Chapter II: DREAM Models and Parameter Replacement

Terminology

Chapters III through IX of this report describe necessary activities in the integrated system solution for PHAiRS. Figure 3 shows these as discreet steps in time to achieve the system solution. To keep the reader oriented, the following terms are used in this report:

"DREAM/MED" refers to the original dust entrainment model developed at the University of Malta by Slobodan Nickovic and his team for use in the Mediterranean region.

"DREAM/SW" refers to the DREAM/MED model after it was adapted for the PHAiRS domain over the southwestern United States (see Figure 4)

"Enhanced DREAM/SW" (*e*D/SW) refers to a series of DREAM/SW model outputs after NASA ESR data sets were assimilated to replace DREAM/SW parameters.

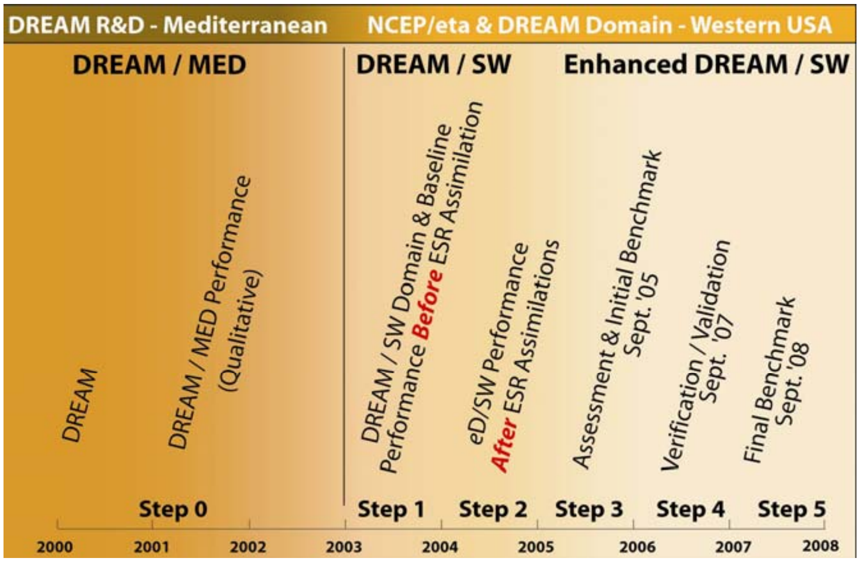


Figure 3. Steps in the evolution of the PHAiRS' integrated system solution for public health.

PHAiRS project goals

The project had three goals. The first focused on assimilating satellite data from MODIS Terra and other sources to create an *enhanced* DREAM/SW (*e*D/SW) model. The aim of this effort was to: (a) verify that advanced satellite measurements could replace DREAM/SW parameters in its new domain; and, (b) validate that parameter replacements led to improved model simulations of dust episodes.

The second goal focused on iterating *e*D/SW model runs using combinations of replaced model parameters to assess incremental improvements. The questions of interest were: (a) how well and to what degree of sensitivity could *e*D/SW simulate dust entrainment and transport? (b) how well could *e*D/SW simulate the speed and direction of moving dust clouds? (c) could sound evidence be generated coupling dust episodes to observed respiratory health responses? and (d), could areas affected by dust clouds be simulated in time to alert health officials and populations at risk?

The third goal involved establishing collaborative relations with public health authorities to assess whether there are statistically valid relationships between dust episodes and increased respiratory complaints. Since it is not often clear what specific exposures and durations led to observed outcomes, developers of SYRIS designed a reporting system that encourages public health officials, air quality monitoring offices, doctors, and clinicians to coordinate their information electronically, and in appropriate ways, to protect patient confidentiality. The system allows group attributes to emerge in geospatially explicit ways that populations at risk could be forewarned. For effective application of satellite observations in this system, physicians and clinicians must be motivated to report non-confidential patient information in ways that reveal emerging spatial and temporal patterns that can be recognized by authorities early in the development of an episode. For this to happen, satellite-based dust simulation models must, in turn, be perceived as a reliable source of information.

The ultimate objective of the project was to contribute to an improved public health decision support system that could 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 project was designed to build the scientific and technological underpinnings of these near-future capabilities.

Figure 4 shows the model domain for DREAM/SW and *e*D/SW. It is large enough to encompass the southwestern United States, centered on 109°W, 35°N.

Figure 4. DREAM/SW and eD/SW model domain.

DREAM/MED

Design

DREAM/MED was developed for use in Europe and North Africa, and was tested using National Centers for Environmental Prediction, eta version (NCEP/eta) data provided by the European Center for Medium-Range Weather Forecast (ECMWF) (Nickovic et al., 2001). NCEP/eta is based on large-scale numerical solutions controlled by conservation of integral properties. It uses a non-linear horizontal advection numerical scheme that preserves energy and squared vorticity and controls non-linear energy cascade. In the eta vertical coordinate, which generates quasi-horizontal model levels, topography is represented by step-like elements (Mesinger et al., 1988). Physical parameterization includes land surface processes, turbulent mixing, convection, large-scale precipitation, lateral diffusion and radiation. Horizontally, the model uses a semi-staggered Arakawa E grid (Arakawa and Lamb, 1977). The E grid spacing between neighboring mass (h) and wind (v) points is 1/3 degree. To take advantage of the higher spatial resolution of ESR land cover data, this spacing was reduced to 1/9 degree. The eta surfaces are quasi-horizontal in both mountain and non-mountain areas. From sea level to 100 hPa (geopotential height), there are 24 half-eta levels. This numerical model provides operational global weather forecasts, but it does not forecast dust.

DREAM is specifically designed to simulate dust entrainment and transport on a regional scale. To simulate these events, it has been nested within NCEP/eta. It is a eulerian model designed for studying atmospheric phenomena connected to transport of air pollutants. Its two major components simulate meteorological fields and terrain fields that govern the dust cycle (Janjic, 1984; Mesinger et al., 1988; Janjic, 1994). The concept behind DREAM/MED is shown in Figure 5.

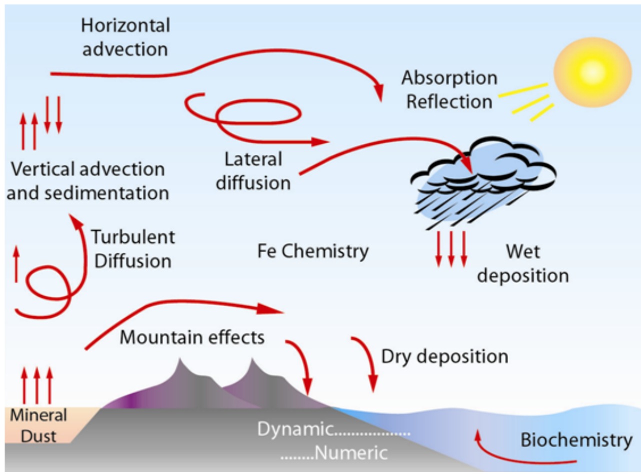


Figure 5. Conceptual model for dust entrainment and transport

DREAM/MED simulates dust entrainment, advection, turbulent diffusion, and both wet and dry deposition (Nickovic et al., 2001; Shao et al., 1993; Georgi, 1986). Its surface terrain parameters (Table 3) are all static, consisting of: (a) texture classes at 2x2 minute resolution; (b) land cover at 10 minute resolution; and (c), topography at 1x1km resolution. DREAM/MED uses soil texture derived from 134 categories of global soil types that were subsequently grouped into nine Zobler soil categories (Zobler, 1986), and converted into sand/silt/clay texture categories that characterize the physical properties of wind-blown dust (Cosby et al., 1984). Dust particles were grouped into four bins according to particle size (0-3.4µm, 3.4-12µm, 12-28µm, and over 28µm). The Olson World Ecosystems (OWE) land cover scheme containing 59 land cover categories has only one "barren class" to estimate dust sources.

Table 3 Surface terrain parameters in DREAM/MED and DREAM/SW

Data Set	Purpose/Properties
ECWMF medium-range weather forecast	Initial & boundary conditions; Res. = 1°
USGS GTOPO30 terrain data	Res. = 1km
Olsen World Ecosystems (OWE)	Land cover; Res. = 10min.; Dust categories = 8, 50, 51, 52
FAO World Soil Map	Res. = 2min.; 134 categories reduced to Zobler/Cosby categories for soil texture

Performance

DREAM/MED's performance was assessed over its Mediterranean domain 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 ground-based 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. Evaluations were performed also to assess the horizontal spread and vertical structure of simulated dust by comparing model outputs with patterns observed by SeaWiFS and as measured by ground-based LIDAR and sun photometers in the region. The results indicated qualitative agreement with SeaWiFS satellite images and LIDAR height-time displays over Barcelona. Dust was present mainly in the 1-5km altitudes, affecting most of the Iberian Peninsula and extending into West-Central Europe.

Passive satellite sensors only show horizontal 2-dimensional features of dust plumes that often are undetectable over continents because sensors cannot distinguish easily between the color of atmospheric aerosols and surface background reflectance. This is especially true 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.

Figures 6 through 10 show the horizontal spread and vertical structure of dust plumes originating in the Sahara. Figure 6 shows the horizontal spread of dust on 14 and 18 June, 2002. Figure 7 illustrates an encouraging agreement between the modeled vertical structure and the observed vertical profiles over Barcelona on 18-19 June, 2002. Figure 8 compares modeled and observed Aerosol Optical Depth (AOD) over Arenosillo, Spain. Figure 9 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 reportedly 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 10 is a SeaWiFS image showing North African dust transport into Eastern Europe and comparing the pattern to the native DREAM/MED output. These qualitative comparisons provided adequate encouragement for the model to be adapted for use by the PHAiRS project.

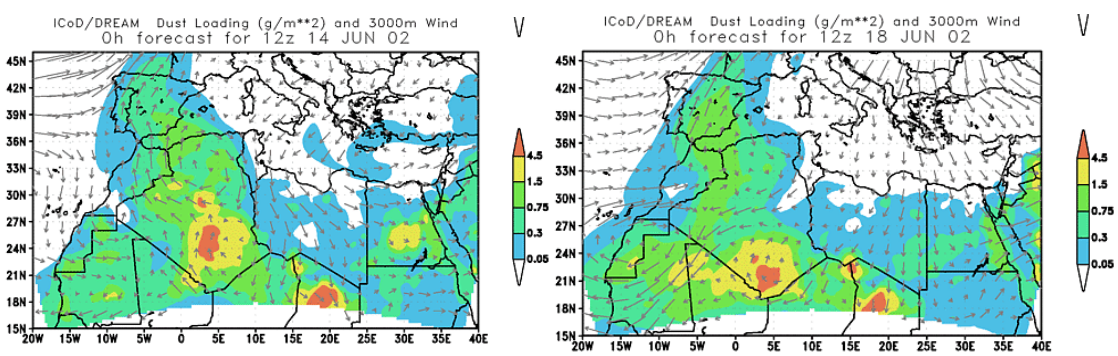


Figure 6. Horizontal spread of a Saharan dust storm. (left): modeled dust loading and winds at 3000m for 14 June, 2002; (right) DREAM/MED output for 18 June, 2002 (Perez et al., 2006).

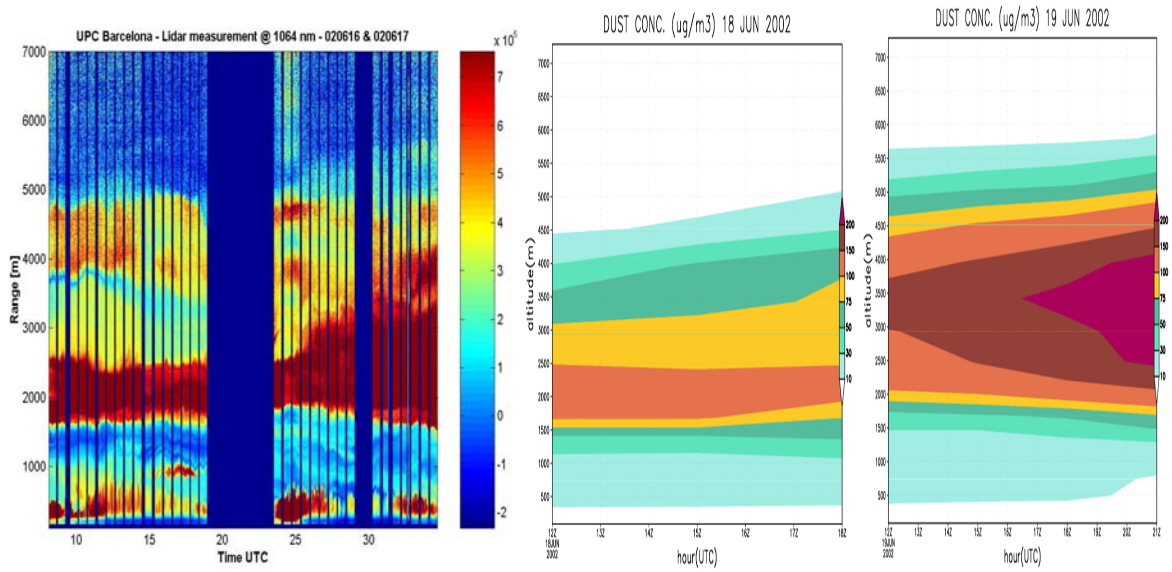


Figure 7. Vertical structure of a Saharan dust storm over Barcelona, Spain, 18-19 June, 2002. (left): LIDAR measurements. Dark blue columns indicate no measurements on the range corrected 1064nm signal (arbitrary units; temporal resolution is 60 sec.); (right) DREAM/MED vertical dust concentration (Perez et al., 2006).

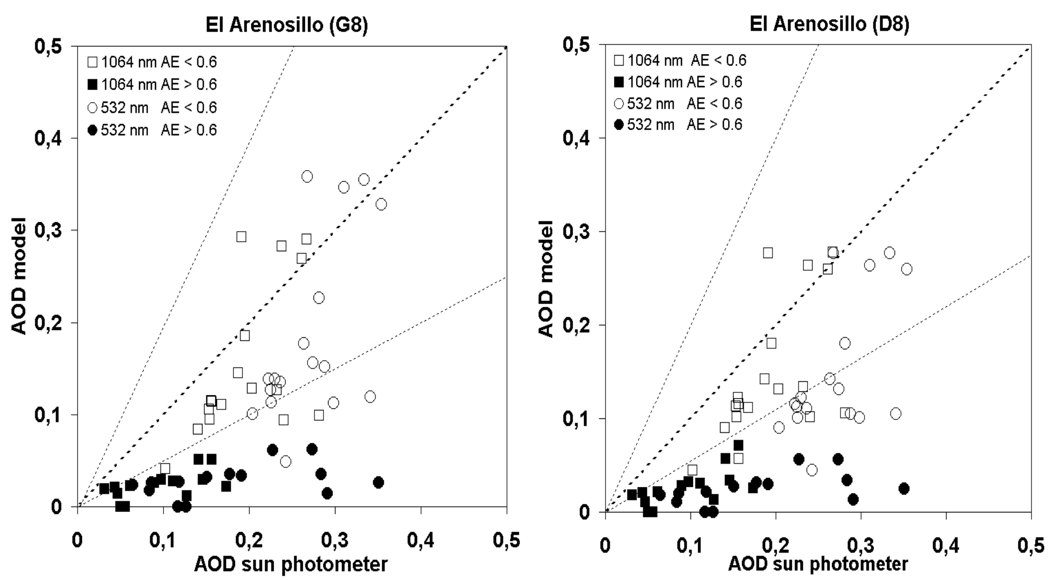


Figure 8. Modeled vs observed Aerosol Optical Depth (AOD) for G8 (left) and D8 (right) at El Arenosillo, Spain (Perez et al., 2006).

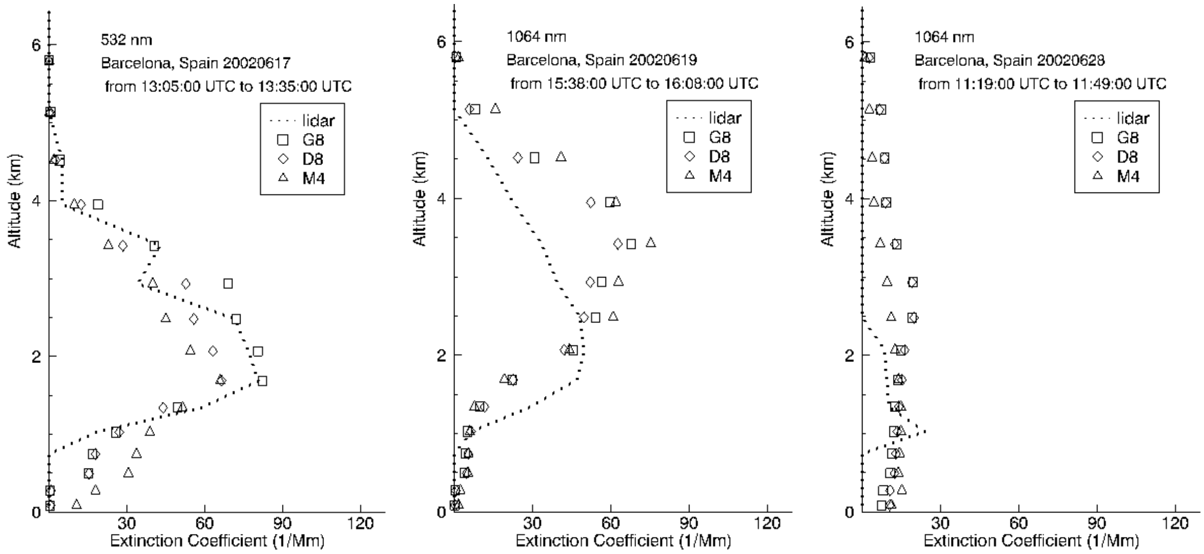


Figure 9. Comparisons of modeled and observed vertical profiles in the extinction coefficient for M4, G8 and D8 over Barcelona, Spain. (left) 532nm on June 17, 2002 between 13:00 & 13:35 hrs; (center) 1064nm on June 19, 2002 between 15:38 & 16:08 hrs; (right) 1064nm on June 28, 2002 between 11:19 & 11:49 hrs UTC (Perez et al., 2006).

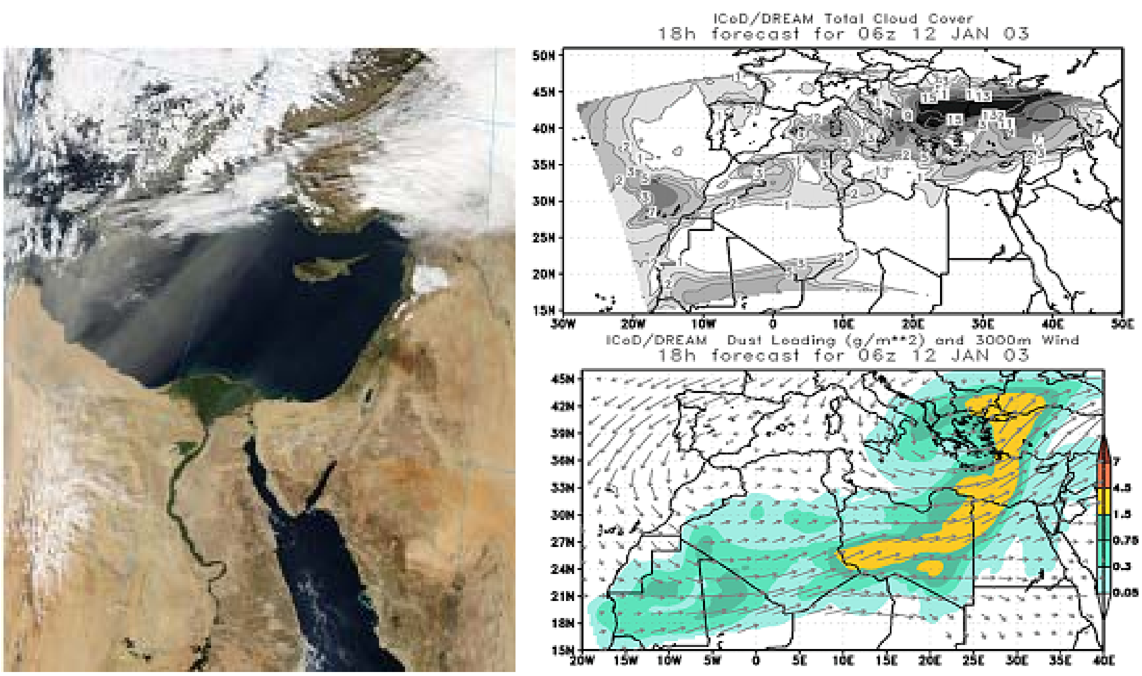


Figure 10. Mediterranean dust transport on 12 January, 2003. (left): SeaWiFS image of dust clouds over the Mediterranean; (right): DREAM/MED output showing dust clouds in orange (Nickovic et al., 2004).

DREAM/SW

Because of its European heritage, DREAM/MED had to be adapted for use in the southwestern United States. Its performance in this new environment also had to be tested and validated using observed weather patterns and dust events before it could assimilate NASA ESR data. The same NCEP/eta numerical model used in DREAM/MED was retained for configuring DREAM/SW. One of the major differences of the two domains is the absence of a large body of water (the Mediterranean Sea) as seen in Figure 10. The absence of water in the DREAM/SW domain, except on the west coast of California and Oregon, restricts use of qualitative visual comparisons of dust plumes moving eastward out of California. Instead, ground-based networks of dust monitoring sites have been substituted allowing more quantitative comparisons, but from only scattered observations.

Parameter replacements

Table 4 is the list of products prepared for assimilation into DREAM/SW. These data sets were selected to replace model surface parameters to achieve finer landscape resolution (land cover and topography), and more dynamic temporal resolution. Moderate Resolution Imaging Spectroradiometer (MODIS) products (MOD12Q1 and MOD15) were selected as replacements for land cover to locate dust source areas. Topography derived from the Shuttle Radar Topography Mission (SRTM) was selected to replace the horizontal steps formulated by Mesinger et al. (1988); and aerodynamic surface roughness length ("z_o") was added to *e*D/SW to estimate a complex attribute of surface/air dynamics that leads to dust entrainment.

Table 4. Initial products for testing improvements of input parameters for DREAM/SW.

Name	Sensor/Product	Original Format	Time Period
Dust sources	MOD12 barren land	HDF	Dec 03
Dust Sources	MOD15 FPAR	HDF	Dec 03
Topography	Space Shuttle	GTOPO30	May 06
Surface Roughness Length	MOD12	Table look-up	Dec 03

Dust sources

Obviously, dust modeling requires knowledge of the changing locations of dust sources through time. Ideally, the pattern of dust sources should be monitored continuously to account for seasonal agricultural practices, large construction and housing developments, the effects of drought, and gradual changes attributable to climate change. Even though ESR data are collected daily, products showing dust sources are derived from complex algorithms that, at best, have been available only after years of effort. There is a high level of uncertainty in translating satellite radiances that are labeled as barren ground into "dust sources," in both the spatial and temporal domains. Nevertheless, PHAiRS has been able to demonstrate that even periodic updates of dust sources improve model performance.

MOD12Q1:

Patterns of land cover are important for *e*D/SW to identify dust source areas. PHAiRS assimilated MOD12Q1, which was created in 2001. It provides high temporal and spatial resolution compared to the 1970's-80's OWE data. It has 17 classes of land cover defined by the International Geosphere-Biosphere Programme (IGBP), five of which are relevant to dust entrainment: open shrublands; grasslands; croplands; urban and built-up; and, barren or sparsely vegetated land. Figure 11 is a comparison between barren categories identified by the OWE and MOD12Q1.

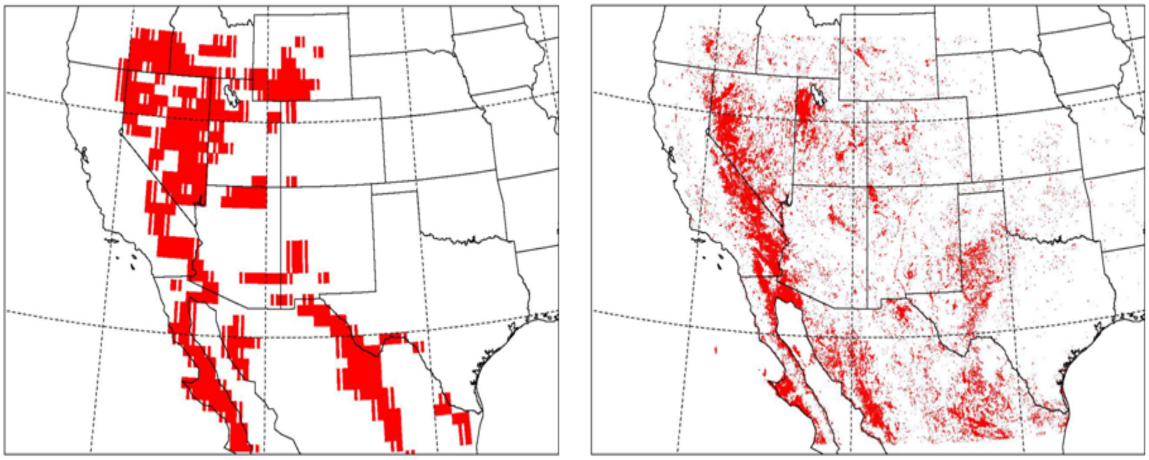


Figure 11. Comparison of the Olson World Ecosystem (OWE) barren category (left) with MOD12Q1 barren categories in the DREAM/SW domain (right). Note the major differences throughout the SW region, especially in west Texas and Mexico.

Land cover categories for MOD12Q1 were produced by NASA's 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 decision tree algorithm while systematically varying the training sample. The training sample is modified through iteration to focus the algorithm on samples that are difficult to classify correctly. This modification is performed by providing a weight for each training sample. 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.

Boosting provides an additive logistic regression that improves probabilities of class membership. These probabilities are used to assess the confidence of the classification results, as well as to incorporate prior probabilities to improve discrimination of cover types that are difficult to separate in the spectral domain. In addition to the classifications and boosting, MOD12Q1 also provides quality control information about whether each pixel has been newly classified or is dependant on a persistent value. Each measurement contains an embedded land/water mask as bit flags in the 8 bit Land_Cover_Type_QC parameter. The major drawback to MOD12Q1 is that values are not updated on a seasonal or annual basis. Since Earth observation data are considered the best way to provide up-to-date, fine resolution dust sources, other data sources were investigated to assess their possible contribution for providing a seasonal update of dust sources. Two candidates were identified: MOD15 FPAR; and, the Southwestern Gap Analysis Program (GAP) classification, which is derived from 30 meter Landsat data.

MODIS FPAR:

Another product (MOD15) is the Fraction of Photosynthetically Active Radiation (FPAR). This was used experimentally to augment identification of dust source areas provided by MOD12Q1. Both products were processed into an ASCII GRID format for input into DREAM/SW.

Visual comparisons of the MOD12Q1 and MOD15 products to commercial 1-meter resolution satellite imagery for sites ranging from California to west Texas suggest that MOD12Q1 over-estimates and MOD15 under-estimates the area of possible dust generation. For example the MOD12Q1 product seems to identify small (~1km) dust source areas where there may be none, especially in eastern New Mexico and west Texas. Both products seem to show credible patterns in the larger dust source areas shown in Figure 12. A benefit of considering MOD15 is its higher temporal resolution (every 8 days). The data set used for the comparison in Figure 12 (right) is from June 2005; MOD12Q1 (center) was last updated in 2001.

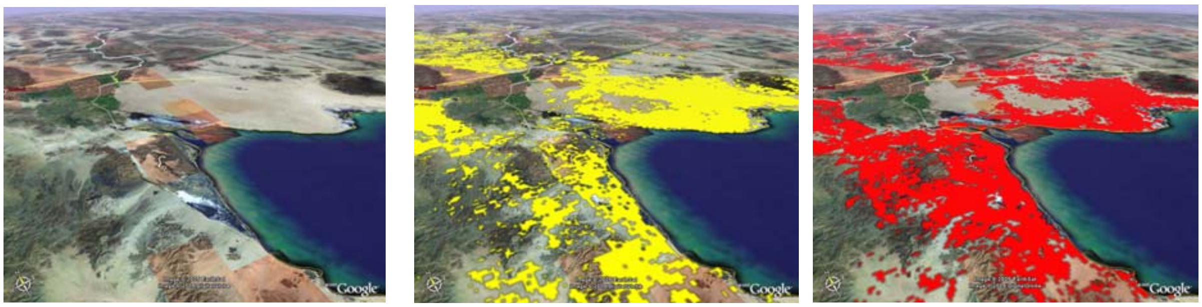


Figure 12. (left) Digital Globe image at the head of the Sea of Cortez (CA, AZ, Mex.) © Digital Globe; (center) MOD12Q1, Class 16 – "barren or sparsely vegetated"; (right) FPAR class 253 – "barren, desert, or sparsely vegetated".

To test the utility of MOD15, data were assimilated as a proxy seasonal update to MOD12Q1. The idea was tested over the White Sands National Monument in New Mexico by substituting MOD12Q1 pixel values with FPAR class 253 values. Qualitative comparisons between the two showed that MOD12Q1 missed several areas of dunes, while correctly classifying surrounding transitional areas. Moreover there were complications arising from desert areas covered by winter snow. Evaluation of class 253 for December 2003 and July 2005 (seasonal opposites) led to a question about whether the FPAR data were being updated along with other non-fill classes. The relationships between MOD12Q1 and MOD15 appear to be complex, and as a consequence, the MOD15 experiment was only used in one model run.

REGAP:

GAP is a national program of the USGS Biological Resources Division (BRD) that maps the distribution of plant communities and selected animal species, and compares these distributions with land stewardship to identify potentially endangered sites. GAP uses Geographic Information System (GIS) technology to analyze biological and land management data to identify areas (gaps) where conservation efforts may not be sufficient to maintain diversity of living natural resources. Previous projects were conducted by individual states. Differences among adjacent states in the methodologies and data used for land cover mapping commonly yielded surprisingly inconsistent vegetation and animal habitat maps. As a means for reducing state boundary discontinuities, a second effort was conducted using eco-regional units as the basis for classifying imagery. The program's first multi-state effort included Arizona, Colorado, Nevada, New Mexico, and Utah. It became know as the Southwest Regional Gap Analysis Project (REGAP).

REGAP included a barren category containing 20 subclasses. These are listed in Table 5. Classes in bold font indicate those most likely to be susceptible to dust generation. The other classes are primarily bare rock, canyon walls, and talus slopes considered to be too small to contribute dust. Before direct comparisons between these new classes and those from MOD12Q1 could be made, it was necessary to include the equivalent classes in the California GAP classification. Since California was not included in REGAP, categories had slightly different names and numbers than those generated in REGAP. A map of these two datasets showed obvious mismatches in color and edges between California and the REGAP states (Figure 13, left). The California GAP classes were analyzed, regrouped, and renumbered so that they matched those of REGAP. The resulting mosaic (Figure 13, right) shows a much better agreement between the two data sets.

Table 5. REGAP barren category sub-classes. Classes in bold font are possible dust sources.

North American Alpine Ice Field	**Inter-Mountain Basins Active & Stabilized Dune**
Rocky Mountain Alpine Bedrock & Scree	Inter-Mountain Basins Volcanic Rock & Cinder Land
Mediterranean California Alpine Bedrock & Scree	**Inter-Mountain Basins Wash**
Rocky Mountain Alpine Fell-Field	**Inter-Mountain Basins Playa**
Rocky Mountain Cliff & Canyon	North American Warm Desert Bedrock Cliff & Outcrop
Sierra Nevada Cliff & Canyon	**North American Warm Desert Badland**
Western Great Plains Cliff & Outcrop	**North American Warm Desert Active & Stabilized Dune**
Inter-Mountain Basins Cliff & Canyon	North American Warm Desert Volcanic Rockland
Colorado Plateau Mixed Bedrock Canyon & Tableland	**North American Warm Desert Pavement**
Inter-Mountain Basins Shale Badland	**North American Warm Desert Playa**

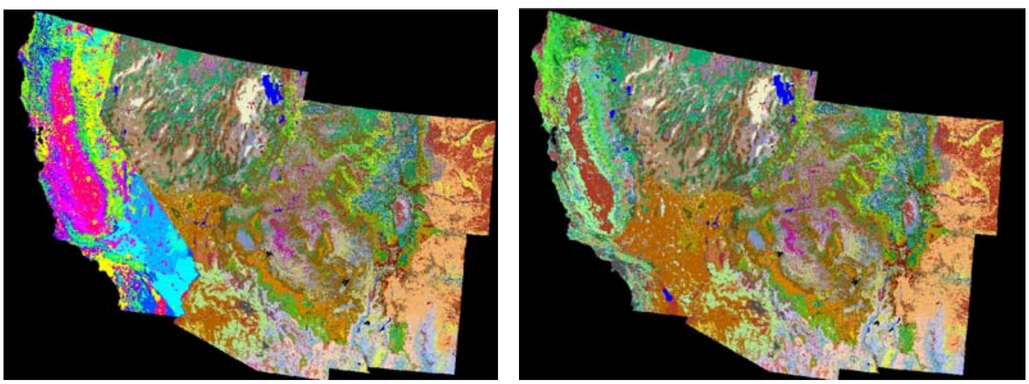


Figure 13. (left) REGAP classes showing edge effects between California GAP and SW REGAP; (right) REGAP classes after adjusting California GAP to REGAP categories.

Barren classes from MOD12Q1 and from REGAP were compared. The REGAP data were reprojected to the MOD12Q1 data and trimmed to match the outer polygon of the REGAP data. The barren classes for each image were highlighted in red and all other classes were colored gray (Figures 14). The two sets of barren classes are somewhat similar but certainly not identical in distribution and in overall size, but there are significant differences along the California/Nevada/Arizona border area, the Salton Sea area in southeastern CA, and in numerous small areas of AZ, NM, and CO.

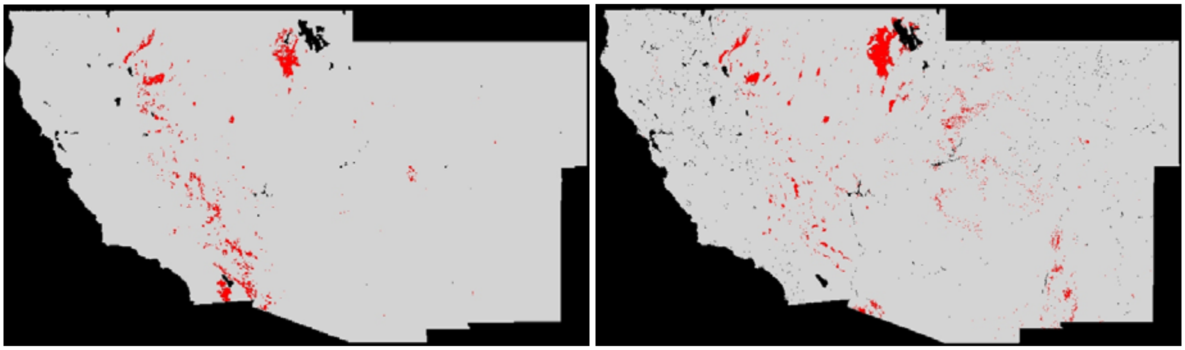


Figure 14. (left) MOD12Q1 barren class (red); (right) REGAP barren classes (red).

Topography

NASA Shuttle Radar Topography Mission (SRTM) data provided a high resolution data set for topography. The Mission generated a nearly complete high-resolution digital topographic database of Earth. It consisted of a specially modified radar system onboard Space Shuttle *Endeavour* in February 2000. The payload consisted of two radar antennas: one located in the Shuttle's payload bay; the other on the end of a 60-meter mast extending from the payload bay. Thirty arcsec [1km] data, released in May 2006, were configured; but, before they could be assimilated, data values had to be inserted to create a contiguous data set with no spikes, wells, or large voids. Voids are caused by geometric artifacts such as specular reflection off water, phase unwrapping artifacts, and complex dielectric constant (Dowding et al., 2004). For the DREAM/SW domain, the primary concern focused on small “salt and pepper” voids consisting of pixels having no signal response, and larger voids representing areas of contiguous pixels. The “salt and pepper” voids were replaced by interpolated values using a 5x5 neighborhood filter. The larger voids were filled using GTOPO30 data, also at 1km resolution (Sanchez, 2007). Figure 15 illustrates the raw and filtered data set for a small part of the model domain.

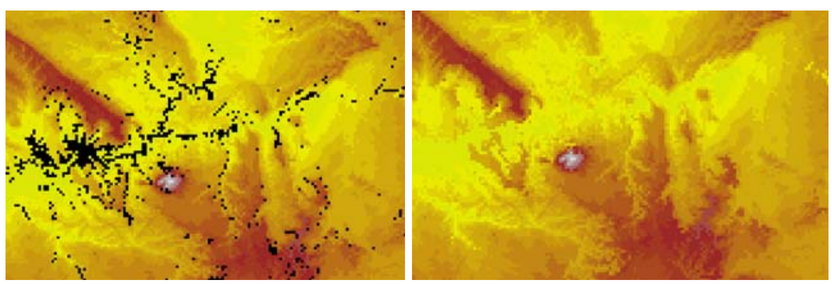


Figure 15. (left) “salt and pepper” voids in SRTM data; (right) appearance after voids were removed using a neighborhood filter.

Aerodynamic surface roughness

SRTM data also were used along with MOD12Q1 to create a surface roughness layer. A program was created at Stennis Space Center to generate aerodynamic surface roughness length (z_0) values using a look up table. These values define the height above 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). For $z_{0,}$ the less water held in a soil, the more prone it is to wind erosion and dust entrainment. Retention of soil moisture is governed by two properties: (a) molecular adsorption on particle surfaces; and (b), interparticle 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; Srivastava et al., 2008).

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 medium scale dust entrainment models does not exist. To overcome this hurdle, Stennis Space Center combined SRTM and MOD12Q1 data sets to simulate a z_0 product. Table 6 is considered a "best practice" estimate of z_0.

Table 6. Look-up values for aerodynamic surface roughness length.

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. Passive microwave emissions can be used to estimate soil moisture in the surface centimeters (NSIDC, 2000). However, there are several interoperability issues to using the data in medium scale dust entrainment models: (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 that is not readily compatible with existing models; (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. Aside from interoperability, one could argue conceptually that rains falling on bare or sparsely vegetated surfaces in arid and semi-arid areas would provide enough soil moisture to retard the entrainment of dust for a day or two depending on soil/air boundary temperatures, surface wind speeds, and duration of wind. DREAM/SW, for example, contains a land surface model (LSM) that treats interactions among soil, vegetation, and atmosphere. It simulates soil moisture and soil temperature variations based on water and heat exchanges on the interface between land and atmosphere, including snow and vegetated areas. When precipitation occurs below zero degrees Celsius, the model counts the precipitation as snow and simulates sublimation and melting processes based on water and heat exchanges at the air/land boundary.

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 aerosols 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 AOD patterns would reveal dust events (Mahler et al., 2006) and discovered that they would not. Candidate dust events were located first in the NASA Earth Observatory archive of natural hazard smoke and dust events. Eight dust storms in the desert southwest were clearly observed in the associated visible imagery. Appropriate MODIS L2 data were examined for all of these cases. Second, a university library data base (Lexis Nexis) of newspaper articles revealed 38 local and regional dust storms, most of which caused fatal pileups on major highways. MODIS L2 data were examined for 26 of these dates. If high winds were reported for a series of days surrounding a dust event, data were ordered for days preceding and following these events in hopes that dust storm evolution could be observed and/or tested in *e*D/SW.

Over land, aerosol optical depth (AOD) is derived using the dark target approach but limited to the moist parts of the continents (Kaufman and Tanre, 1998). A dynamical aerosol model was selected to describe the aerosol size distribution, refractive index, single scattering albedo, and the effect of non-sphericity. Models are derived from analysis of ground based remote sensing of the ambient column aerosol size distribution and in situ measurements. Measured radiance from the satellite is converted into aerosol optical thickness, volume/mass concentration, and spectral radiative forcing.

The expected result was that the AOD product would show well-defined areas of elevated dust concentration in the vicinity of the dust event. However, this did not occur. Horizontal distributions of dust in the regions of dust storms were ill defined by the MODIS AOD product. However, AOD data did show dust over some parts of the reported dust event. Furthermore, AOD products for the desert SW appear geographically incoherent in most cases. The AOD data seem to be interspersed with many pixels of no data.

There are several inherent problems in using MOD04 AOD data despite the fact that AOD has been validated in the literature (Kaufman and Tanre, 1998; Ichoku et al., 2005; Remer, 2005). Primary among these is that the units of measurement are incompatible with PHAiRS needs. That is, MOD04 expresses AOD as a column-mass concentration (ug/cm^2), whereas ground reporting networks report concentrations in ug/m^3. Assuming a vertical column of 10 km yields one cubic meter of air, these units should compare directly. However, ground-based concentrations approaching 10 ug/m^3 did not compare favorably with AOD concentrations of only 1ug/cm^2. In most cases there was little optical distinction between aerosol and desert background. In addition, dust events that were accompanied by cloud cover were not visible to MODIS. A decision was therefore made to focus on particle sizes in the $PM_{2.5}$ to PM_{10} range.

Chapter III: DREAM/SW Performance

The model was tested for two dust storm events. One storm occurred December 8-10, 2003 (hereafter, Case 1); the other on December 15-17, 2003 (hereafter, Case 2). Comparisons were made between the observed and modelled patterns to measure how well the system simulated the meteorology, and to assess the performance of the dust entrainment module. Simulated meteorological fields were evaluated against measurements and data from 95 surface synoptic sites, 663 surface Meteorological Aerodrome Report (METAR) sites, and 77 upper air radiosonde sites. Dust field patterns and relative dust concentrations were compared with satellite images, measured visibility distributions, and surface $PM_{2.5}$ and PM_{10} observations obtained by the U.S. Environmental Protection Agency (EPA) Air Quality System (AQS), and the Texas Commission on Environmental Quality (TCEQ). Graphical measures, such as pattern comparison, site against site time series, vertical profile comparison, and statistical metrics, were used.

Emphasis in this report is on Case 2 results. Case 1 results are used primarily to validate model performance. The Case 2 dust storm was triggered by a Pacific cold front that swept through the region bringing gale force winds and dry conditions, and causing one of the worst dust storms in recent years. Since it was installed in 2001, one of the Continuous Air Monitoring Stations (CAMS) in Lubbock, TX measured a one-hour $PM_{2.5}$ average at the storm's peak of 485.6 $\mu g/m^3$ between 1300-1400 hrs Central Standard Time (22:00 GMT). It also measured a daily average $PM_{2.5}$ of 76.7 $\mu g/m^3$. The PM_{10} daily average concentration of 384 $\mu g/m^3$ was estimated to be five times higher, than is considered "healthy" by the EPA.

Modeled meteorological fields for both surface and 500 hPa (geopotential height) level, agreed 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 that the model performed reasonably well in simulating the measured values.

Surface patterns

For Case 1 (Figures 16, left), note that the 12-hour simulated forecast of the high pressure location over northern Utah matched well with the observed location in Figure 16 (right). Atmospheric pressure differences between the modeled and the measured fields were less than a few millibars. Since pressure patterns reveal something about wind direction and speed, it can be inferred also from these data that wind forecasts were realistic. The model saw this storm coming at least twelve hours in advance.

The 12-hour model forecasts of precipitation along the northern California and southern Oregon coasts match observations. Most important, for Lubbock, TX the location and intensity of the low pressure area, the location and timing of frontal passages, and centers of precipitation were predicted quite well; but, the dust model spread precipitation more generously across central Colorado and western Nebraska than the observations show.

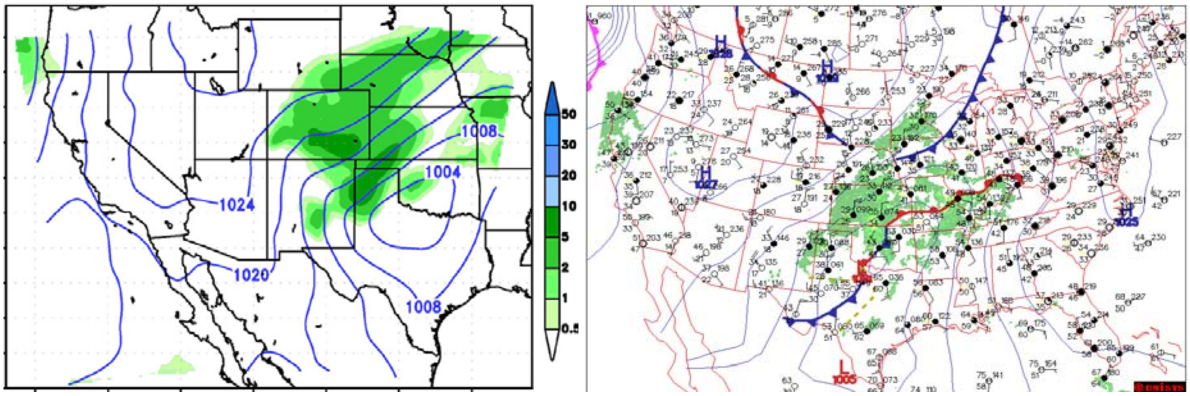


Figure 16. (Case 1): (left) model-generated precipitation (in shades of green) and pressure (blue isobars); (right) observed precipitation and pressure patterns.

The Case 2 storm was also reasonably well predicted (Figures 17). The modeled surface patterns (Figure 17, left) locates the Utah high pressure center at 12Z on 16 December, 2003 slightly north of the observed position (Figure 17, right), and is lower in pressure by perhaps five millibars. The forecasted precipitation pattern is slightly south of that observed, but the frontal passage and pressure (correlated with wind) patterns are reasonably 'forecasted' twelve hours in advance.

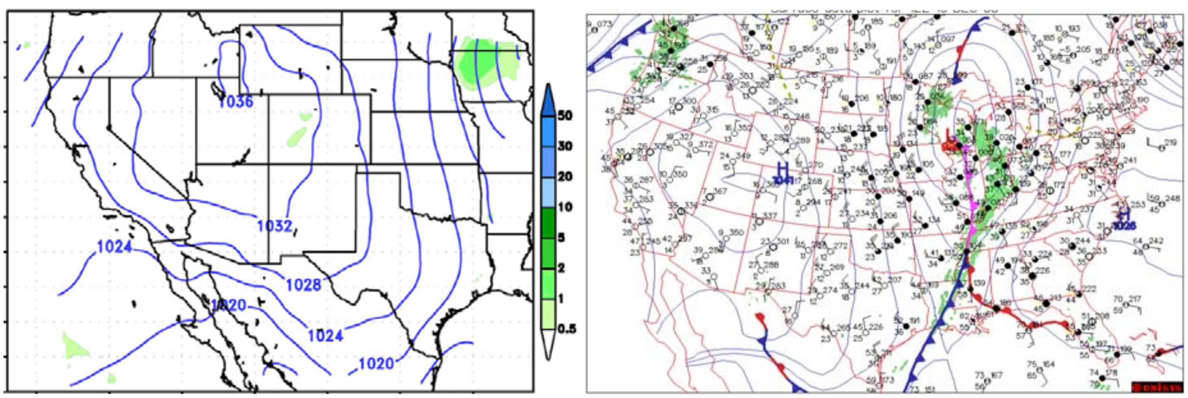


Figure 17. (Case 2): (left) model-generated precipitation (shades of green) and pressure (blue isobars); (right) observed weather patterns.

Upper air patterns

Figures 18 (left) and 19 (left) show the modeled upper-air geopotential height and temperature patterns for both Case 1 and Case 2, respectively. Figures 18 (center and right) and 19 (center and right) show the measured conditions. For Case 1 (Figure 18, center) the cold air intrusion with a low trough located in northern Texas, and the general temperature and height patterns (Figure 18, right), compare reasonably well with the simulated 12-hour forecast with the (after-the-fact) ground observations.

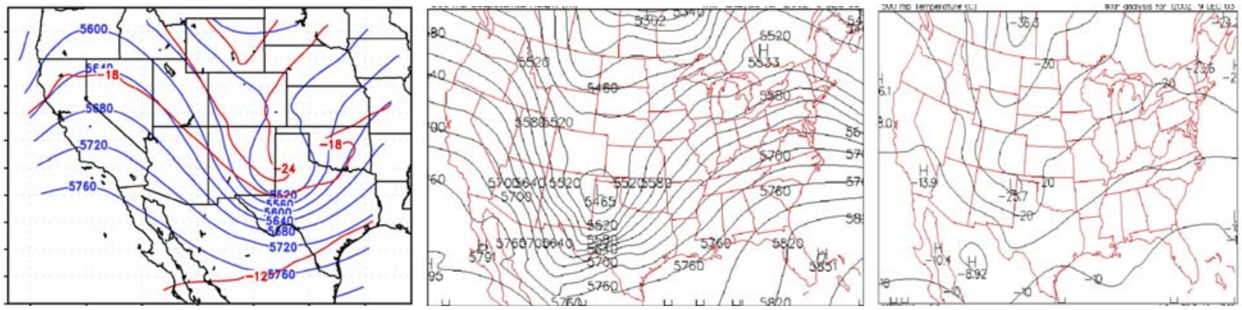


Figure 18 (Case 1): (left) model simulation of pressure (blue isobars) and temperature (red isotherms) at 12Z 9 December, 2003; (center) observed geopotential height; (right) observed temperature

Comparing the same meteorological fields for Case 2, the upper-level trough depicted in geopotential height appears in both the modeled and measured patterns (Figure 19, left) compared to 19 (right). The observed temperature field (Figure 17c) shows fine agreement with the model simulation.

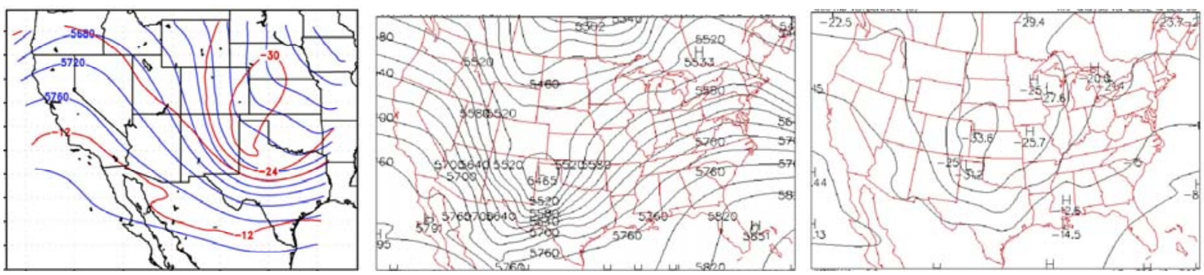
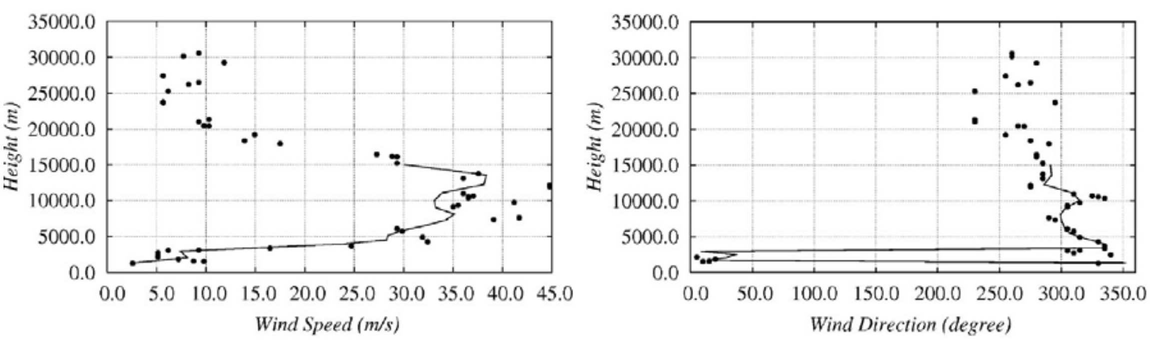


Figure 19 (Case 2): (left) model simulation of pressure (blue isobars) and temperature (red isotherms) at 12Z 16 December, 2003; (center) observed geopotential height; (right) observed temperature.

Profiles through the atmosphere

DREAM/SW and *e*D/SW provide three dimensional dust forecasts Dust concentrations can be calculated for any point, at any height, at any time. Thus, it is important to know if meteorological variables in three-dimensions, including vertical profiles, are being simulated correctly. Figure 20 shows how closely the forecast simulations match the observed variables.



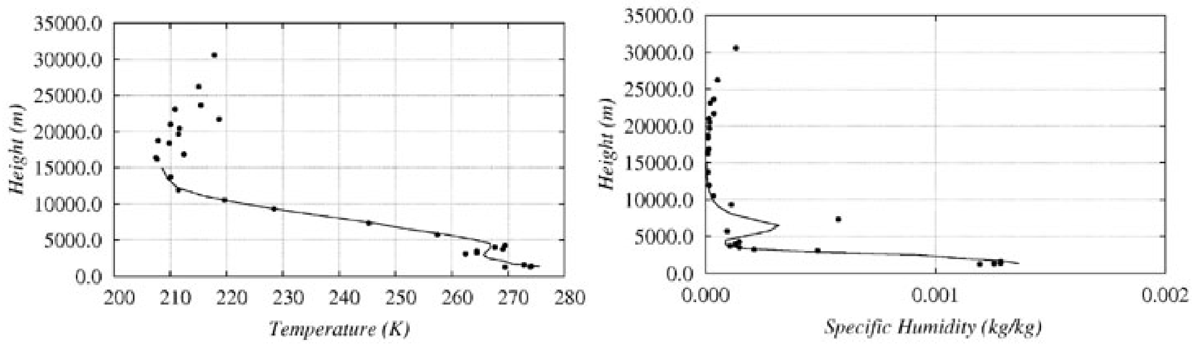


Figure 20. (Case 2): Vertical profiles for wind speed, wind direction, temperature, and humidity at Santa Teresa, NM at 12Z 16 December, 2003. Dots = observed values; lines = modeled values.

Surface wind and temperature

Wind and temperature at a given location change over time as weather systems pass over. If the model cannot simulate these changes, it is unlikely the dust forecast will be correct. Table 7 lists the performance statistics of modeled surface meteorological variables for Case 2. These statistics were calculated using modeled data and hourly measurements from 95 surface synoptic stations and 633 surface METAR stations throughout the domain. In addition to other statistics, the agreement of indices for wind speed, wind direction, and temperature all exceed 0.7. These performance statistics show that DREAM/SW simulates surface wind and temperature reasonably well.

Table 7. Performance statistics of DREAM/SW surface wind and temperature for Case 2.

Metrics	Wind Speed	Wind Dir (degree)	Temp (K)	Definition M=modeled O=observed
Mean Observed	5.53	231.40	276.74	$\frac{1}{N}\sum_{i=1}^{N} O_i$
Mean Modeled	4.65	226.60	275.56	$\frac{1}{N}\sum_{i=1}^{N} M_i$
Mean Bias	-0.88	-4.80	-1.20	$\frac{1}{N}\sum_{i=1}^{N} (M_i - O_i)$
Mean Error	1.97	51.76	4.09	$\frac{1}{N}\sum_{i=1}^{N} \lvert M_i - O_i \rvert$
Agreement Index	0.74	0.74	0.71	$1-\frac{\sum_{i=1}^{N}(M_i-O_i)^2}{\sum_{i=1}^{N}\left(\lvert M_i-\overline{O}\rvert+\lvert O_i-\overline{O}\rvert\right)^2}$

Modeled & storm-generated dust cloud

Figure 21 (left) shows the dust cloud for Case 2 as imaged by the MODIS sensor on Terra and (at center) as modeled by DREAM/SW. Siliceous, calcareous, and ferric dusts all are visible in the cloud-free area. Such high concentrations of airborne dust particles cause severely reduced visibility at the surface. Figure 21 (right) shows the prevailing visibilities from METAR observations in the imaged region for 20Z December 15, 2003. A Cressman analysis (Cressman, 1959) was performed to generate the visibility distribution. An impact radius of one degree was applied to ensure that the gridded data maintain good representation of the site data. White areas in the plot indicate that no measurement data exist. Unfortunately, there are not enough measurement data from Mexico to allow credible comparison with the modeled results in Mexico. Looking only at New Mexico and Texas, the most reduced visibility areas are the Texas Panhandle, specifically the Lubbock and El Paso areas, which correspond well to the locations of dust imaged by MODIS.

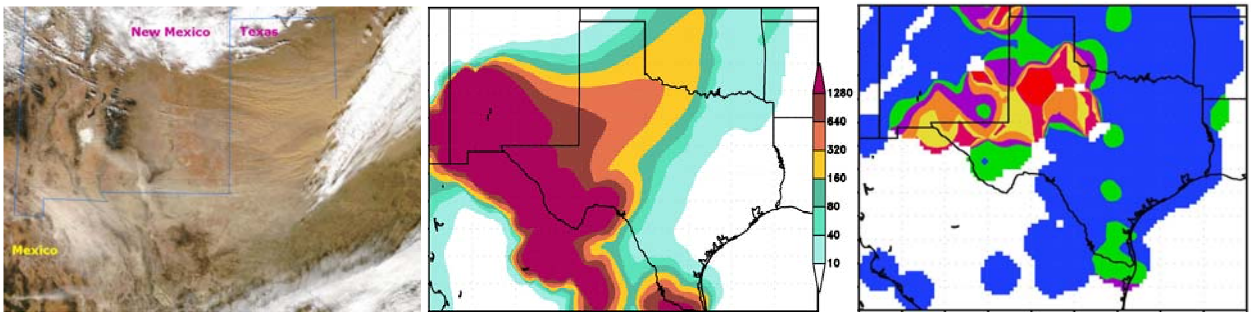


Figure 21. (left) Terra MODIS image, 15 December, 2003. The bright circular spot at left-center is White Sands National Monument.; (center) DREAM/SW dust concentration distribution for 20Z, 15 December, 2003; (right) METAR prevailing visibility in miles for 20Z 15 December, 2003 (greater visibility for blue and green colors; less visibility for red and maroon colors; white areas = no data).

The modeled dust distributions in Figure 21 (center) compare favorable with the MODIS image (left) and the reduced visibility patterns (right). However, there are discrepancies between the details of the modeled and observed dust cloud distributions. The output from DREAM/SW shows that the dust cloud stretched across southern and southeastern New Mexico to eastern Texas. Although the model captured the general areas of dust, there is apparent bias between modeled and observed dust and reduced visibility. The highest dust concentration areas did not coincide well with the most reduced visibility areas.

Point by point comparison

The observed $PM_{2.5}$ data from 40 ambient air monitoring stations in New Mexico and Texas that were affected by the Case 1 and Case 2 dust storms were used to test DREAM/SW performance. Figure 22 shows the locations of some of the sites. The measurement data were obtained from EPA's AQS and TCEQ. They are real time, hourly measurements from Tapered-Element Oscillation Microbalance (TEOM) samplers. The question is "how well does the DREAM/SW model simulate the timing, magnitude, and duration of a dust storm event at these air monitoring stations?

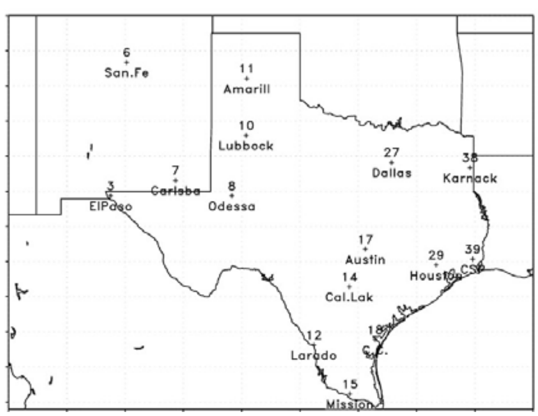


Figure 22. Location of surface $PM_{2.5}$ sites.

Peak hour and peak concentration ($PM_{2.5}$):

The timing of the event is defined as the peak hour; that is, the hour when the maximum $PM_{2.5}$ value occurred at each site. There is reasonable correlation (Case 1 $R^2 = 0.77$; Case 2 $R^2 = 0.76$) between model and *in-situ* timing in both cases when all data (n = 40 sites) are considered collectively. Figure 23 (left) shows the peak hour model performance. Figure 23 (right) shows that there is almost no correlation between the two data sets for Case 2 during the dust-generating stage of the dust storm, and only a weak correlation on December 16 during the transport and deposition stages of the episode. The model over estimated the fine-particle aerosol mass generated in source regions during the onset of dust events in both test cases. Modeled peak concentrations ranged between 1–1185 ug/m^3 compared to the *in-situ* values that, in Case 2 ranged from 11-168 ug/m^3. In both cases, the model appears to underestimate background levels.

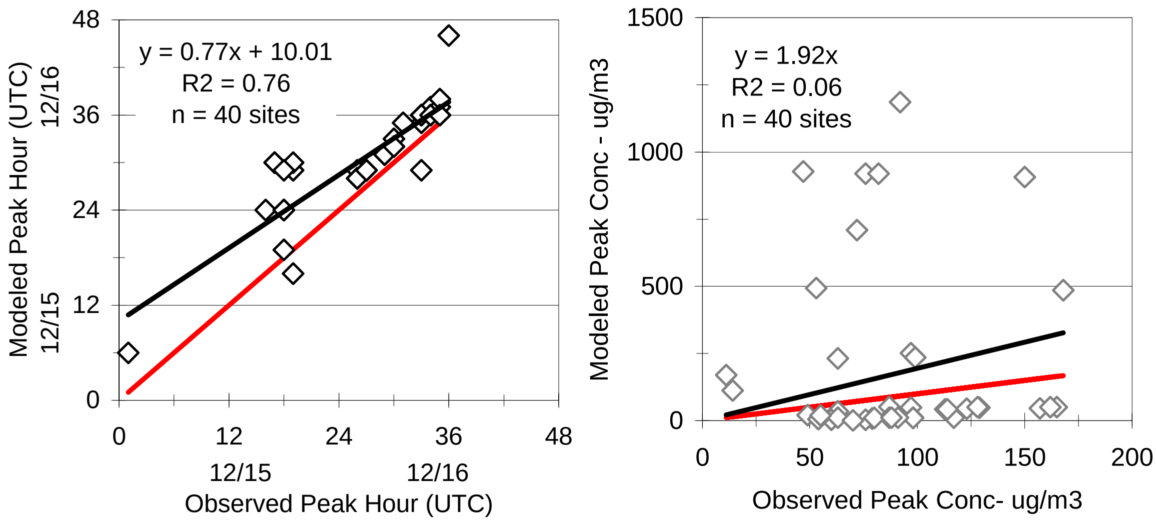


Figure 23 (left). Model run 1a correlation between DREAM/SW $PM_{2.5}$ and observed peak hours for Case 2. Perfect correlation line ($R^2 = 1.0$) in red; (right) DREAM/SW vs. observed peak concentrations.

Dust episode duration:

The duration of the episode is defined as the period of time when measurements at one or more of the *in-situ* monitoring stations exceeded the EPA National Ambient Air Quality Standards (NAAQS). In general, the duration of an episode is a function of the amount of dust generated. The duration of the event was exaggerated by the model at most sites in both test cases.

Of more interest to public health are the periods of time that monitoring sites exceed the primary fine-particle ambient air standard; that is, in excess of a 24-hr average of 65 ug/m^3. EPA primary standards set limits to protect public health, particularly the health of "sensitive" populations such as asthmatics, children, and the elderly. Only one site was found for the two test cases that exceeded the primary fine-particulate standard (Lubbock, TX 12/15/03; 78 ug/m^3). However DREAM/SW predicted eight exceedances in eastern New Mexico and west Texas during the same time period. This effect is related to the over-estimated magnitude of the event throughout the model domain.

Table 9 lists the performance statistics of modeled $PM_{2.5.}$ They were calculated using modeled and observed dust concentrations at 40 sites in the affected areas. The average modeled $PM_{2.5}$ concentrations at these sites are more than 3 times higher than the measured average, possibly because DREAM/SW 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 ESR data sets were assimilated into the model to replace baseline parameters.

Table 9. Performance statistics of modeled surface $PM_{2.5}$ concentrations

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

Chapter IV *Enhanced* DREAM/SW (*e*D/SW) performance

ESR data assimilation

Data assimilation is a multifaceted process hampered by the general absence of metadata. One must first compare the attributes of existing model inputs and of possible satellite data replacements. Like DREAM/SW, many models currently used for Earth system science were designed without benefit of remotely acquired data sets. Data compatibility issues therefore have to be considered, including: (a) measurement units, (b) x,y,z resolution, (c) temporal frequency, (d) map projection and ease of re-projection to fit model requirements, (e) file formats, (f) error and error propagation, and (g) validity of the replacement data in terms of enhancing or improving model outputs. Table 10 lists the original DREAM/MED and DREAM/SW parameters in the left column and the *e*D/SW parameters in the right column. The next step was to iterate the model runs with different groups of products to measure incremental improvements in model outputs.

Table 10. DREAM/SW parameters and candidate *e*D/SW parameters.

DREAM/MED and DREAM/SW Parameters	*e*D/SW Parameters
Land Cover: Olson World Ecosystem 10-min, (19km) res.	MOD12Q1 1km res.
Elevation: USGS 1km terrain data	SRTM-3 arcsec (90m) terrain data re-sampled to 30 arcsec, (1km) res.
Aerodynamic roughness length predicted using 12 SSiB land cover types	Look-up table linked to MOD12Q1 land cover, 1km res.
Additional dust source areas	FPAR "Fill" class 254-255, 1km res.
Soil Moisture: simulated using a land surface model	AMSR-E, 70km res.

Model runs and comparative agreement indices

Table 11 lists key model runs and the ESR data that were assimilated (marked with Y). Model run1a is the DREAM/SW (pre-assimilation) run. Mod12Q1 (barren class) was a standard replacement set in all the model runs and was the only parameter replacement in run 2c. Run 4a assimilated the barren class and digital elevation from SRTM; runs 5a and 5b added z_0; run 6a was a test of Mod12Q1 with Mod15 FPAR substituted for part of the domain; run 15a assimilated barren ground with AMSR-E soil l moisture without digital elevation; and run 10a assimilated barren ground, SRTM, z_o, and AMSR-E soil moisture. Model run 20a replaced only OWE with REGAP data

Table 11. Sample model runs with varying combinations of assimilated ESR data. Model run 1a is DREAM/SW, all others are *e*D/SW runs

Run #	MOD12	SRTM	Surface roughness length	FPAR	AMSR-E	REGAP
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	
Run 20a						Y

Figures 24 show the agreement indices of modeled surface wind speed and direction, temperature, $PM_{2.5}$ and PM_{10} concentrations compared to observations. Although impacts vary with different model runs, in general, the assimilation of NASA ESR data improved DREAM/SW's performance measurably.

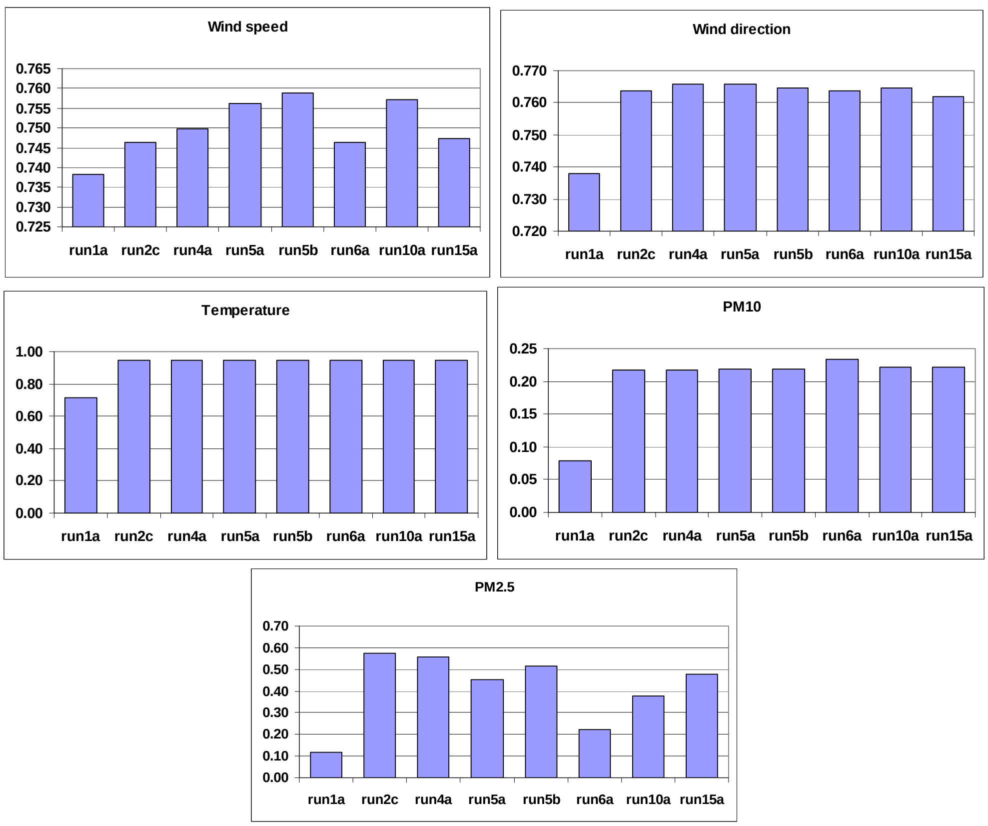


Figure 24. Agreement indices for eight model runs of surface wind speed and direction, temperature, and $PM_{2.5}$ and PM_{10} concentrations compared to observations. REGAP assimilation (model run 20a) is not included.

*e*D/SW Performance

Model performance after MOD12Q1 data were assimilated show that surface weather patterns (sea level pressure, 500 hPa potential height, and temperature) match well with the observed weather patterns. This suggests that the MOD12Q1 product for dust sources had little noticeable affect on the performance of atmospheric fields. The primary difference between the two sets of model results is seen in sea level pressure fields, although these differences did not affect the overall pattern. The upper-air fields were *not* affected by the data set replacements.

Among the vertical profiles for wind, temperature, and specific humidity, only slight differences were seen after data assimilation, except for differences in the near-ground wind speed. This seems reasonable since the OWE land cover data used in DREAM/SW had much coarser spatial resolution. Even though both data sets resulted in fairly good model performance, one expects vegetation height and density to add incrementally to topography's influence on surface wind speeds; and, in turn, to influence surface roughness length, soil moisture status, and the ability of wind to entrain dust.

The performance statistics of the modeled surface meteorological variables using MOD12Q1 showed a significant improvement in 2m (height above surface) temperature compared to DREAM/SW performance. Model performance for 10m wind speed and direction showed slight improvement using assimilated data.

Surface patterns

Figure 25 shows the modeled sea level pressure and 12-hour precipitation at 12Z for Cases 1 and 2, respectively. Detailed visual comparison of these outputs with results in Figures 16 and 17 (page 22) show that sea level pressure contours are not as smooth as in the DREAM/SW results but that, overall, the sea level pressure and precipitation patterns are very similar. These patterns match the measured patterns.

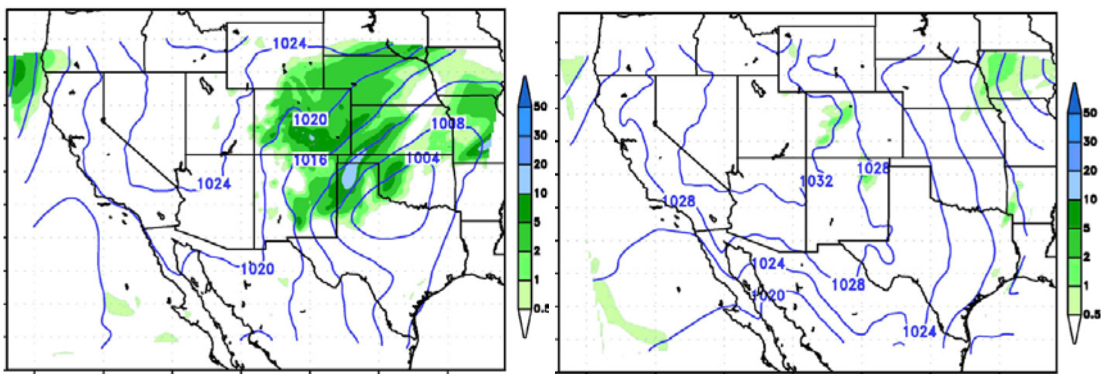


Figure 25. (left) modeled sea level pressure (blue isobars) and precipitation (green shaded areas) at 12Z for Case 1; (right) for Case 2.

Upper air patterns

Figure 26 shows modeled 500 hPa geopotential height and temperature fields for Cases 1 and 2. Compared with DREAM/SW results in Figures 18 and 19 (page 23), the modeled geopotential height and temperature fields in *e*D/SW are virtually identical. Again, the upper air fields were not affected.

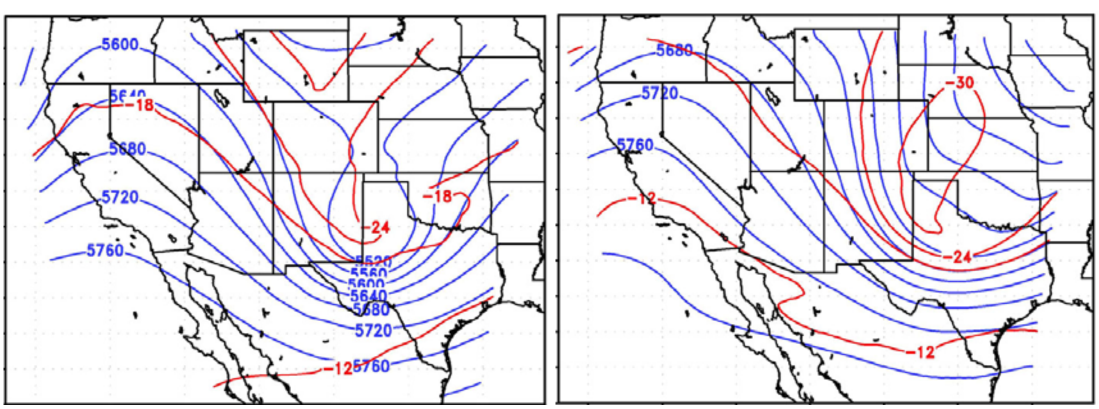


Figure 26: (left) modeled 500 hPa (blue isobars) and temperature field (red isotherms) at 12Z for Case 1; (right) for Case 2.

Profiles through the atmosphere

Visually, the modeled variables in Figure 27 match quite well with the observed values. Compared with Figure 20 (page 38), the modeled wind direction, temperature, and specific humidity profiles are almost identical to the DREAM/SW performance; however, as indicated earlier, there are very important differences in near-surface wind speeds.

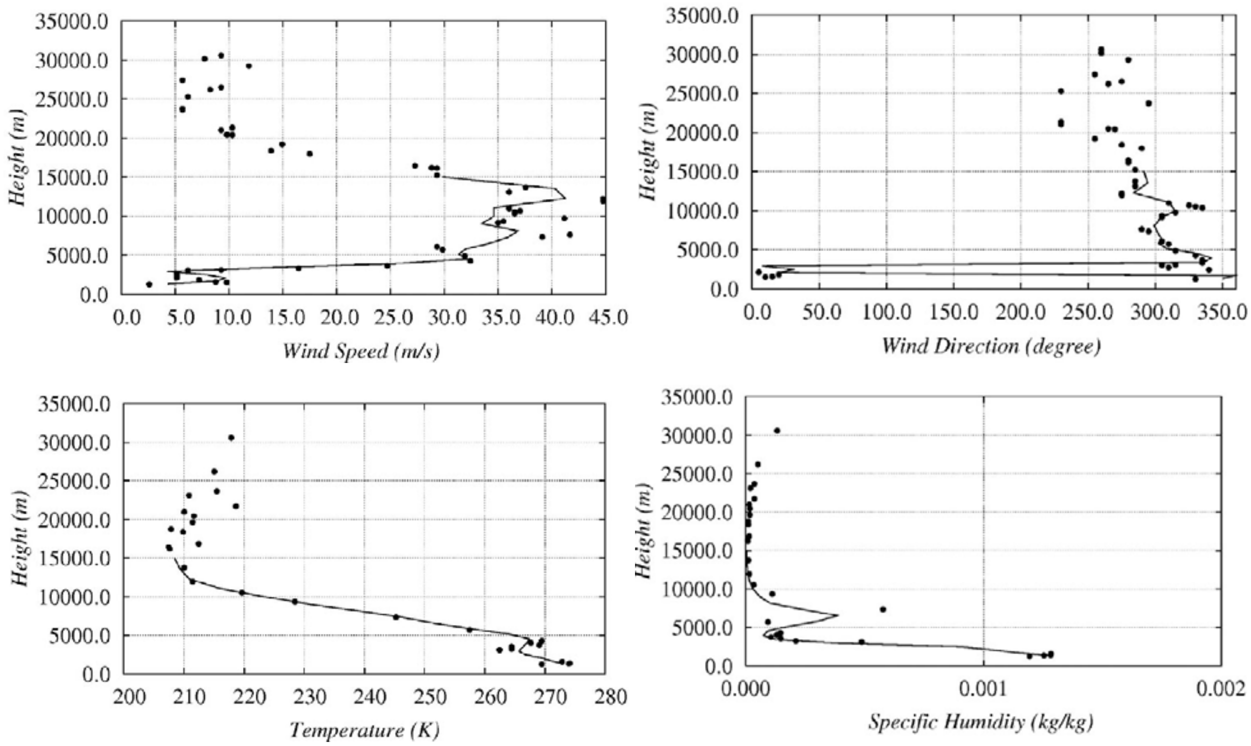


Figure 27. Vertical profiles at Santa Teresa, NM at 12Z 16 December, 2003. Dots = observed values; lines = modeled values.

Surface wind and temperature

Table 12 lists the performance statistics for model run 1a and 4a (DREAM/SW vs. *e*D/SW) for surface variables. The biggest difference between the two runs occurred in temperature at 2m height above surface. The agreement index after ESR data assimilation was 0.95, compared with 0.71 in DREAM/SW. This is an important measurable improvement in model performance. The mean bias and mean error after parameter replacement are lower than those for DREAM/SW. The agreement index for 10m wind direction and speed was slightly better after parameter replacements, but the mean bias and mean error were actually slightly higher than those obtained from DREAM/SW parameters.

Table 12. Comparative performance statistics for DREAM/SW and *e*D/SW surface wind and temperature for Case 2. DREAM/SW values are in blue; *e*D/SW values in red

Metrics	Wind Speed	Wind Direction (degree)	Temp (K)	Definition M=modeled O=observed
Mean Observed	5.53	231.40	276.74	$\frac{1}{N}\sum_{i=1}^{N} O_i$

Mean Modeled	4.65 4.37	226.60 230.38	275.56 277.48	$\frac{1}{N}\sum_{i=1}^{N} M_i$
Mean Bias	-0.88 -1.16	-4.80 -1.02	-1.20 0.72	$\frac{1}{N}\sum_{i=1}^{N} (M_i - O_i)$
Mean Error	1.97 2.03	51.76 47.85	4.09 2.67	$\frac{1}{N}\sum_{i=1}^{N} \lvert M_i - O_i \rvert$
Agreement Index	0.74 0.75	0.74 0.76	0.71 0.95	$1-\frac{\sum_{i=1}^{N}(M_i-O_i)^2}{\sum_{i=1}^{N}\left(\lvert M_i-\overline{O}\rvert+\lvert O_i-\overline{O}\rvert\right)^2}$

Modeled & storm-generated dust cloud

Parameter replacements in model run 4a had a major impact on modeled dust concentrations. The near ground dust concentration distribution matched better to the satellite observed dust cloud and reduced visibility distribution. Figure 28 is a replication of Figure 21 (page 24) but with the *e*D/SW modeled dust cloud replacing the DREAM/SW image. The dust pattern is much better defined than in DREAM/SW. The modeled high dust concentration areas also correspond better to the ground-based visibility observations (compare with Figure 21 [right]).

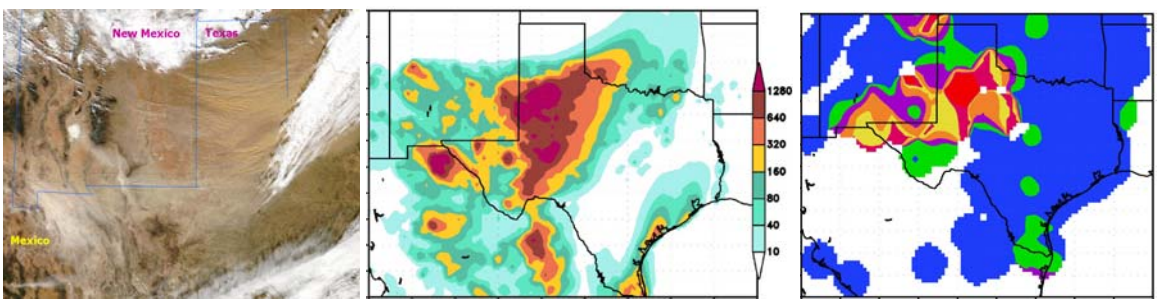


Figure 28. *e*D/SW output from model run 4a. Compare with patterns shown in Figure 21 (page 39).

Point-by-point comparison

Major improvements were achieved by assimilating ESR data into *e*D/SW. The peak hour correlation was least affected by the change. However major improvements were made in the magnitude comparison. The enhanced model simulated the magnitude of the dust storm event better than DREAM/SW at almost all locations in the model domain. The dust episode in Lubbock, TX was also better simulated. The improved performance indicates no false alarms in either test case. This result begins to illustrate the potential use of *e*D/SW as a tool for health communities and local governments to use for unhealthy dust level alerts in the Southwest.

Peak hour and peak concentration (PM2.5)

Correlations between model and observed timing when all sample sites are considered are shown in Figure 29. The *e*D/SW model results appear slightly more scattered than results from DREAM/SW, but the difference between the two outputs is within a reasonable margin of error. There was a substantial improvement in R^2 values for the $PM_{2.5}$ peak concentrations (DREAM/SW, $R^2 = 0.06$; *e*D/SW $R^2 = 0.29$).

The range of modeled values is in good agreement with measured *in-situ* values (14-168 ug/m^3). Although the correlation is better, there is much room for improvement.

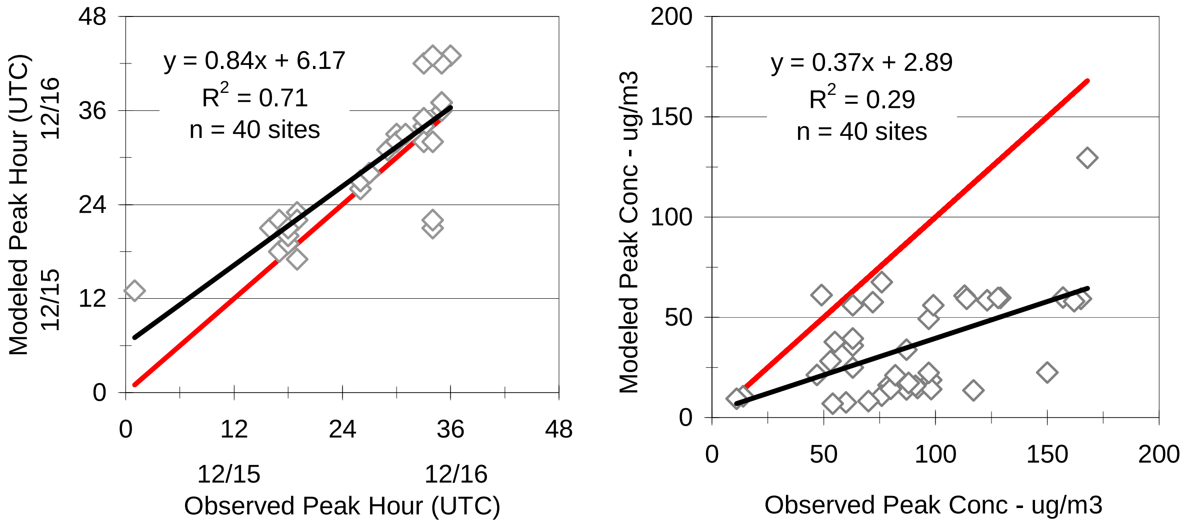


Figure 29. (left). Model run 4a correlation between *e*D/SW $PM_{2.5}$ and observed peak hours for Case 2. Perfect correlation line (R^2 = 1.0) in red; (right) *e*D/SW vs. observed peak $PM_{2.5}$ concentrations.

Dust episode duration

Performance improved also for simulation of the duration of dust episodes. The *e*D/SW output accurately predicted the exceedance for fine-particle ambient air standard observed at Lubbock, TX on 15 December, 2003. For the two case studies DREAM/SW simulated exceedances eight times in western NM and east TX. The *e*D/SW had no 'false alarm' predictions.

Southwest REGAP experiment

Products from REGAP were assessed for their contribution to improved resolution (Landsat TM vs. MOD12Q1) for the barren ground categories. In Figures 30(a-d), "*in-situ*" refers to ground-based measurements of $PM_{2.5}$. from 40 measurement sites. Figures 30 (a) and (b) compare observed concentrations with results from DREAM/SW (i.e., model run 1a without ESR enhancements). In Figure 30 (b) the observed values are compared to results from *e*D/SW model runs 4a (enhanced with MOD12Q1) and 20a (enhanced with REGAP) to assess the impact of assimilating a prototypical seasonal dust source data set. Figures 30 (c and d) repeat the comparisons for peak hour. These comparisons clearly indicate that *e*D/SW with MOD12Q1 and REGAP data performed much better than the DREAM/SW model run. Model run 4a results using MOD12Q1 and 20a using REGAP data are quite similar.

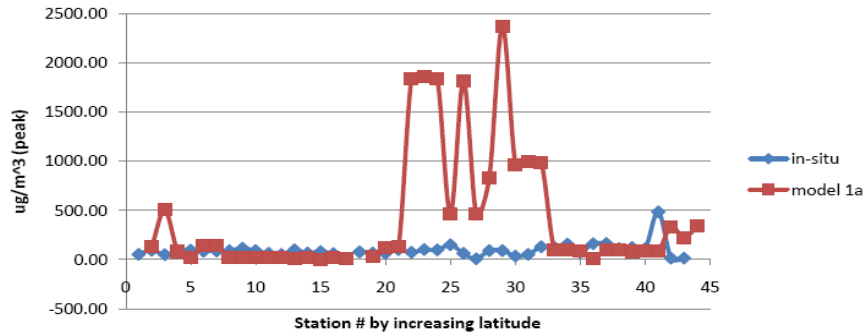


Figure 30(a). Observed (blue) vs. DREAM/SW (red) peak hour concentrations of $PM_{2.5}$ (model run 1a) using the OWE barren class.

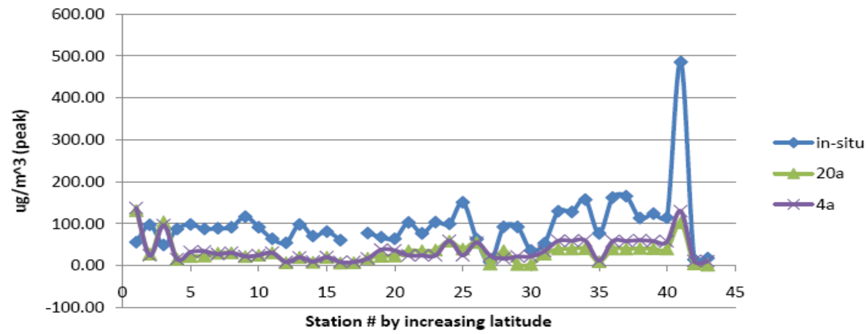


Figure 30(b). Observed (blue) vs. *e*D/SW peak hour concentrations of $PM_{2.5}$. The models used barren ground data from MOD12Q1 (model run 4a [purple]) and REGAP (model run20a [green]).

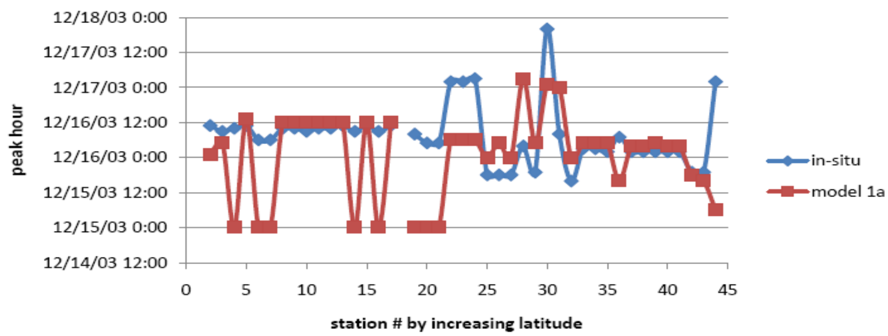


Figure 30(c). Observed (blue) vs. DREAM/SW peak hours (model run 1a) using the OWE barren class.

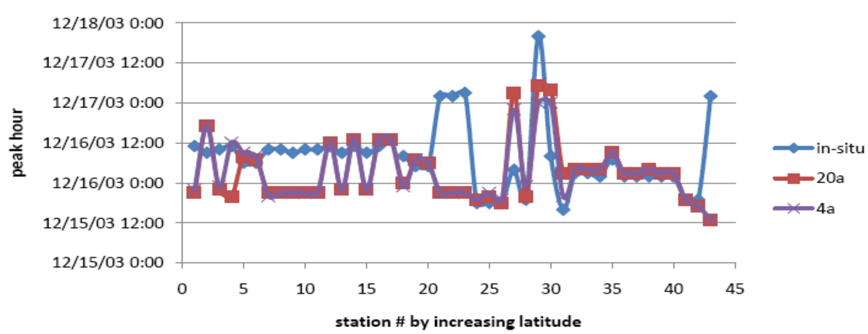


Figure 30(d). Observed (blue) vs. *e*D/SW peak hours. The model used barren ground data from MOD12Q1 (run 4a [purple]) and REGAP (run 20a [red]).

Chapter V. Verification and & Validation

Approach

The following approach evolved according to the steps outlined in Chapter II.

For model performance, simple correlations (peak hour and concentration) were used between hourly observed and modeled outputs to assess how well *e*D/SW predicted dust events. Verification and validation was done by making qualitative and statistical comparisons of model outputs with *in-situ* dust concentrations reported by EPA's AIRNow network.

A set of statistical metrics was defined for DREAM/SW by Yin et al (2005), and used to assess which parameter replacements led to improvements in *e*D/SW. These metrics related to surface meteorological parameters important in dust entrainment, and resulted in agreement indices between observed and modeled data sets.

It was discovered that this statistical approach was not wholly adequate. Agreement indices tend to decrease with increasingly large data sets, so hourly point-by-point comparisons are not sufficient to verify model performance.Results suggest that there are lags in model timing and concentration averaging, which help improve verification of model performance. 24-hour averages were compared to test the model's ability to predict exceedances of the EPA health standards for $PM_{2.5}$ and $PM_{10.}$ Finally, a weather forecaster's approach was used to model verification by using the WRF Model Evaluation Tools (WRF MET, 2008). Skill and threat scores were calculated in much the same way meteorologists predict rainfall patterns. These show great promise as *e*D/SW verification tools.

AIRNow data

Historical AIRNow data (hourly $PM_{2.5}$ and PM_{10}) for the entire modeling period (2006-present) are available through the DataFed's *AIRNow Web Coverage Service (WCS)*. These data are acquired daily as a CSV file for all EPA stations within the model domain for the previous 60 days (Figure 31). 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 development 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 *e*D/SW model outputs with well-defined ground observation times.

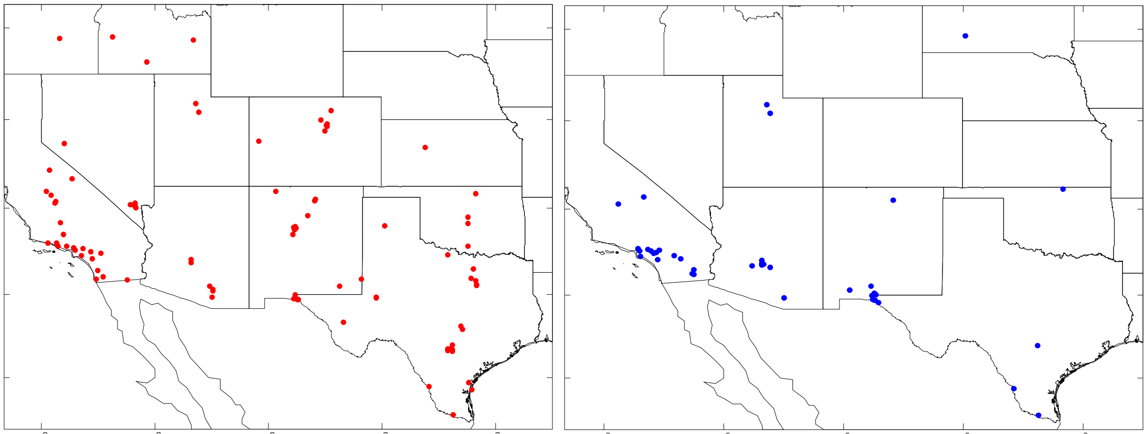


Figure 31. (left) $PM_{2.5}$ monitoring sites; (right) PM_{10} monitoring sites.

Dust storm cases (2007)

Model runs were compared to observational PM_{10} data during dust events that occurred in Texas and southern California. One of these occurred on 5 January, 2007, when a severe wind and dust storm grew near Barstow, CA. A minivan collided with 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 high winds, and several big-rigs toppled or jackknifed on freeways across the region (*High Winds Aren't Over Yet*, L.A. Times, 6 January, 2007, p. A1). Similar high winds occurred across the desert southwest as far east as Texas. The *e*D/SW outputs were compared in a hind-casting mode with subsequent data from AIRNow PM_{10} monitoring stations. Four stations were located in Southern California (Burbank, Riverside, Palm Springs, Indio), and three others in Texas (El Paso, Mission, Selma).

Model statistics

Figure 32 is a plot for each station 24 hours before, during, and after the event (4-6 January, 2007). The stations are plotted geographically from west on the left to east on the right. Southern California was most affected by this event. The data show that the dust event occurred around 23:00 UTC on January 5th at most stations. Both the observed and modeled data show a spike in the dust gradient with the exception of Riverside, where no significant event was recorded by the AIRNow station. A modest model improvement between model run 15a and 20a is suggested by decreased magnitudes in the data for Burbank, Riverside, and Palm Springs. 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.

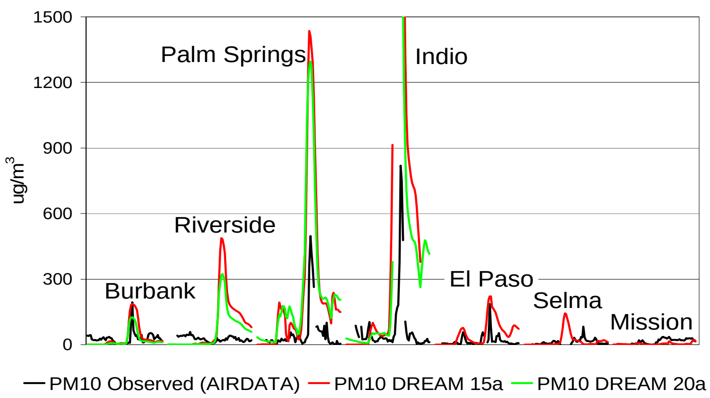


Figure 32. Modeled and observed PM_{10} magnitudes at seven AIRNow stations for 4-6 January, 2007 for model runs 15a (MOD12Q1 and AMSR-E) and 20a (REGAP).

Figure 33 represents the correlation between modeled and observed dust magnitudes for 5 January, 2007. 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 model's tendency to over-predict dust magnitudes. Model improvements are indicated in the improved correlation from model run 15a to 20a. Table 13 is a statistical analysis performed for the seven sites using the observations and output from model run 20a.

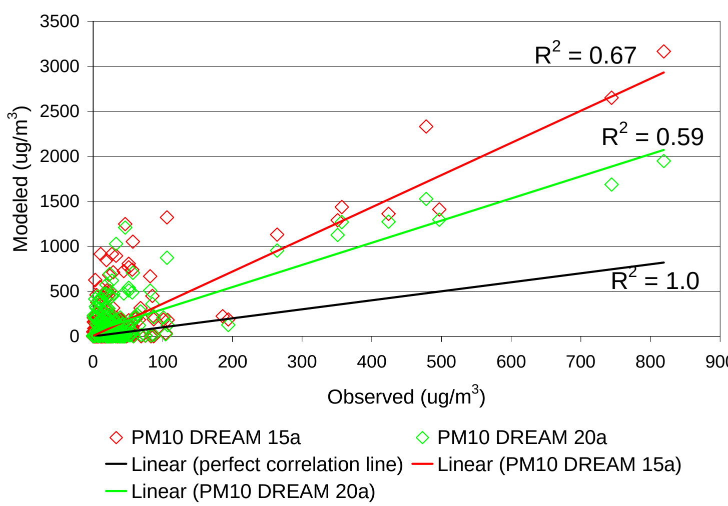


Figure 33. Magnitude correlation (R^2) for seven sites, (N = 443) during the 4-6 January, 2007 episode.

Table 13. Statistical analysis of seven test sites for the 4-6 January, 2007 event.

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 peak hour correlations for the 4-7 January, 2007 test case is shown in Figure 34. The x-axis is a 72-hour event clock; the y-axis is a plot of the modeled vs. observed peak hour. Several sites had more than one peak hour during the three-day event. A plot of daily peak hours 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 run 20a show an improvement over previous model runs published in earlier work ($R^2 = 0.76$, Yin et al, 2005).

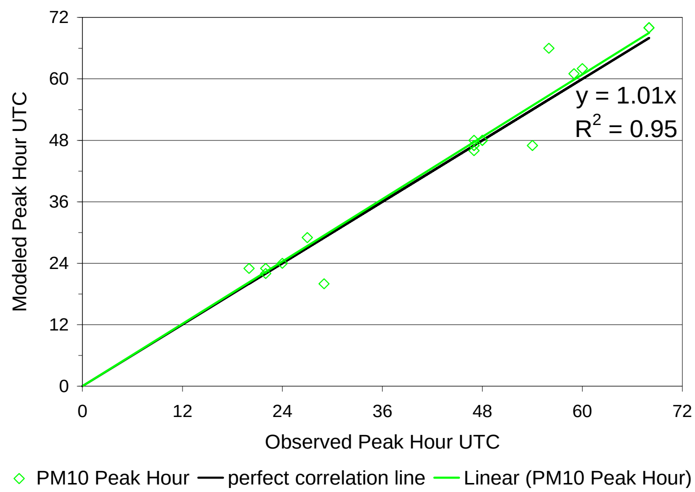


Figure 34. Timing correlation (N=18 peak hours, seven sites) for 4-6 January, 2007.

Another test case was evaluated 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 33). 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 the AQI scale.

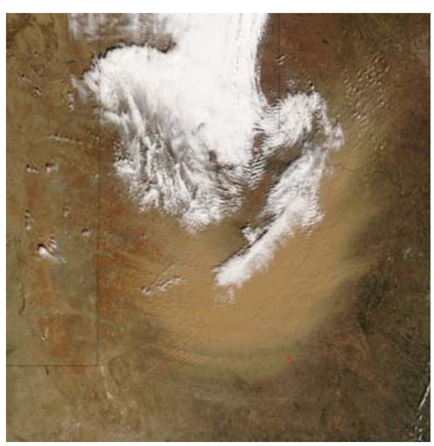


Figure 35. MODIS Terra 500m resolution image of a dust storm in Texas on 24 February, 2007.

Figure 36 is a three day plot of dust magnitudes for seven stations before, during, and after the 24 February, 2007 dust storm that occurred around 00:00 UTC. The stations are plotted as in Figure 32 (page

38) for model runs 15a and 20a. Results under-predicted the event at Burbank, Riverside and 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 model outputs were in fairly good agreement for these sites.

Considerable effort was spent improving the *in-situ* data sets by filling in missing observed data and adding monitoring stations to the verification process. Table 14 exemplifies how the statistics for the dust event of 23–25 February changed when the number of sites was increased. What became evident was that, as the data sets became larger and more complete, the agreement indices showed less agreement. Some of the reasons for this may be that: (a) the number of 'non-events' and model 'false alarms' was increased; and (b), when the model predicted a dust event to occur and the observed data also showed the event (but at a different peak hour), the agreement index lowered even though the model hind-casted a dust event within a few hours of the observation.

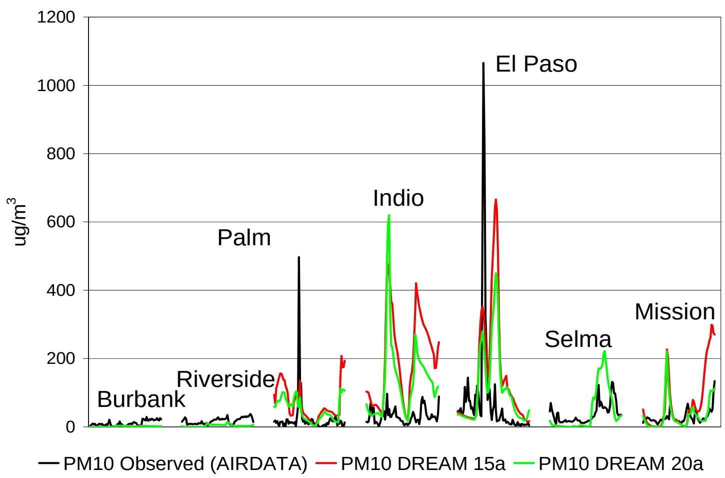


Figure 36. Modeled vs. observed $PM_{2.5}$ magnitudes at seven AIRNow stations for the 23-25 February, 2007 dust event for model runs 15a and 20a.

Table 14. PM_{10} agreement indices for the 23–25 February, 2007.

# Sites	4	7	9	13
N (hourly observations)	276	472	590	1157
Mean Observed	41.5	33.9	52.8	60.5
Mean Modeled	124.7	57.3	70.5	44.8
Mean Bias	82.9	23.4	17.6	15.9
Mean error %	106	51.1	70.4	64.0
Normalized mean bias	66.8	40.9	25.0	26.2
Normalized mean error	68.0	64.6	63.4	63.7
Fractional bias	81.0	-12.7	-2.9	27.4
Fractional error %	116	115	113	124
Index of Agreement	0.84	0.59	0.53	0.37

Figure 37 illustrates the magnitude correlation between modeled and observed data for the 23-25 February, 2007 test case. Correlations for both model versions were poor for this test case ($R^2 \sim 0.1$), due primarily to the Palm Springs and Indio data discrepancies. In spite of this, the timing correlations (Figure 38) showed excellent agreement between observed and modeled peak hour.

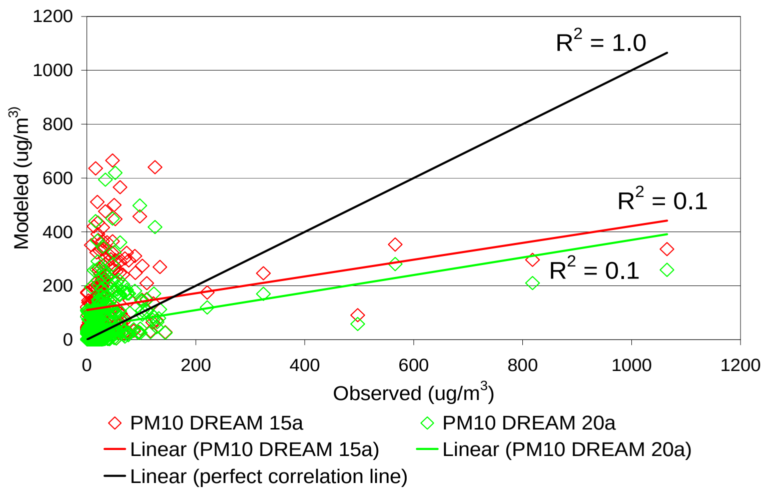


Figure 37. Magnitude correlation ($R^2 = 0.1$) between observed and *e*D/SW modeled data, 23-25 February, 2007.

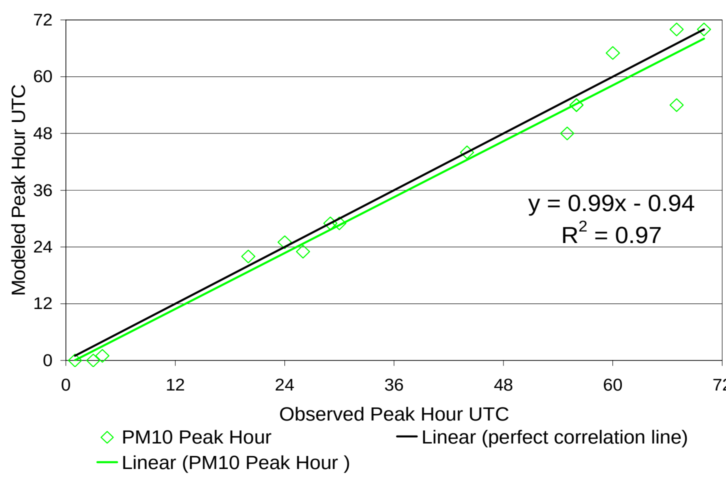


Figure 38. Peak hour correlation, 23-25 February, 2007 (N=16 peak hours).

The same statistical analyses that included seven sites using model 20a is shown in Table 13 for the February 24, 2007 event. The statistics indicate that the model had a negative bias, or under-predicted the event. The January test case had a positive bias and a much better index of agreement (0.63 vs. 0.42).

Table 15. 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

The Point-Stat tool

To augment point-by-point correlations to verify *e*D/SW performance, an alternative statistical V&V approach was adopted. Methods used by the weather forecasting community to evaluate climate model performance led to the Weather Research & Forecasting (WRF) Model Evaluation Tools User's Guide (MET) developed at the Development Testbed Center, National Center for Atmospheric Research in Boulder, Colorado. The primary goal of MET development is to provide a state-of-the-art verification package to the numerical weather prediction community. Currently, MET is a set of tools that can be applied easily by any user on their own computer platform. Although it was specifically designed for application to the WRF model, MET can be used for the evaluation of simulations from other models such as *e*D/SW.

The MET Point-Stat tool provides categorical verification stats for modeled forecasts at observation points. It matches gridded forecasts to point observation locations using several different interpolation approaches. One is intrinsic, as in the case of rainfall, where the observation points either have rain or no rain; another uses a 'rain threshold' such as 0.01" to verify the model's ability to predict measurable rainfall.

In order to evaluate *e*D/SW's performance as a predictive tool for the public health community, the EPA's 24-hour standards for $PM_{2.5}$ (35 ug/m^3) and PM_{10} (150 ug/m^3) particulates were used as a 'dust threshold' in the statistical verification process. The verification stats are formulated using a contingency table as shown in Table 16. "M" represents the modeled hourly forecasts of $PM_{2.5}$ or coarse dust concentrations and "O" represents the AIRNow hourly observations; the two possible M and O outcomes are represented by zero (no) if the EPA standard was not attained and one (yes) if the outcome exceeds the 'dust threshold'.

Table 16. Contingency table for observed and modeled dust concentrations

Modeled forecast	Observation (yes)	Observation (no)	Total
	O = 1	O = Ø	
M = 1 (yes)	n_{11}	n_{10}	$n_{11} + n_{10}$
M = Ø (no)	n_{01}	n_{00}	$n_{01} + n_{00}$

Total (T)	$n_{11} + n_{01}$	$n_{10} + n_{00}$	$n_{11} + n_{10} + n_{01} + n_{00}$

Where the counts n_{11}, n_{10}, n_{01}, n_{00} are called 'hits', 'false alarms', 'misses', and 'non-events' respectively. From this table, all of the statistics for the Point-Stat tool can be generated. A few of the many performance measures are described in Table 17.

Table 17. Sample performance measures provided by a modeled vs. observed contingency table

Statistic	Definition	Answers the question
Accuracy	$(n_{11} + n_{00})/T$	What portion of forecasts were either hits or non-events?
Probability of detection (POD)	$n_{11}/(n_{11} + n_{01})$	What fraction of events was correctly forecasted?
Probability of false detection (POFD)	$n_{10}/(n_{11} + n_{00})$	What fraction of forecasted events did not occur?
Threat score (TS)	$n_{11}/(n_{11} + n_{10} + n_{01})$	What fraction of events was successfully modeled? (ignores non-events)
Skill score (SS)	POD - POFD	How well does the model discriminate between events and non-events?

Although the EPA 24-hour standards are calculated from midnight to midnight local standard time, a rolling 24-hour average was used to account for model timing lags (usually zero as indicated by the superb timing agreement indices, but occasionally one to three hours between modeled and observed peak hour). If a dust threshold was exceeded at any time during the day, it was defined as a "yes" condition. This method allows the earlier performance stats (i.e., timing and magnitude agreement indices for either or both the $PM_{2.5}$ and PM_{10} aerosol) to be combined into one set of verification statistics yield probabilities described below.

Performance stats were calculated for the 2007 test cases using the Point-Stat tool as indicated in Table 18 only ten 'hits', or exceedances, of the 'dust threshold' occurred over the entire model domain during these events.

Table 18. Modeled vs. observed hourly dust forecasts of dust concentrations over the entire *e*D/SW domain during the CA,AZ,NM, TX events of 4–6 January, and 23–25 February 2007.

Case	Fraction	N	Hits	Misses	False alarms	Non-events
Jan '07	pm10	27	2	0	3	22
Feb '07	pm2.5	267	4	0	15	248
Feb '07	pm10	52	4	4	2	42
Combined	pm2.5	267	4	0	15	248
Combined	pm10	79	6	4	5	64

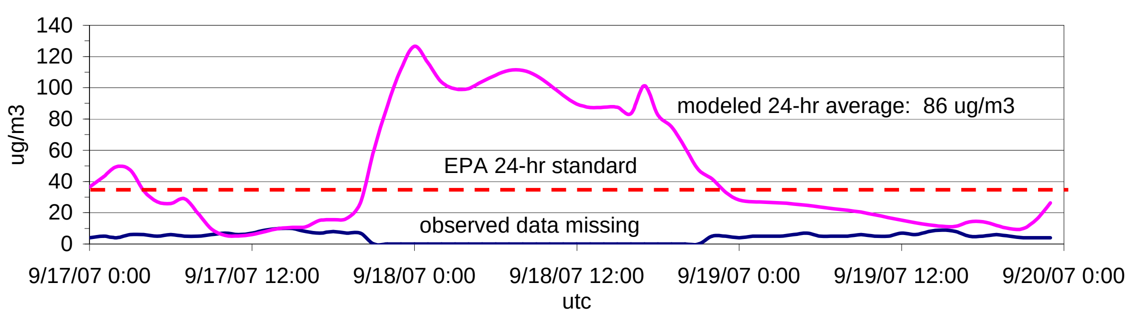
Combined	Combined	346	10	4	20	312

Using these combined values yields the following performance stats with the Point-Stat tool. These results indicate that *e*D/SW successfully forecasted 71% of the hourly averages and only 29% the hours exceeding the dust threshold during these two events. Although a new tool for evaluating model performance has been introduced, the point-by-point comparisons are still being used. The approach has been modified using the Phoenix metro area having multiple observation points (Table 19).

Table 19. Point Stat Evaluation – Phoenix Metro Area

Accuracy	What portion of forecasts were either hits or non-events?	0.93
POD	What fraction of events was correctly forecasted?	0.71
POFD	What fraction of forecasted events did not occur?	0.06
TS	What fraction of events was successfully modeled? (ignores non-events)	0.29
SS	How well does the model discriminate between events and non-events?	0.65

The Phoenix metropolitan area had seven $PM_{2.5}$-particulate monitoring sites on the EPA's AIR-Now network during early 2007. Performance statistics were computed for the time period January–April 2007 using the Point-Stat variables described above, which yielded 111 daily averages where both observed and modeled data were available for comparison. If on any given day one or more of the seven monitoring sites observed a 'yes' condition, and the model indicated a 'yes' condition at one or more of the seven sites, these days were considered 'hits'. Several days were discovered where the model indicated extremely high dust levels and the observed data were missing from the AIRNow data set. Possible reasons include: (a) occasionally, if $PM_{2.5}$ collection instruments are overwhelmed, they will shut down or render the sample useless; (b) the local air quality agency may have determined that the data did not meet verification requirements; or (c), the dust collector may have been inoperative at the time for one of numerous other reasons such as power failures during high wind days. The DYSART $PM_{2.5}$ monitoring site in Phoenix had at least three such days in 2007 when *e*D/SW predicted a dust episode that exceeded the $PM_{2.5}$ threshold (35 ug/m3). These three days are included as 'hits' in the analyses and the events are shown in Figure 39.



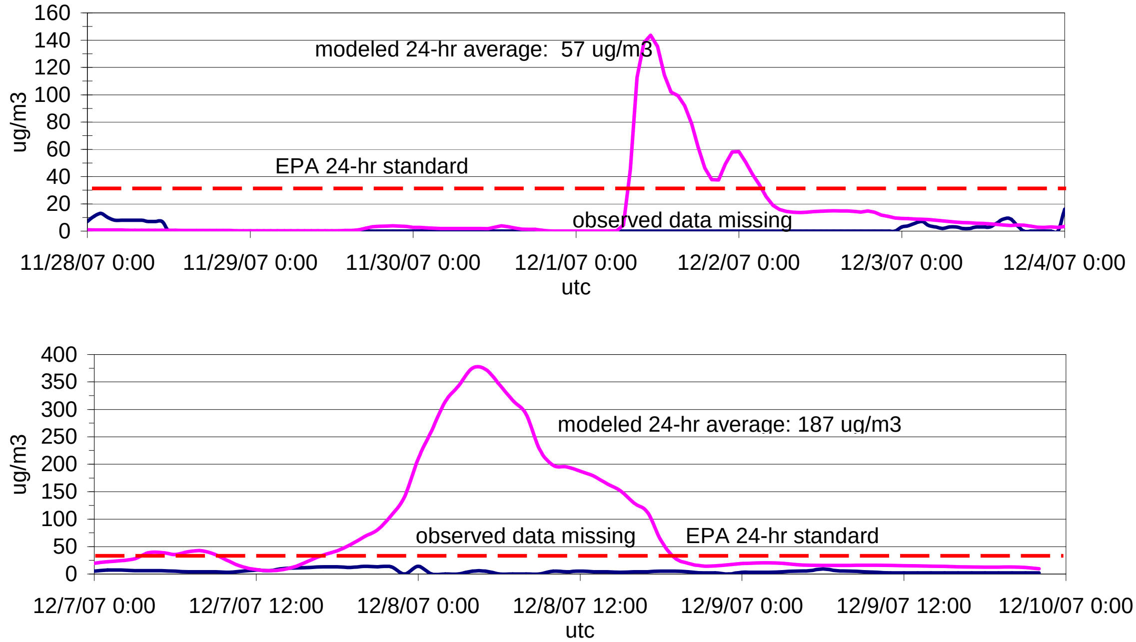


Figure 39. DYSART $PM_{2.5}$ Monitoring Site, Phoenix Arizona

Results of the Point-Stat evaluation of the Phoenix metro area, January – April 2007, are shown in Table 20.

Table 20. Performance stats for the Phoenix metropolitan area for January – April 2007

N	Hits	Misses	False alarms	Non-events
111	64	24	10	13

Accuracy	What portion of forecasts were either hits or non-events?	0.69
POD	What fraction of events was correctly forecasted?	0.73
POFD	What fraction of forecasted events did not occur?	0.13
TS	What fraction of events was successfully modeled? (ignores non-events)	0.65
SS	How well does the model discriminate between events and non-events?	0.60

Of the 111 days where both observed and modeled data were available, 64 days were 'hits' (over half the days in this period exceeded the $PM_{2.5}$ dust threshold). Significantly better *e*D/SW performance statistics resulted from using numerous observed data points that can be obtained simultaneously over a single metropolitan air shed such as the Phoenix valley. These results indicate that *e*D/SW can be used in much the same way weather forecasters predict the possibility of rainfall over metropolitan areas. The *e*D/SW had a threat score of 0.65 from January–April 2007, which implies that the odds of hind casting a dust event somewhere in the Phoenix metro area during that time period were roughly two-out-of-three.

Improvements to the model will yield better statistical correlations between *in-situ* and modeled data at more stations in the model domain.

Chapter VI: Data and Information Systems

SYRIS™

Key to mitigating disease epidemics is situational awareness, both before an outbreak occurs and during an outbreak. Physicians, veterinarians and their assistants see many of the first cases. A large number of professionals who, though not usually considered to be part of the "clinical" community, also see ill people or animals and collect related data, which is of great value to public health officials (PHOs). Emergency medical technicians, school nurses, animal control personnel, laboratory technicians, and medical investigators fit into this "other" clinical group.

The Syndrome Reporting Information System™ (SYRIS) is a JAVA-based, platform independent system that runs on most PCs and laptops, and does not require a Web browser. SYRIS™ supports two-way disease information reporting and data sharing for these medical professionals. It provides a fast, reliable, portable method for reporting suspicious or novel symptoms that may be part of a known disease or disease-complex. Reporting is based on symptom complexes known as syndromes. These can be defined with a high degree of specificity (e.g., hemorrhagic fever syndromes) or can be made more general, reflecting common medical care parlance.

SYRIS functions mainly as a data integration tool: data from all SYRIS users is summarized for PHOs to view as temporal graphs and map layers. This architecture greatly facilitates identification of epidemic disease factors and locations, and also provides a means of distributing medical alerts to all SYRIS users.

The design protects patient confidentiality for all types of data reporting because it does not use patient-specific information. PHOs may privately contact a physician or other health professional to obtain patient-specific information if they believe such information is needed to protect public health.

Among other contagious and communicable diseases SYRIS detects many high-risk respiratory diseases including, but not limited to: anthrax, influenza, SARS, West Nile Virus, and avian influenza; as well as six categories of human infectious diseases, specifically: influenza-like illness (e.g. Hantavirus Pulmonary Syndrome), fever with skin rash, severe diarrhea, and severe respiratory distress.

RSVP-EDAC collaborations augmented the early system with geospatial capabilities in the pre-2003 timeframe; the SYRIS-PHAiRS collaboration has augmented the system since 2003 with access to *e*D/SW model outputs and other system capabilities described below for prototypical health alerts and combined health / environmental data analyses. Upon login, the map displays the geographic area that is associated with the local PHO's jurisdiction. When Details are selected for a syndrome, the map view changes to display the geographic area where cases have occurred that are associated with the syndrome. Navigation buttons allow users to zoom in or out, or to move up, down, left and right. The map can be viewed as a series of layers, some of which are more or less transparent, some of which are colored in various shades, and some of which represent borders of counties, states and zip codes. Various features of the map can be turned on or off in "groups" (e.g., all of the human syndromes), or individually. The default setting (when the map is first displayed) is to have all features turned ON.

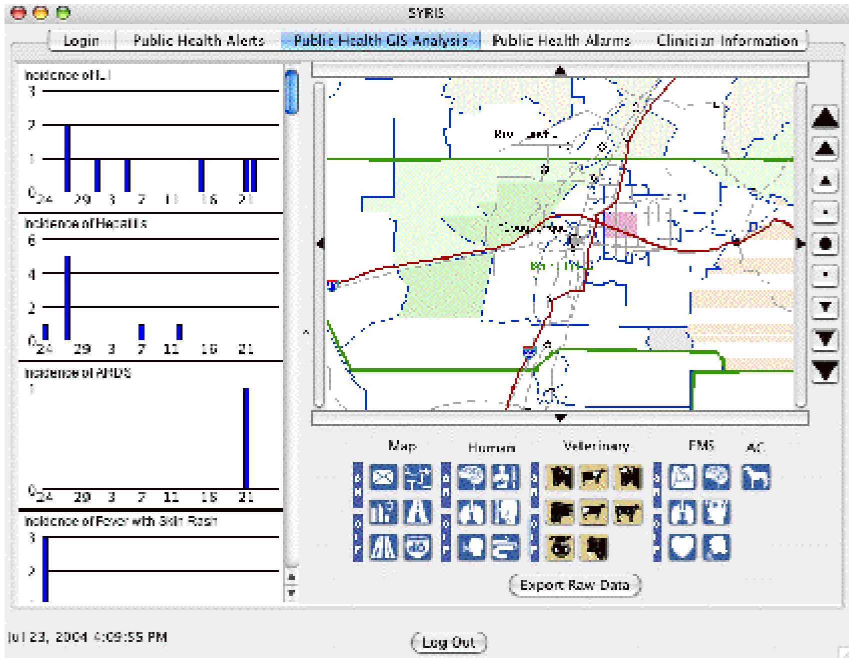


Figure 40. Typical screen for SYRIS™ health reporting. In this example the section for reporting a veterinarian case has been turned OFF, as indicated by the gold color below "veterinary." The data in this graphic are notional and do not represent actual patient or public health data from *any* jurisdiction.

PHAiRS system

*e*D/SW output archive

From the outset of PHAiRs, the intention was to generate a modeling system that incorporated ESR data and that would enhance an existing public health decision support system. Three tasks were designed to generate an archive of *e*D/SW forecasts. These are true forecasts in the sense that they are not based on comparative data from AIRNow or other observation sites. Comparative data are not available for the forecast period because by definition they are not available before the fact. They are used only for testing model performance, for verifying and validating the outputs, and as a supplemental data source for historical trend analyses.

The first task was to generate an archive of 48-hour dust concentration forecasts over the domain using model run 4a. This included a retrospective *twice daily* model run (also using model run 4a) as an ongoing daily model run for the current day. Thus the archiving system is designed to execute three model runs per day (two retrospectives and one current 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 30, 2008, there were 820 separate 48-hour forecast datasets in the archive beginning on January 1, 2006. These are stored on the data servers at EDAC. Most model runs reach completion, but there are days when they do not. On 51 occasions (6.22%) the runs did not complete and had missing data.

Data management and web services

The second task was to develop web services that permit DSS developers and health-care users to search for, access, and download dust concentration data generated by *e*D/SW, as well as data collected *in-situ* by EPA's AIRNow network when they become available from DataFed. Both the historical observations 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 (Figure 41) allows users to search for and download both EPA AIRNow $PM_{2.5}$ and PM_{10} particulate data, as well as 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 combined AIRNow and modeled 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 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 (i.e., organic and non-anthropogenic) and anthropogenically-produced particles.

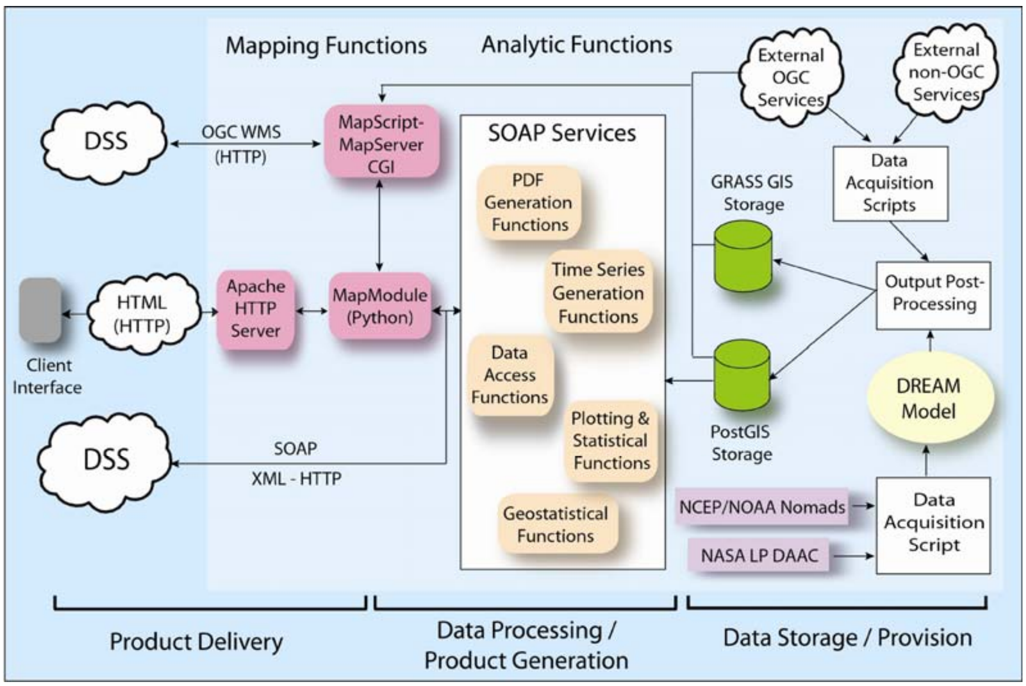


Figure 41. PHAiRS system architecture for providing, storing, processing, and delivering *e*D/SW dust event forecasts and animations, and for retrieving archival data for health research.

Statistical measures

The third task was to create web services and analytical tools that allow developers to generate statistical measures and indices. One of these, the *DREAM Data Access and Statistical Wizard* provides hourly *e*D/SW output from 2006 to present and *in-situ* PM_{10} and $PM_{2.5}$ data from DataFed, and allows one to extract modeled dust values for specified x-y coordinates at specified times and to combine them with AIRNow values to generate statistics. 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.

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 model run, or for a date range specified by the user.

The importance of having a vigorous model performance and verification program is to grow acceptance by users that the modeled forecasts are reliable, accurate, and independent of ground-based instrumentation. Many sites have missing data for lengthy periods, especially for days of known dust events. It is sometimes possible to obtain data from the AIRNow website itself rather than through the *e*D/SW web interface. Also, there are important gaps 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 would improve validation statistics at US stations (Yin et al., 2007), yet to date there are no *in-situ* measurements from Mexico.

Systems integration

Technology integration consisted of two components: 1) development of modularized visualization tools that integrated model outputs, statistical methods and additional geospatial reference information; and, 2) development of automated data acquisition, processing, and integration technologies that streamline assimilation of ESR data inputs and *e*D/SW outputs. A combination of Open Source technologies is being used to meet the requirements of both components.

The visualization component integrates MapServer (an Internet mapping application), R (a mathematical programming and statistical analysis package), and GRASS (a GIS application) through custom Python scripts to present to the user a combination of mapping and analysis tools via a web interface. The web interface provides a standards-compliant (validated against the W3C HTML 4.01 Standard) user interface that may be run on a variety of computer platforms through any standards-compliant browser. The extensive use of server-side processing and scripting also produces a very lightweight client that may be accessed over a wide range of network bandwidths, including dial-up.

MapServer as the internet mapping server application greatly facilitates the deployment of OGC Web Services (initially WMS, with WFS and WCS if needed), through its integrated support for these service architectures. The integration of R and GRASS into the mapping module permits using the wide range of statistical and geoprocessing functions provided by these applications. All three software environments have been successfully integrated into a client interface.

Data acquisition and processing automation have been accomplished through a combination of Python and Bash shell scripts. Shell scripts process MODIS products acquired as multiple HDF files, into mosaiced ArcASCII Grids for assimilation into the *e*D/SW as GRASS rasters for analysis, and GeoTIFF's for data download. More recently, through work performed in collaboration between the PHAiRS team and NASA with George Mason University, MODIS land cover data have been acquired via WCS from the LP-DAAC for use in model initialization. Python scripts have been developed that automatically acquire current GRIB formatted forecast data and reprocess these data into ArcASCII and GRASS raster grids, both for use as *e*D/SW inputs and as data for analysis and visualization within the user interface. These scripts provide the foundation for automated acquisition and processing of ESR and other data sets assimilated into the model.

Key outputs from the *ESR/ESMF Data Provider* include statistical data relating to estimated uncertainty in model outputs (e.g., error bars associated with over/under predicting concentrations). These data are derived from comparisons between observed values from ESR products, ground observations from AIRNow, AERONET, and other contributing *in-situ* networks. Comparisons generate location-specific error estimates for ground observation locations, and continuous fields of model uncertainty derived from extrapolated ground observations and synoptic satellite observations.

In total, these technology developments facilitate timely acquisition of data and provide the tools and interfaces needed to improve public health decision-making using SYRIS^{TM}.

Chapter VII: Final Benchmark

The PHAiRS integrated system solution 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. RSVP was PHAiRS' proposed decision support system. However, between 2003 and 2005, this prototype system morphed into a commercial version called 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. It is suspected that initial 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.

Given the performance statistics from *e*D/SW model runs, the PHAiRS project team is satisfied that Earth observations data can be used to improve dust episode forecasting in the Southwest. The team and its partners are encouraged that these improvements will lead to more timely forecasts so that health authorities can issue early warning alerts. The team is also confident that the mapping services module developed for the New Mexico Department of Health under separate funding from the Centers for Disease Control and Prevention (CDC) will support SYRIS™ users in their day-to-day clinical reporting of respiratory diseases.

Improvements to *e*D/SW

Previous sections have shown that assimilating NASA data into DREAM/SW to create improves *e*D/SW simulations improve the timing correlation (peak hour) of dust storms that sweep CA, AZ, NM, NM, and TX. In consultation with Arizona and New Mexico health and air quality offices, experts in epidemiology and the effects of airborne particles on human respiratory systems, and developers of public health decision support systems, the project team has focused on model output improvements that are most important to these stakeholder communities. They need accurate, reliable, and understandable forecasts and simulations of dust events such as (a) time of arrival and duration of elevated levels of airborne dust, (b) expected concentrations, (c) particle size discrimination, especially $PM_{2.5}$, and (d) the time-dependent spatial extent of dust plumes.

The *e*D/SW using various combinations of ESR parameter replacements improve simulations of the timing and duration of major dust events, at least to a level of probability that respiratory health interventions can be considered by PHOs. Table 21 lists the performance statistics for modeled surface variables. The biggest differences between results from DREAM/SW and *e*D/SW are for air temperature at 2-meters above terrain. The agreement index after NASA data assimilation was 0.95, compared to 0.71 using DREAM/SW parameters. This is a significant model improvement.

Table 21. Performance statistics of modeled surface wind and temperature.

Agreement Index	Wind Speed	Wind Direction (degree)	Temp (K)
DREAM/SW	0.74	0.74	0.71
*e*D/SW	0.75	0.76	0.95

The project has demonstrated the additional value of the Point Stat Tool as an alternative approach to verification and validation of *e*D/SW's performance. The team has tested also the use of animated 2-D and 3-D visualizations of model outputs to improve user/stakeholder understanding of model capability and to identify types of outputs most useful for user applications. This has generated several suggestions from users.

The test cases illustrated here indicate that the model can accurately predict the timing of major dust events. The timing of modeled peak hours (maximum hourly concentration) during dust events is typically within three hours of the observed peak hour in all case studies and at all locations in the model domain.

*e*D/SW typically overestimates the magnitude of both $PM_{2.5}$ and PM_{10} hourly averages during dust events observed in the model domain.

Performance stats derived from point-by-point comparisons can be misleading; timing lags and concentration averaging can be used to improve model verification.

The MET Users Guide (ref) provides model verification tools ideal for use with *e*D/SW. The Point-Stat method for evaluating model performance shows great promise.

Performance stats were calculated for the 2007 test cases using the Point-Stat tool. As indicated in Table 18 (page 44), only ten 'hits', or exceedances, of the 'dust threshold' occurred over the entire model domain during these events.

The Phoenix case study Table 19 (page 43) indicated that the *e*D/SW can be used in much the same way weather forecasters predict the possibility of rainfall over metropolitan areas. *e*D/SW had a threat score of 0.65 from January – April 2007, which implies that the odds of hind-casting a dust event somewhere in the Phoenix metro area during that time period were roughly two-out-of-three.

Using these combined values yields the following performance stats with the Point-Stat tool. These results indicate that *e*D/SW successfully forecasted 71% of the hourly averages and only 29% the hours exceeding the dust threshold during these two events. Although a new tool for evaluating model performance has been introduced, the point-by-point comparisons are still being used. The approach has been modified using the Phoenix metro area having multiple observation points.

Improvements to SYRIS™

Dr. Alan Zelicoff, principal developer of RSVP and of its successor SYRIS™, is excited by the dust forecasting capabilities produced by *e*D/SW. "The new version of SYRIS contains extensive modeling and disease prediction tools, including environmental diseases. The latter is especially important in daily clinical practice (in both veterinary and human disease) as dust particulates ($PM_{2.5}$), nitrous and sulfur oxides, and ozone clearly increase the incidence of acute lung disease and respiratory symptoms in a given area. Distinguishing such environmental illness from infectious diseases is a very difficult clinical challenge. Thus, atmospheric data combined with a dust model may be very useful for clinicians in their daily practices. Such predictive models may enable emergency rooms and clinics to prepare for an increase in patient visits or may enable public health officials and physicians to contact patients who may be advised to change medication or behavior in anticipation of an environmentally induced exacerbation of chronic lung or cardiac disease.

As an outcome of the 4th Annual Review (March 31-April 2) an effort was agreed upon to deploy the historical and ongoing model runs into SYRIS™. A roadmap for this activity has been developed and a meeting with Texas Public Health officials was held with the goal of bringing the TX Public Health Officials online to access the dust forecast models. The potential for enhanced capabilities was well received. It was decided that, to facilitate the delivery of model products through SYRIS, an MOU would be developed between the TX Public Health Officials, the University of Arizona, and ARES Corporation to define clearly the appropriate use and guidelines for integrating dust model animations into the SYRIS system. The combination of animations and associated syndromic data from SYRIS will be used to examine correlations between exacerbation of COPD and asthma with forecasted dust events. While the study will extend beyond the end of the PHAiRS project, preliminary results will be included in the final report for the PHAiRS project, tentatively scheduled for January 2009.

Other health system improvements

The project team actively engaged stakeholders in New Mexico, Arizona, and Texas to help develop a dust forecasting module that enhances their syndromic surveillance systems. Representatives from health and air quality offices in these states have participated actively in projected activities that will improve their toolset and decision-making capabilities.

NM Department of Health (EPHTS)

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 *e*D/SW model outputs and AIRNow measurements for the Lubbock and Midland/Odessa areas. 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 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 *e*D/SW 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 network-accessible resources 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 *e*D/SW 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 capturing and presenting air quality and health data in a consistent and statistically valid manner. An initial capability for the summarization of *e*D/SW outputs for counties has been developed as part of the PHAiRS SOA, providing daily summaries for regional model outputs (as opposed to single model cell/pixel).

Arizona Department of Health Services

The Office of Public Health Emergency Preparedness and Response at the Arizona Department of Health Services (AZ/DHS) detects and responds to natural or intentional disease events. Funded by the Centers for Disease Control and Prevention (CDC), the Office is composed of several program areas, one of which is Electronic Disease Surveillance. Under this program, the Office is developing a web-based application to enhance disease surveillance and to detect bioterrorism events in Arizona,.known as the Medical Electronic Disease Surveillance and Intelligence System (MEDSIS). AZ/DHS is very interested in using outputs from *e*D/SW to enhance their electronic surveillance tools.

Dr. Trujillo offered these comments: "*As a syndromic surveillance epidemiologist, I am always searching for useful sources of data to t*rack syndrome illnesses that I can add to my pr*ogram. One of the problems with disease surveillance in general is that we do not know when and where events are going to take place and therefore we are reactive, not proactive. Another problem specific to sy*ndromic surveil*lance is that, with the non-traditional data sources commonly used in syndromic surveillance, there is no common user interface. We must use many different programs and softwares to visualize and analyze the data. Based on the demonstration* we at Arizona Department of Health Services were shown, the *[enhanced] Dust Regional Atmospheric Model has the potential to add to existing data sources for syndromic surveillance. First, the dust storm model can help predict when and where respiratory ill*nesses are potentially going to i*ncrease, which is a much needed addition to disease surveillance tools. Being forewarned about the possibility of dust storm-related illnesses will help health officials better cope with the resulting illnesses. Second, the* model seems simplistic enough to integrate into existing programs i*nstead of requiring its own user interface and program. I understand that it will be possible to format this model to be added as an extra button/tab built into existing visualization sys*tems. This aspect alone will increase the utility of the model for sy*ndromic surveillance. If the PHAiRS program can help us prepare for events and be integrated into current program operations with such ease, it will be a very welcome and useful tool."*

Pima County AZ

The Pima County Department of Environmental Quality is interested in improving forecasts of airborne dust events that affect human health. Visualization techniques would promote this effort. After experiencing a prototype virtual reality 3-D visualization of *e*D/SW output, Beth Gorman wrote: "*The visualization of the data was an exciting way to see the numbers on a page come to life. It was especially intriguing to watch changes in the dust plume over time and* from different perspectives. Our department is looking forward to continued coordination with the U of A and others to develop a method of forecas*ting airborne dust events to protect the many individuals who are at risk in our community.* Wayne offered further comments: *The DREAM model visualization was quite interesting. It provided a virtual look at the formation of a dust event with indications of the originating area. I believe with some modifications it might prove useful in pinpointing sources of dust* events which could prove useful in remediation. It would be more useful if the values for dust content, elevation, and wind speed could somehow be indicated in the visualization. Overall, I think it is a good begi*nning."*

Users outside of the medical community also find promise for the dust forecasting potential of *e*D/SW's output. Margaret Fowke from NOAA wrote in an email message to Dr. Sprigg: "*If you or other colleagues are interested, I would love to have your involvement in developing health impact s*tatements related to dust issues that potentially could be delivered on the air by professional broadcast meteorol*ogists and/or warning coordination meteorologists. I have been working with another public health group affiliated with Tufts University and Un*iversity of Colorado focused on increasing physical activity a*ccording to weather.*

City of Lubbock Health Department

Since early 1999, the City of Lubbock Department of Health has evaluated "syndrome-based" disease surveillance systems (SBDSS). That office has provided a preliminary summary (Appendix 3) of its assessment of SBDSS, in meeting the following needs of public health services. The key points of their assessment are itemized below:

In theory, SBDSS's, by virtue of their timeliness and volume of information flows, could assist in meeting these central public health responsibilities. However, in practice, the specific designs, underlying technical features, scientific approaches, and ease-of-use are dramatically different across the dozens of SBDSS's currently in existence. Some of these have been implemented only in narrowly defined demographic settings or other limiting service areas. The promise is often not met in real-world use.

It is also important to note that the overwhelming majority of SBDSS data gathering focuses solely on human patients, despite the fact that in all significant outbreaks of novel diseases over the past decade in North America, animals were the primary source of the diseases. In particular, very large or economically significant disease outbreaks among humans had animal sources.

We [the Lubbock Health Department] found that all of the "automated" SBDSS systems (that is, data mining systems, as opposed to active syndromic surveillance systems, were problematic in several key areas: timeliness and accuracy. Of importance to PHAiRS was that information was nearly always reported in tabular or textual format without accompanying geospatial tools for analysis.

RSVP (later SYRIS) was the only active SBDSS available to compare with passive systems. Both defined six common syndromes worded in the daily parlance of medicine and public health, and further provided an electronic interface that operated on virtually any computer connected to the Internet (Zelicoff et al., 2001). It also provided primitive, but useful, geographic mapping tools. Experience was generally positive. Contrary to the belief that physicians would not take time to enter cases, physician compliance was high because the number of cases of seriously ill patients who fit into one of the syndrome categories was, on average, one case per month per physician (except during large epidemics). Further, the system provided information of immediate clinical importance to physicians, thus increasing their cost-effectiveness in practice. Finally, on rare occasions, the system enabled public health officials to contact doctors within minutes of a case report when the data suggested unusually worrisome symptoms that might require immediate contact investigation. Thus, RSVP/SYRIS reduced the time from initiation of a case report to syndromic information from days to mere minutes.

Model verification issues

The *e*D/SW Data Access and Statistical Wizard web interface provides hourly model output from 2006 to present and the corresponding observed PM_{10} and $PM_{2.5}$ data obtained from the EPA's AIRNow network via the DataFed system. Within the present *e*D/SW domain, the website has 41 sites from which to download PM_{10} (modeled and observed, side-by-side for comparison) and 94 $PM_{2.5}$ sites. Several errors and discrepancies were observed during verification efforts. Time lags existed between modeled and observed concentrations that were explained by time zone and corresponding time stamp corrections. Initial verification required manual adjustment of downloaded data to match *e*D/SW output (in UTC) with observed (local time) by the hour.

Many monitoring sites have missing data for long periods in the time series, especially for days of known dust events. It is suspected that sensors fail for extreme concentrations and/or the reporting of these events takes longer than the usual automated system. Consequently, it is impossible to verify certain dust storms at locations where there is a lack of data. It is unclear how many sites within the observed network have this problem, but it is often observed 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 the DREAM web interface. There is an obvious gap in site coverage for PM_{10} in central Texas, a region known to experience widespread dust events. The sensors are located primarily in urban cities, making verification for rural areas difficult. It has been shown that increased land cover data sets that include Mexico in the model domain improve verification statistics at US sites yet there are no observed measurements from Mexico to date.

It has been shown also that the MOD12Q1 data for Mexico (within the modeling domain) would improve validation statistics at US stations. However, to date there are no *in-situ* measurements from Mexico.

New technology development

An adjunct interoperability project funded by NASA's Geoscience Interoperability Office (GIO) was initiated during the second quarter of 2007. The PHAiRS team partnered with George Mason University to improve services by enhancing interoperability capabilities between the PHAiRS project and NASA's data services. The project developed a high-performance computing version of the *e*De model for execution in grid and High Performance Computer frameworks which, in turn, speed delivery of products from the system. This includes development of an initial KML file that makes use of the time-enabled WMS services for the current collection of *e*De outputs. The KML file was successfully used to view an interactive animation for a specific dust event early January 2008 in Google Earth. An automated process for routine acquisition and processing of updated land cover data via a WCS service hosted by the LP DAAC was successfully developed. A collaboration web site was established as part of the PHAiRS Interoperability project that is providing a complementary workspace for the interoperability team. Membership in that collaboration site significantly overlaps with the PHAiRS team.

During the first two years of the PHAiRS project, the team developed technologies and scripted programs to support both the research and the applications development components of the project. The new technology development activity has become so integrated with the other components of the project that it was no longer reported as a stand-alone activity. Many of the technological enhancements are incorporated in this report. The team continues to engage in technology advancements as project needs arise and will report them accordingly.

Data uncertainties

Replacement data sets

Data assimilation, itself, is a multifaceted process hampered by the general absence of metadata. One must first compare the attributes of existing model inputs and of possible satellite data replacements. Like DREAM/SW, many models currently used for Earth system science were designed without benefit of remotely acquired data sets. Data compatibility issues therefore have to be considered, including: (a) measurement units, (b) x,y,z resolution, (c) temporal frequency, (d) map projection and ease of re-projection to fit model requirements, (e) file formats, (f) error and error propagation, and (g) validity of the replacement data in terms of enhancing or improving model outputs.

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. This requires a separate investigative report.

Dust speciation

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 would be to verify and validate these fractions continuously on the basis of individual species' concentrations; but, this would require a separate investigative report.

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 more quickly. Because *e*D/SW 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 *e*D/SW. Anthropogenic concentrations often are present when *e*D/SW 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 wind-generated PM_{10}. During non-windy conditions, it is still possible to observe other sources of PM_{10} that *e*D/SW 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, both anthropogenic (as exhaust) and natural (by plants), and others react in photochemical reactions that act as sources and sinks of particles. Elemental carbon in $PM_{2.5}$, also called black carbon or soot, is produced by combustion. Trace metals are produced via human factory emissions 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. The *e*D/SW is concerned only with mineral dust, but other components of $PM_{2.5}$ complicate measurement of particulate concentrations, and therefore model performance (Shaw, in press). Total $PM_{2.5,}$ as referred to here, is the net concentration of all species in the air for that size range. The *e*D/SW 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. The *e*D/SW can only model the former, so distinguishing the two using speciation would be 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 can be used for speciation, so *e*D/SW can be validated discretely only at this frequency. Continuous hourly data are ideal, but are probably not feasible due to cost and time constraints. The cities in EPA's Speciation Trends Network (STN) are mainly large metropolitan areas that monitor anthropogenic species. The soil component usually is small in proportion to other species at these sites, but it is assumed to be larger in non-urban areas that are routinely exposed to desert dust rather than urban pollutants. 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), 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 investigation.

Health data

Uncertainties in public health data far surpass those for environmental measurements and modeling. Health data are dogged by internal genetic and eternal environmental unknowns that, for the most part are not controllable. 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, school nurses, ER personnel, physicians, and others to treat the case along prescribed best practices aided by patient history.

Exacerbating these uncertainties 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 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 (ICD-10) 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.

Uncertainties in health statistics begin with exposure information and propagate throughout the analyses and interpretation of aggregate findings. Uncertainties in ESR data and associated models lie in sensor design, the algorithms defining sensor products, and assimilation requirements. In the first category PHAiRS' approach to reduce uncertainties was to create and analyze large datasets statistically over longer periods of time. In the second category the approach was to use NASA sensor data to reduce two of the largest modeling uncertainties; emissions and the initial and boundary conditions.

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.

Chapter VIII: Outreach & Transition

Outreach efforts were to engage user communities and to gain their confidence in adopting PHAiRS products and services into future routine practice. They were also targeted at the scientific community to help verify and validate the model's performance and to establish a groundwork for collaborative efforts. The former type of outreach focused on local and state epidemiologists, asthma registry personnel, school nurses, and other public health professionals and stakeholders. The latter focused on the science behind the results such as the modeling approach, utilizing Earth science results, verifying and validating techniques, and developing a sophisticated, yet easy-to-use data delivery model. In the PHAiRS project, outreach began early with team members participating in, and presenting at, local symposia and conferences, and progressed into national and international meetings. At each of these opportunities, team members presented material on PHAiRS that described phases of the project, such as modeling results, verification and validation results, data delivery architecture, and stakeholder involvement. Outreach also included dialogue with expert modelers, health community professionals, weather forecasters, and environmental scientists, all of whom were eager to learn more about PHAiRS and how its methodology and results might benefit their work.

Engaging stakeholders

As PHAiRS began, a strategy was mapped for transitioning dust episode forecasts into health services. Four steps were identified: (a) adapt the simulation and forecast model over the southwest US where airborne dust is a health hazard, a concern of local and State air quality and health offices; (b) develop interest in the technology within health and air quality offices of local, State, and Federal agencies; (c) develop a user-friendly, client interface between the operating model and end users; and (d), help local and State offices use the client interface in simulated operations to test the dust forecast system and products. Arizona's Pima County Department of Environmental Quality and New Mexico's Department of Public Health have been involved from the very beginning of PHAiRS. The National Institutes of Health and the Centers for Disease Control and Prevention have been more casually interested and, hence, less frequently engaged. As the model outputs began to demonstrate dramatic improvement in dust forecasting, and the web-based client server became functional, other end users in the region became interested. Stakeholders have been invited to, and participated in, project team meetings and annual reviews. They also have been included in the distribution of quarterly reports, the Initial Benchmark Report, and the Verification and Validation Report.

As the project progressed, team members visited local and State health and air quality offices to demonstrate PHAiRS results and to discuss how these products could be integrated into their decision support systems and practices. Visits were made to the City of Lubbock Public Health Department, Texas Public Health District 1, Pima County (AZ) Department of Environmental Quality, and the Arizona Department of Health Services. Meetings were held with epidemiologists in the New Mexico Department of Health to discuss sources of health data in general and in the region, and to coordinate project progress and plans with PHAiRS partners at UNM. Additionally, meetings were held with the City of Albuquerque Environmental Health Department/Air Quality Division and the Albuquerque Public Schools Asthma Registry, and Asthma Allies. The Arizona Department of Health Services participated in a meeting in Tucson to explore use of the model in addressing Valley Fever, which is caused by ingesting spores borne aloft in dust storms. Arizona State University (ASU) has become interested in collaborating with PHAiRS partners. Show-and-tell discussions were held in Phoenix to survey potentially complementary skills and aims. Dr. Jim Anderson at ASU, for example, is very well respected for sampling and chemically analyzing airborne particles. Others at ASU use the EPA CMAQ for simulating airborne particulate matter, particularly for regulatory purposes.

The project team presented its work at technical meetings throughout the term of the project. One of the first was the Research Association of Medical and Biological Organizations (RAMBO), a local gathering of medical doctors and scientists from Los Alamos National Lab, University of New Mexico, Sandia National Labs, State offices, and private practice. PHAiRS was represented at monthly RAMBO meetings, and team members have presented stages of the project at annual RAMBO conferences. An outcome of interaction with RAMBO led to contact with stakeholders in Lubbock, Texas, specifically the city of Lubbock and the Health Sciences Center at Texas Tech University.

These are but a few of the outreach actions the team has taken. An analysis of why some early adopters are slow to embrace this new technology reveals several inhibiting factors, common to all, overcome by only a few. Offices must contend first with proof of performance and reduction of risk in taking on the new technology. Additionally, in varying degrees, they must contend with understaffing, the "not-invented-here" syndrome, commitment to other systems, fear of the unknown, fear of change, and protecting status quo. The team believes that public knowledge and familiarity with new technology can overcome all but the first of these inhibitors. It must be emphasized however, that all technology transfer must be predicated on "proof of performance and skill."

Collaborations at the national level

As the outcomes of the PHAiRS project have been demonstrated, it has attracted the attention of scientists and communities of practice nationally and internationally. This recognition opened doors for new opportunities to extend PHAiRS capabilities to other air quality and health applications.

One of the desired outcomes of the project was to produce results and products that are extensible from a regional to a national application. To this end, the team extended outreach activities to include presentations at, and participation in, national and international symposia, conferences, and workshops; as well as visiting key health and environmental offices and agencies. These efforts netted responses from the US-Mexico Border Health Commission, the Center for Hydrometeorology and Remote Sensing at the University of California – Irvine, the Naval Research Laboratory (NRL), the National Weather Service (NWS)/National Centers for Environmental Prediction (NCEP), and NASA's Geoscience Interoperability Office to name a few.

Though interactions with these groups occurred at national meetings, most of the potential for adapting or extending PHAiRS modeling and products were aimed at international applications. For example, dialogue with Dr. Lawrence Kline, Commissioner of the US-Mexico Border Health Commission, was initiated as a result of a presentation given by the team at the International Conference of the American Thoracic Society in San Diego, CA. Dr. Kline would like to see PHAiRS perform quasi-operational dust forecasting throughout the border region. Discussions with Dr. Kline are continuing. The Center for Hydrometeorology and Remote Sensing at the University of California – Irvine wants to collaborate with PHAiRS for dust modeling and applications of the model in arid and semi-arid regions around the world.

The team was invited to the Naval Research Laboratory (NRL) in Monterey, CA to share ideas about dust storm modelling, simulations and predictions. The NRL develops dust forecasting tools and provides operational dust forecasts for U.S. military operations worldwide. These discussions led to an agreement that NRL would provide operational fields of global and regional atmospheric aerosol loading to PHAiRS to help solve one of the team's problems: "fugitive" and "background" dust that "contaminate" the model's predictions of dust concentrations; especially important since these results are compared to EPA dust measuring sites for verification and validation. The NRL agreed also to participate in a Pan-Am Centre, providing the same operational products.

Collaboration was initiated with the NWS regarding delivery of health services based on environmental information. NWS formed an interagency task group to examine plans for establishing a health office within their agency. The interagency leader requested input from the PHAiRS project to create a "health effects of dust" component for a Public Health Training aid. Through this collaboration, NASA was invited to participate in the interagency task group. NWS and the PHAiRS team also outlined joint strategies for dust model development, testing, evaluation, and eventual use by the NWS field offices.

A cooperative effort stemming from the Earth Science Information Partners (ESIP) Federation included the PHAiRS team, NASA's Geoscience Interoperability Office, and George Mason University. These partners developed a separate, but related, project to improve the flow of data into the model and visualization system by integrating NASA interoperable data services for more current data (i.e. land cover). This integration allows faster delivery of the model outputs.

Testing the PHAiRS client

Training workshops

The PHAiRS team conducted a half-day training workshop for stakeholders and partners in Albuquerque, November 2007. Goals of the workshop were to demonstrate the capabilities of the dust forecasting tool and data/information delivery system to public health officials and air quality experts in the local area, and to provide a "hands-on" session for exploring the toolset. Attendees represented the City of Albuquerque Air Quality Office, the NM Department of Health, UNM Health Sciences, and ARES Corporation (developer of SYRIS™). Each participant brought their own laptop, and through a wireless connection was able to access the PHAiRS online client and database. A resource book and DVD-based movie illustrating how to use the PHAiRS dust forecasting tool were provided to each attendee. A few suggestions were made by participants for improving some elements of the tool, but most participants felt the toolset was adequate for its intended purpose at this time. Plans were to extend the training workshop to stakeholders in Arizona, but due to conflicts and difficulties in scheduling, those workshops never took place.

Lubbock test

Health data are being assembled to review and analyze retrospective dust events from the panhandle of West Texas. Dust from Eastern New Mexico is such a perennial problem in West Texas that validating its health impacts on populations at risk is a core goal of this project. Records have been kept in West Texas for the past two decades. Over 100,000 records of respiratory illnesses have been drawn from a variety of sources and aggregated to the census block level. These records include detail on asthma, influenza, mortality, behavioral- and risk-factor surveys, clinic files, and hospital discharges.

Team members met with Texas Public Health officials who are using SYRIS™ currently, to present the PHAiRS products, describe the model and its outputs, discuss the V&V results, and identify steps for bringing the TX Public Health Officials online to access the dust forecast model products. Attending the meeting were officials from the City of Lubbock Public Health Department and the Texas Public Health District-1. The potential for enhanced capabilities was well received, and it was decided that effort would be made to deploy the *e*D/SW outputs into SYRIS for access by the Public Health officials licensed to use SYRIS. To facilitate the delivery of model products through SYRIS, an MOU was developed between the TX Public Health Officials, the University of Arizona, and ARES Corporation clearly defining the appropriate use and guidelines for integration of *e*D/SW output into SYRIS. The combination of dust model outputs and associated syndromic data from SYRIS would be used in a prospective study examining the correlation between exacerbation of COPD and asthma and forecasted dust events. While the study would extend beyond the end of the PHAiRS project, preliminary results will be included in the final reports for the PHAiRS project. To perform this analysis SYRIS needs to have additional human and veterinary syndromes added to the system, including exacerbation of COPD and asthma for humans.

Future opportunities

The modeling capabilities, output products, and delivery system exhibited by the PHAiRS project offer a unique package that has attracted the attention of several international bodies, especially those in Asia and Europe. These interests represent opportunities for collaborations beyond the life of the original project. Some of these contacts were made as a result of presentations given at conferences, while others occurred through direct dialogue with individuals in specific agencies and organizations such as the U.S. Geological Survey (USGS), the Centers for Disease Control and Prevention (CDC), the China Meteorological Center, the World Meteorological Organization, and the Group on Earth Observations (GEO).

While the domain and boundary conditions for the DREAM model have been set for dust forecast modeling over the southwest United States, requests have been directed to the team for expanding the domain to include other geographic regions. The USGS would like to see the domain broadened to include all of the western United States and the CDC is interested in extending the forecasting capabilities to a national level. The ground work established by the PHAiRS project will be used in part to enhance the CDC's Environmental Public Health Tracking System (EPHTS).

Opportunities in China

The Chinese are very interested in the PHAiRS project and invited team members on several occasions to participate in conferences and symposia, as well as to give lectures on the project to audiences within their water, climate, and atmospheric agencies. Lectures were given also at a few of the Chinese universities. The Chinese Ministry of Water Resources (MWR) invited team members to assess current issues of climate variability and change in the Haihe River basin which includes Beijing and Tianjin, eight provinces, and over three percent of China's territory. The river system is seriously compromised because of drought and increased pollution due to agriculture, industry, urbanization and wind blown dust contaminated by various industrial pollutants. It is probably true that dust blown in from the Taklamakan and Gobi deserts, as well as dust sources in and around Beijing and Tianjin, bring heavy metals and other industrial contaminants into surface water systems and cisterns affecting potable water supplies. It is assumed, however, that other water polluters are overwhelmingly accountable for most of the water pollution in the basin. Several meetings took place in Tianjin (home of the Haihe River Conservancy Commission, hosts) and Beijing (with headquarters staff of the Ministry). No follow-up action, other than tracking progress in dust simulation, forecast, and public information access, was proposed.

Visits to other institutions were arranged to share information and strategies with others involved in soil erosion, analyses of airborne particulates, dust storm and particulate pollution forecasting, regional dust modeling in global climate models, and the climatic consequences of airborne dust. Visits also were arranged with the China Meteorological Center (meeting with the Beijing Climate Center and National Climate Center) and the Chinese Academy of Sciences' Institute of Atmospheric Physics (CAS/IAP). The visit revealed that Chinese scientists are more than adequately funded and that the link between research and government operations is much improved over previous visits. The ministries are funding much of the work in pollution measurement and control, but substantial funding also comes from the Chinese National Science Foundation that supports basic research in aerosol science, dust storm prediction, and airborne dust forecasting. The NASA PHAiRS project leads the way. The MWR, the CAS Institutes, and the universities all would like to work with the PHAiRS team. They have followed PHAiRS publications and presentations.

In July 2008, taking advantage of the ISPRS Congress in Beijing several executives met to discuss adaptation of PHAiRS technology to regional particulate air quality problems in Asia. It appears that shared responsibility would: (a) simulate and predict generation, entrainment and downwind concentration of dust in and near the Asian source regions at high resolution using a modified *e*D/SW; (b) simulate and predict PM diffusion at moderate resolution from an Asian domain modified from *e*D/SW for the service area of China's SDS WAS Centre; (c) simulate and predict PM transport from Asia across the Pacific

Ocean, North America and around the World; and (d), develop the PHAiRS web-based client server to satisfy Asian air quality and public health services' needs, using Korea and India as the first prototype.

World Meteorological Organization

Another significant outreach effort has been with the World Meteorological Organization (WMO). PHAiRS technology has been integrated into the draft Implementation Plan for the WMO and the Group on Earth Observation (GEO) "Sand and Dust Storm Warning Advisory and Assessment System." PHAiRS' Co-PI is a member of the SDS WAS Steering Committee and chaired the IP drafting group. PHAiRS attracted attention because it is the only forecast system of its kind. Its unique qualities are that it has been: 1) initialized by satellite surveys, 2) verified against actual *in-situ* measurements, 3) integrated into a national weather service for long-term dependability, and 4), developed in partnership with local public health services.

The first draft of the SDS WAS plan (15 March 2007) includes a Global Assessment Centre (GAC) that draws from PHAiRS' successful collaboration between Earth system science and public health services. The GAC, if funded, would assess global atmospheric particle concentrations as background for dust forecasts, blending global, hemispheric, and regional analyses from model-generated and satellite-based sensors. The regional dust forecast models would then initialize the forecast period with estimates of dust imported from outside the model domain, and improve correlations between forecasted and observed dust concentration. Initially the new GAC would take advantage of PHAiRS modeling and existing NASA Earth observations but would anticipate taking advantage of improved data from NASA's GLORY instrument. Since SDS WAS is likely to be a very long running WMO program, the demand for GLORY data should persist for the life of the new NASA instrument. And, PHAiRS would have proven that space-based remotely sensed data applied in dust simulation models improve public health, not only in the United States, but around the World. The GAC also would assemble a global picture of dust source regions, seasonally adjusted, to assist SDS WAS Regional Forecast Centres and the global and hemispheric analyses of atmospheric dust. Source regions vary throughout the year, and PHAiRS studies demonstrate the importance of refreshing the source scene. Current and future NASA satellite constellations for monitoring vegetation state and precipitation would be used extensively in this GAC task.

In early November 2007, the draft SDS WAS plan was presented to an assembly of more than 100 scientists assembled by the WMO and GEO to review steps taken to date in launching the dust warning system. Lessons learned through PHAiRS have been integrated into the international program plan: (1) high reliance on satellite sensing for detecting storms, monitoring dust plumes, and measuring and monitoring ground cover and dust source regions; (2) a multi-disciplinary approach that includes guidance at from the global and international health services communities who would have continuing responsibilities in reviewing and steering the SDS WAS; and (3) keeping an eye on new opportunities afforded by the "A-Train." Through this exercise, Sprigg was asked by WMO and GEO to form a Pan-American dust Centre as one of the core four SDS WAS "Centres" to cover the globe. Implied by this invitation is that PHAiRS exemplifies the activity within these proposed Centres: (1) developing simulation and forecast models; (2) demonstrating use of NASA satellite remote sensing data to improve range, diversity, and accuracy of model products; (3) developing networks of collaboration with national and international entities to speed technology transfer into (health) applications; and (4) developing and demonstrating information services that would benefit human health, disaster response, highway and airline safety, and many other environmental and economic sectors.

GEO/GEOSS

The Group on Earth Observations (GEO) is coordinating efforts to build a Global Earth Observation System of Systems (GEOSS) in response to outcomes of the 2002 World Summit of Sustainable Development and the Group of Eight (G8) leading industrialized countries. Nine societal benefit areas (SBA) have been defined in the GEO 10-year Implementation Plan. The GEO User Interface Committee's (UIC) mission is to engage the communities of practice in each of these SBAs to develop and implement GEOSS so that it provides the data and information they need to address local, regional, and global problems. Health is one of these SBAs. The PHAiRS project has elements that complement the goals of GEO and has caught the attention of the UIC. Team members from PHAiRS were invited to participate in the UIC, and have given several presentations on PHAiRS at Committee meetings with specific emphasis on engaging the public health community of practice. With some modifications to the model, PHAiRS products could be extensible to the GEO user community. On behalf of GEO/GEOSS, and in collaboration with the International Electronic and Electrical Engineers (IEEE), the Open Source Geospatial Consortium (OGC), and the International Society for Photogrammetry and Remote Sensing (ISPRS), the PHAiRS team organized two international GEOSS workshops that focused on environmental impacts on human health; one in Goa, and one in Beijing.

Goa:

The first was a one day workshop on Applications in Public Health that was held in Goa, India. It featured the GEOSS architecture and how it can meet needs in public and environmental health. The workshop provided a high-level forum for addressing the benefits and challenges of advanced global information system implementation for societal benefits.

Beijing:

The second workshop was held in Beijing, China and focused on Air Quality and Human Health. The purpose of this workshop was to introduce ministerial-level decision and policy makers to appropriate and affordable space-based technologies. The workshop was attended by 25 people. The role of GEO and GEOSS including GEONETCast was very prominent in the workshop. Participants saw how easily this existing program could enhance cooperation around the globe. Frequent references were made to applications that could transfer observational and modeled data and information. A demonstration of GEO and information sharing provided several examples of how the user-friendly data access system could be applied in many different ways at affordable and sustainable costs. There was interest in obtaining the online demonstration so that this information could be provided to people who were unable to attend. Considerable progress has been made over the past eight to ten years where today quite reasonable depictions of dust mobilization at the surface of the earth, entrainment into the atmosphere, and downwind dispersal of the dust are now almost routine. Presentations were given on modeling dust processes in Japan, Korea, China, and the US. Many other modeling efforts in other countries were referenced.

The examples given here are testimony of the impact the PHAiRS project has at both national and international scales, and fulfils the requirement that the project is extensible to national and global issues and priorities.

Media interest in PHAiRS

Dr. Sprigg was invited to present a paper at the AAAS Annual meeting in Boston, 14-19 February 2008. He presented progress in PHAiRS and its relevance to the global problem of airborne dust. The print media, the Washington Post, the Arizona Republic and the Arizona Daily Star, ran articles based on interviews with Dr. Sprigg following his presentation. Arizona Public Television (KUAT) aired a short piece on "Arizona Illustrated," which was followed by an interview broadcast on Public Radio. The German newsmagazine, *Der Spiegel*, interested in the global context of dust storms and current science and technology to address them, sent a reporter and photographer to interview Dr. Sprigg; the article should appear around September 2008.

Media interest in PHAiRS, air quality, and human health has peaked. Several media services contacted project team members for interviews and information. On February 6, 2008 The Washington Post ran an article by Staff Writer Doug Struck, *Dust Storms Overseas Carry Contaminants to U.S.* Both Stan

Morain and William Sprigg were interviewed for this piece. In another instance, David Coles requested information and graphics to support a story he was developing for The News Hour with Jim Lehrer. It is unclear whether or not this story was aired. Local television stations in Arizona also followed suit, contacting William Sprigg for interviews. The Assistant Editor of *South Pacific Science Press International* and *Position Magazine* and *Spatial Business Newsletter* has requested follow-up interviews from Beijing workshop concerning PHAiRS, the SDS WAS, and dust storms.

Publications and presentations

The PHAiRS project was promoted through publications in refereed journals and proceedings of professional conferences, symposia, and seminars. Articles appeared in several issues of *Atmospheric Environment*. Papers appeared in proceedings of conferences worldwide, including venues in the United States, Russia, China, India, Costa Rica, Saudi Arabia, and Italy. A list of all the publications is presented in Appendix 3.

Over 40 technical presentations were given by the project team at conferences, symposia, and seminars around the globe. While many of these were oral presentations, posters also were prepared and presented at these some of these events as well as to the New Mexico State Legislature. A comprehensive list of presentation and poster titles is given in Appendix 4.

Chapter IX: Summary and Conclusions

V&V summary

V&V efforts began in 2004 by generating a time series of aerosol contour maps and visually comparing these to DREAM/SW aerosol loading contour maps. The generated contours included aerosol data, but no meteorological or geographical data, from thirteen EPA AIRNow continuous air monitoring stations in NM and TX. These results were poor as there was little detail in aerosol contour maps and just as little agreement between these and the modeled contour maps.

Since 2005 more AIRNow stations were added (n = 40) and maps were generated with the hourly station data plotted as points rather than contours. Detail was greatly improved. It was easier to superimpose aerosol data points onto the model outputs, but correlation was still poor. Consequently, a visibility analysis tool using both aerosol point measurements and meteorological parameters was identified to illustrate visibility impairment due to blowing dust. This tool was used also by the TCEQ. The contour maps were generated with DREAM/SW, and were compared to the visibility contours available from TCEQ during the widespread Texas dust storm of 15 December, 2003 (Case 2). The two model outputs showed good agreement with the visibility contours in both the location and timing of the dust event and subsequent dispersal across Texas over the following 24 hours. However, there were still difficulties in correlating visibility observations from DREAM/SW simulations, most probably because the model does not account for ambient air quality particulates. These particulates degrade statistical comparison with the model outputs.

As DREAM/SW transitioned into *e*D/SW, a variety of replacement parameters were used to improve model performance. These resulted in model ‘runs’ with different input parameters and different output data sets. These data (hourly $PM_{2.5}$ and PM_{10} particle concentrations) were compared to the corresponding observed data during test cases, and a set of performance indices were calculated for each model run. The first set of indices was included in the *Initial Benchmark* (Morain and Sprigg, 2005). Each successive model run was compared to assess whether the input resulted in improved model performance.

Model results from two test cases in December 2003 using different land cover data sets were compared in detail (Yin, 2007). Measurements and observational analyses such as surface synoptic data, surface Aerodrome Reports (METAR), upper-air radiosonde, satellite radiances, measured visibility distributions, and surface $PM_{2.5}$ and PM_{10} data were used in model comparisons to show differences in model performance.

The major effort on verification focused on *e*D/SW’s ability to forecast unhealthy dust episodes. As such, verification of the timing and magnitude of dust events is critical. The *e*D/SW model accurately predicts the timing of dust events throughout the model domain. Changes to input parameters had little or no effect on the timing agreement. The most significant improvements to the magnitude agreement occurred between runs 1a and 2a (the replacement of the OWE data set with MOD12Q1) and from run 2a to 4a (the addition of SRTM terrain data). Differences between the two land cover data sets are due to several possible reasons. The 10-min spatial resolution OWE data set was compiled in the early 1990s, based mainly on data sources of 1970s and ‘80s. The 1km resolution MOD12Q1 data set is based on 2001 satellite data but is not temporally updated to record seasonal changes in land cover resulting from agricultural activities, urbanization, drought, wildfires, and other climate variations.

Different land cover data had a much larger and systematic influence on the modeled dust concentrations. The *e*D/SW near-ground dust concentration distribution simulated satellite-observed dust clouds and TCEQ reduced visibility distributions better than results from DREAM/SW. Comparisons of modeled and observed surface $PM_{2.5}$ time series also showed that *e*D/SW results were better than those from DREAM/SW. This was confirmed by performance statistics for modeled $PM_{2.5}$ concentrations and by comparisons of modeled vs. observed $PM_{2.5}$ fields. Figure 42 shows relative improvements in *e*D/SW model runs for the two test cases. The most important improvements occurred in model runs 2a and 4a (addition of MOD12Q1 and SRTM terrain data respectively.

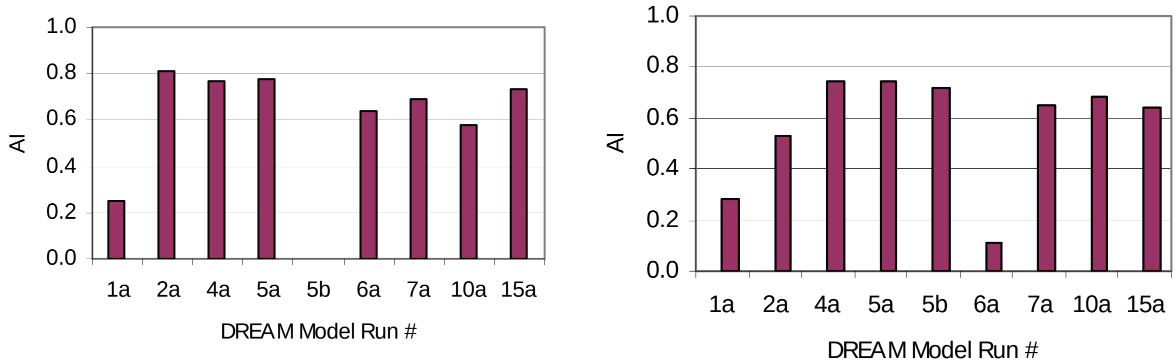


Figure 42. Agreement indices for the December 2003 dust events: (left) Case 1; (right) Case 2.

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 with DREAM-generated PM_{10} and $PM_{2.5}$ concentrations. Furthermore, these thin sampling networks are concentrated in densely-populated urban areas subject to PM_{10} and $PM_{2.5}$ sources generated by human activity, as in construction and combustion. There are too few speciated particle sampling sites 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 dataset/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.

While enhanced DREAM/eta simulated meteorological patterns well, it has mixed performance predicting the extent of dust events. Further improvements might be obtained if better estimates of aerodynamic surface roughness length (z_0) were obtainable through ESR data. Understanding and measuring this parameter is crucial for understanding surface friction and the ability of wind to lift dust from a surface.

The project team is convinced from early results that there is ample room for improving the model with better resolution satellite data for surface parameters.

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.

References

Anonymous. 1997. Pro-ACT & TM. *Journal of Nursing Administration*, 27(2):37.

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

Arakawa, A., V.R. Lamb. (1977), Computational design of the basic dynamical processes of the UCLA general circulation model. *Methods in Computational Physics.* Academic Press, 17: 173-265.

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

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

Binder, S., A.M. Levitt, J.J. Sacks, J.M. Hughes. 1999. Emerging Infectious Diseases: Public Health Issues for the 21st Century. *Science* 284: 311-313.

Bravata, D.M., K.M. McDonald W.M. Smith, C. Rydzak, H. Szeto, D.L. Buckeridge, C. Haberland, and D.K. Owens. 2004. Systematic Review: Surveillance Systems for Early Detection of Bioterrorism-Related Diseases. Ann Intern Med. 140:910-922.

CDC (Centers for Disease Control and Prevention). 1998. Surveillance for Asthma—United States, 1960-1995. MMWR. 47(SS-1).

CDC. 2005. *Synd*romic Surveillance: Reports from a National Conference, 2004. MMWR 2005:54 (Suppl). 208 pages.

Chu, D.A., Y.J. Kaufman, G. Zibordi, J.D. Chern, J. Mao, C. Li, and B.N. Holben. 2003. Global Monitoring of Air Pollution Over Land from the Earth Observing System-Terra Moderate Resolution Imaging Spectroradiometer (MODIS). *Journ. Geophys. Res.*, 108(D21): 4,661.

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.

Cressman, G.P. 1959. An Operational Objective Analysis System, *Monthly Weather Review*, 87: 367-374.

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

Derbyshire, E. 2005. Essentials of Medical Geology: Impacts of the Natural Environment on Public Health. Elsevier, London, pp. 459-480

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.

Fairchild, A.L. and R. Bayer. 2004. Ethics and the Conduct of Public Health Surveillance. *Science* 303 (30 January): 631-632.

Fauci, A.S., N.A. Touchette, G.K. Folkers. 2005. Emerging Infectious Diseases: a 10-Year Perspective from the National Institute of Allergy and Infectious Diseases. *Emerging Infectious Diseases* 11(4): 519-525.

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.

Georgi, F. 1986. A particle dry-deposition parameterization scheme for use in tracer transport models. *Journal of Geophysical Research*, 91: 9794-9806.

Goudie, A.S. and N.J. Middleton. 2001. Saharan Dust Storms: Nature and Consequences. *Earth Sci. Rev.*, 56: 179-204.

Grousset F.E., P. Ginoux, A. Bory, and P.E. Biscaye. 2003. Case Study of a Chinese Dust Plume Reaching the French Alps. *Geophys. Res. Letters,* 30(6): 10.1029/2002 GL016833.

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.

GSFC. 2000. EOS Data Products Handbook. Volume 1: TRMM, Terra, Data Assimilation System. 258 pages.

GSFC. 2003. EOS Data Products Handbook. Volume 2. ACRIMSAT, Aqua, Jason-1, Landsat 7, Meteor 3M/SAGE III, Quiksat, QuikTOMS, VCL. 253 pages.

Gu, Y., W.I. Rose, J.S. Bluth. 2003. Retrieval of Mass and Sizes of Particles in Sandstorms Using Two MODIS IR Bands: A Case Study of an April 7, 2001 Sandstorm in China. *Geophys. Res. Letters*, 30(15): 1805:7-1 to 7-4.

Hadler, J.L., A. Siniscalchi, and Z. Dembek. 2005. Hospital Admissions Syndromic Surveillance—Connecticut, October 2001-June 2004. In: *Syndromic Surveillance: Reports from a National Conference, 2004*. MMWR 2005; 54 (Suppl):169-173.

Ichoku, C., L. Remer, and T. Eck, 2005. Quantitative Evaluation and Intercomparison of Morning and Afternoon Moderate Resolution Imaging Spectroradiometer (MODIS) Aerosol Measurements from Terra and Aqua. *Journal of Geophysical Research*, 110: D10S03.

Janjic, Z. I. 1984. Non-linear advection schemes and energy cascade on semi-staggered grids. *Monthly Weather Review*, 118: 1234-1245.

Janjic, Z. I. 1994. The Step-mountain Coordinate Model: Further Developments of the Convection, Viscous Sublayer and Turbulence Closure Schemes. *Monthly Weather Review*, 122: 927-945.

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.

Kaufman, Y. and D. Tanre, D.,1998. Algorithm for Remote Sensing of Troposheric Aerosol from MODIS. MODIS ATDB Product ID: MOD04.

Kaufman, Y.J., A. Karnieli, and A. Tanre. 2000. Detection of Dust over Desert Using Satellite Data in the Solar Wavelengths. *IEEE Trans. Geosci. & Rem. Sens.*, 38(1): 525-531.

Kaupp, V., C. Hutchinson, S. Drake, W. van Leeuwen, D. Tralli. 2004. Assimilation of NASA Earth Science Results and Data in National Decision Support Systems: A Guidebook. Draft technical report. 56 pages.

Kaya, S., J. Sokol, and T.J. Pultz. 2004. Monitoring Environmental Indicators of Vector-borne Disease from Space: A New Opportunity for RADARSAT-2," *Can. Journ.* Rem. Sens., 30(3): 560-565.

King, M.D., Y.J. Kaufman, D. Tanre, and T. Nakajima. 1999. Remote Sensing of Tropospheric Aerosols from Space: Past, Present, Future." *Bull. Am. Met. Soc.*, 80(11): 2229-2259.

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.

Lee, T. 1989. Dust Tracking Using Composite Visible/IR Images: A Case Study. *Weather and Forecasting*, 4: 258-263.

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

Liverman, D., E.F. Moran, R.R. Rindfuss, and P.C. Stern. 1998. *People and Pixels: Linking Remote Sensing an*d Social Sciences. National Academy Press, Washington D.C., pp. 28-51; pp. 197-203.

Lohr, S. 2008. Most Doctors Aren't Using Electronic Health Records. New York Times. June 19, 2008.

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.

Mathur, M.L. and R.C. Choudhary. 1997 Desert Lung in Rural Dwellers of the Thar Desert, India. *J. Arid Environ.*, 35: 559-562.

Mesinger, F., Z.I. Janjic, S. Nickovic, D. Gavrilov and D.G. Deaven (1988), The Step-mountain Coordinate: Model Description and Performance for Cases of Apline Lee Cyclogenesis and for a Case of an Appalachian Redevelopment. *Monthly Weather Review,* 116: 1493-1518.

Miller, S.D. 2003. A Consolidated Technique for Enhancing Desert Dust Storms with MODIS. *Geophys. Res. Letters*, 30(20): 2071.

Moorman, J.E., R.A. Rudd, AC.A. Johnson, M. King, P. Minor, C. Bailey, M.R. Scalia, L.J. Akinbami. 2004. National Surveillance for Asthma—United States, 1980-2004. CDC, Morbidity and Mortality Weekly Review: 56(SS08): 1-14; 18-54

Morain, S. and A. Budge, 2005. Engineering Satellite Data for Environmental Health Issues. *Remote Sensing A*rabia: For the Betterment of People. Riyadh, Saudi Arabia. May 7-11. Proceedings on CD-ROM.

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.

NASA, 2006. Earth Science Division Applied Sciences Program, Public Health Program Element FY2007-2011 Plan (final draft). Science Mission Directorate. 35 pgs.

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 (D16):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.

Njoku, E.G. 1999. AMSR Land Surface Parameters: Algorithm Theoretical Basis Document Version 3.0. Jet propulsion Laboratory, Cal Tech. 47 pages

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.

Oxford J.S., R. Lambkin, A. Sefton, R. Daniels, A. Elliot, R. Brown, and D. Gill. 2005. A Hypothesis: the Conjunction of Soldiers, Gas, Pigs, Ducks, Geese, and Horses in Northern France Provided the Conditions for the Emergence of the 'Spanish' Influenza Pandemic of 1918-1919. *Vaccine,* 23(7): 940-945.

Parsons, L.C. 1997. "Delegation Decision Making," *Journal of Nursing Administration*, 27(2):47.

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.

Prospero, J.M.. 1999. Long-Term Measurements of the Transport of African Mineral Dust to the Southeastern United States: Implications for Regional Air Quality. *J. Geophys.Res.*, 104(15): 917-927.

Remer, L., 2005. *Journal of Atmospheric Sciences*. April issue.

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.

Shao, Y., M.R. Raupach and P.A. Findlater. (1993). "Effect of Saltation Bombardment on the Entrainment of Dust by Wind," *Journal of Geophysical Research.* 98: 12719-12726.

Shaw, P. 2008. Application of Aerosol Speciation Data as an In situ Dust Proxy for Validation of the Dust Regional Atmospheric Model (DREAM). *Atmospheric Environment*. 42(31):7304-7309.

Srivastava, H.S., P. Patel, Y. Sharma, and R.R. Navalgund. 2008. Retrieval of Surface Roughness Using Multi-polarized Envisat-1 ASAR Data. *Geocarto Intern*. 23(1):67-77.

Stefanov, W.L., M.S. Ramsey, and P.R. Christensen. 2003. Identification of Fugitive Dust Generation, Transport, and Deposition Areas Using Remote Sensing. *Environ. & Engin. Geoscience,* 9(2): 151-165.

Taubenberger, J.K., A.H. Reid, and T.G. Fanning. 2005. Capturing a Killer Flu Virus. *Scientific American,* (January): 62-71.

Varmus, H. R. Klausner, E. Zerhouni, T. Acharya, A.S. Daar, and P.A. Singer. 2003. Grand Challenges in Global Health. *Science,* 302 (17 October): 398-399.

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.

Ward, T. 2005. email communication from the City of Lubbock, TX Health Department to the list serve of the Research Association of Medical and Biological Organizations (RAMBO). 1-page.

Development Testbed Center. 2008. Weather Research & Forecasting Version 1 Model Evaluation Tools User's Guide. National Center for Atmospheric Research. Boulder, CO.

Wiggs, G.F.S., S.I. O'Hara, J. Wegerdt, J. van der Meers, I. Small, and R. Hubbard. 2003. The Dynamics and Characteristics of Aeolian Dust in Dryland Central Asia: Possible Impacts on Human Exposure and Respiratory Health in the Aral Sea Basin. *Geo*g. J., 169: 142-157.

Woerden, J. van. 1999. Data Issues of Global Environmental Reporting: Experiences from GEO-2000. UNEP/DEIA & EW/TR, 99.3, 52 pgs.

World Bank Group. 2002. Public Health Surveillance Toolkit, http://survtoolkit.worldbank.org; accessed June 10, 2005.

Wright, K. 2005. Blown Away. *Discover Magazine*, 26(3): 32-37.

Xu, X.Z., X.G. Cai, and X.S. Men. 1993. A Study of Siliceous Pneumoconiosis in a Desert Area of Sunan County, Gansu Province, China. *Biomed. Environ. Sci.,* 6: 217-222.

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.

Zelicoff A., J. Brillman, D.W. Forslund, J.E. George, S. Zink, S. Koenig. 2001. The Rapid Syndrome Validation Project (RSVP). Albuquerque, NM: Sandia National Laboratories.

Zelicoff, A.P. and M. Bellomo. 2005. *Microbe: are We Ready for the Next Plague?* New York: Amacom (a division of American Management Association). 273 pages.

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.

Appendices

Appendix 1: Zelicoff / Forslund emails

Zelicoff – ARES Corporation, Albuquerque

“ESSENCE is a "data-mining" system and, as such generates enormous amounts of data that have no sensitivity to clinically significant events. It has been tested (along with RODS, Redhat and [other] data-mining systems in Texas -- which is why they have chosen SYRIS. [I] don't want to get into a big discussion about this, rather just to point out the distinction with a difference that is data-mining vs. clinician-driven surveillance. A short paper from the Lubbock DOH is available if you'd like to see it. In the end, the market of public health folks will determine what is useful and what simply generates more data-without-insight.

“It should be noted that RSVP (the predecessor to SYRIS) kept the Butler event from turning into a bioterrorism scare. According to Lubbock public health officials, the syndrome surveillance system enabled them to keep the situation under complete control and not a penny of either public health or medical diagnostic funds were [spent] trying to "rule out" plague. This episode was in direct contra-distinction to the anthrax scares in NYC and Washington DC where literally millions of dollars were [spent] trying to "prove a negative," because of the complete absence of any clinical context into which to put the information

“I certainly agree that the underlying "statistic" determines the utility of the [information] system. In most diseases of public health importance, and in ALL of the diseases of bioterrorism importance (especially, but not limited to smallpox), th[e] underlying statistic is N= 1. That is what SYRIS is designed to have high fidelity for. That's also very problematic for data-mining systems that collect information like ER "chief complaints," which are almost always 2 or 3 word descriptors of the chief complaint, for example,"skin rash." How many of those occur in a typical busy ER every week? (answer: many dozens to hundreds). How many are of clinical public health importance (answer: usually zero). So, if one is looking at distributions for underlying statistics (say a standard deviation or two in a time series), data-mining systems are exquisitely designed to fail in terms of timeliness where hours matter. Perhaps this explains part of the reason that people were "scared" despite having ESSENCE or RODS.

“Put another way, there is a practical difference between lots of data (like numbers in a phone book) and the piece of data that is true knowledge (e.g. the phone number you want). SYRIS isn't "perfect" in this regard (nothing ever will be), but it is a big advance over what currently exists.”

Forslund – Los Alamos National Labs

“I'm not keen on most of the medical surveillance systems being deployed including ESSENCE or RODS, but ESSENCE and RODS do get syndromic data in large quantities [that are] useful for [assessing] anthrax outbreaks (as well as others). There are a number of excellent papers in the literature on the sensitivity and selectivity of syndromic surveillance systems. We have found that getting data electronically in large volumes is better than having doctors enter data, which results in statistically unknowable sensitivity to events even though specific information for an event may be more clinically appropriate. This is largely due to the undeterminable sampling statistics of clinician entered syndromic data. It may be able to see a specific clinical event, but there may be no way to know the pervasiveness of such an event. There are ways of getting large amounts of clinically relevant data without adding a single second to the time that a provider is dealing with a patient, i.e., without having to have "double-entry" of data. This would be the ideal. In any case, my only point is that there was a "clinical context" in the DC area for evaluating the presence of Anthrax. But this isn't sufficient to keep people from being "scared" of anthrax given a thought-to-be positive lab test. There was no lack of clinical context, but the response was out of fear of Anthrax based on earlier history. Rationality isn't always part of the process in responding to a potential BT event. I actually don't know if people even looked at the clinical data in this situation. There does seem to be a problem linking Biowatch data with medical surveillance data which DHS is at least beginning to address.

"Perhaps you have other data than I'm aware of, but there has been a "clinical context" in DC for some time from the use of ESSENCE in the National Capitol Region that should rather easily see Anthrax. As I understand [it], the problem [in] DC was an "error" in some data at USAMRIID. A "positive" of Anthrax in a lab test will cause a scare of significant magnitude especially in a heavily populated area even if there is lots of clinical surveillance data showing nothing."

Appendix 2: Lubbock beta test of RSVP

Experience with Syndrome-based Disease Surveillance in Lubbock, Texas: 1999 – Present

Tigi Ward, BSN, MS, City of Lubbock Health Department

e-mail: tward@mylubbock.us; Phone: (806) 775-2941

Tommy Camden, MS, RS, Health Director, City of Lubbock Health Department

e-mail: tcamden@mylubbock.us; Phone: (806) 775-2899

Introduction

Since early 1999, the City of Lubbock Department of Health has evaluated several "syndrome-based" disease surveillance systems (SBDSS). This brief paper is intended as a preliminary summary of our experience focusing on the utility of SBDSS in accomplishing the following primary goals of public health services:

Prevent epidemics and the spread of disease

Protect against environmental hazards

Prevent injuries

Promote and encourage healthy behaviors and mental health

Respond to disasters and assist communities in recovery

Assure the quality and accessibility of health services

In this summary, we focus on infectious diseases (both communicable and noncommunicable) of public health importance. In theory, SBDSS by virtue of their timeliness and volume of information flows could assist in meeting these central public health responsibilities. In practice however, the specific designs, and underlying technical features and scientific approach and ease-of-use is dramatically different across the dozens of SBDSS currently in existence, some of which have been implemented only in narrowly defined demographic settings or which have other limiting features. The promise is often not met in real-world use.

All SBDSS fall into two basic categories (Brevata et al., 2004):

"Automated" or "passive" surveillance systems that seek to exploit existing data streams and employ various statistical algorithms to detect the presence of infectious disease. Some of the data sources that are "tapped" by these passive systems include: pharmacy sales (including over-the-counter medications), total volume of nurse "hot-line" calls, brief "chief complaint" summaries from emergency room logs, and school and work absenteeism; and, "Active" or "clinical" surveillance system that depend on selected reporting from physicians, veterinarians, EMS services and other healthcare providers based on the clinical judgment when assessing severity of illness among patients (whether animal or human).

It is also important to note that the overwhelming majority of SBDSS data gathering features focus solely on human patients, despite the fact that in all significant outbreaks of novel diseases over the past decade or more in North America, animals were the primary source of the diseases. In particular, the following very large or economically significant disease outbreaks among humans had animal sources:

Hantavirus Pulmonary Syndrome in the Four Corners Area (1991)

West Nile Fever (1999, 2000) Human plague in New York City in visitors from New Mexico (2001)

Cryptosporidiosis in Milwaukee (1996) in which 400,000 people were sickened

Monkey pox in the midwest (2003)

SARS (2003)

H5-N1 Avian influenza in humans (1997, 1999, 2005)

Tularemia transmission from prairie dog-to human in Texas

We would further emphasize that all of the CDC's Class A and Class B bioterrorism diseases (with the sole exception of smallpox) are animal diseases (sometimes also called zoonotic diseases). Thus, it is highly likely that if there ever is a large-scale bioterrorism event, animals will almost certainly become ill in large numbers and probably with classical syndromes recognized easily by the veterinary community.

Past Experience with SBDSS in Lubbock

Because public health offices are charged with wide-ranging responsibilities yet are relatively under-funded, the City of Lubbock Health Department began to explore means of leveraging limited resources by utilizing electronic SBDSS in1999. Although advertised as easy-to-implement and low-cost, we found that all of the "automated" SBDSS systems were problematic in at least four areas:

The vast majority of cases reported from hospitals and ER-s (based on chief complaints, billing codes or simple census information) resulted in a very large amount of "noise" (data that [were] of little utility) and which created a serious liability because of the possible need to respond to "spikes" that were merely manifestations of statistical randomness.

Pharmacy-sales data were inherently delayed or complicated by items being "on sale" at large pharmacy chains.

Information is almost always reported in tabular or textual format without mapping (geographic information system) tools for analysis.

In all cases, since the historical background was largely unknown for any of the data streams, comparisons to identify "true positive" deviations from normal was impossible.

At the same time as we were reviewing the automated disease-surveillance systems that were proliferating across the US, we identified one "clinician-based" or "active" SBDSS called the "Rapid Syndrome Validation Program" (RSVP™) developed by Alan Zelicoff, MD (then at Sandia National Laboratories). RSVP defined six common syndromes worded in the daily parlance of medicine and public health, and further provided an electronic interface that operated on virtually any computer connected to the Internet. It also provided primitive, but useful geographic mapping tools. Key to the RSVP design philosophy was the central notion of "clinical judgment" in which participating physicians (some 10% of all of the practicing physicians in Lubbock) were asked to report those individuals seen in emergency rooms, clinics, and private offices where the patient was assessed as seriously ill (an assessment that clinicians make routinely) and who fit into one of six syndromes strongly suggestive of infectious disease of public health importance:

Fever with influenza-like illness

Fever with skin rash

Fever with mental status change or neurological change

Severe diarrhea

Hepatitis (presumed to be non-alcohol and non-drug related)

Adult Respiratory Distress Syndrome

Only 15 – 30 seconds of physician time is required for reporting a case, and all new reports are immediately reflected on maps of the local public health jurisdiction along with the ability to analyze data using GIS tools. RSVP also allowed Lubbock public health officials to send out alerts on the "front page" of RSVP instantaneously to physicians.

Our experience with RSVP was uniformly positive. Physician compliance was high (contrary to the popular, but incorrect belief that physicians will not take time to enter cases) because the number of cases of seriously ill patients who fit into one of the syndrome categories was, on average, a case per month per physician (except during large epidemics). Further, RSVP provided information of immediate clinical importance to physicians thus increasing their cost-effectiveness in practice. Finally, on rare occasions, RSVP enabled public health officials to contact doctors within minutes of a case report when the data suggested unusually worrisome symptoms that might require immediate contact investigation. Thus, RSVP cut down the time from initiation of contact investigation from days to mere minutes.

Our criticisms of RSVP were as follows: (1) Because it was a 'web-browser' based system, some particular operating systems or web-browsers would not fully accommodate the RSVP code and some of its features were inaccessible for certain users; (2) Mapping functionality, while useful, was slow and cumbersome; (3) There was no ability to report key veterinary syndromes (see above) that would often presage human disease; (4) Statistical analysis via RSVP was somewhat difficult because of the nature of the database where all information was stored; and (5) It was unclear to us that RSVP was NEDSS compliant.

Despite these criticisms, we had two very important public health successes with RSVP. We were able to manage the threat of a plague bioterrorism event in January of 2002 when it appeared that strains of the organism were stolen from the Texas Tech University Health Sciences Center by monitoring respiratory disease cases on literally a minute-by-minute basis and providing diagnostic information via RSVP to clinicians. Panic was completely avoided, and there was no unnecessary diagnostic testing to waste public health resources. We predicted via RSVP that we were dealing with a false alarm and that there were no public health concerns – exactly as turned out to be the case.

Our second success was in early 2003 when we discovered, based on clinical symptoms, the need for earlier-than-usual testing for influenza. This resulted in finding influenza cases in our community approximately three weeks earlier than would otherwise have been possible, probably mitigating much morbidity in the population.

Current Experience

RSVP™ was a useful and highly successful "alpha" product, and the Lubbock City Health Department completed its beta testing of this product. We are currently employing a SBDSS from ARES Corporation in Albuquerque called SYRIS™ - The Syndrome Reporting Information System. In distinction to RSVP and all of the passive SBDSS in the marketplace, SYRIS addresses all of our critiques of past systems and offers the following:

It is completely platform-independent and does not require a web-browser. Thus, it will run on virtually any Internet-connected device including many handheld devices. SYRIS is comprehensive including all critical "health care providers"

Physicians, physician-assistants, nurse practitioners and nurse clinicians

School nurses (who report absenteeism and commentary)

EMS professionals (reporting transport-cases by syndrome)

Veterinarians (who have 9 separate syndromes covering all major domestic, agricultural and exotic animal species)

Coroner/Office of the Medical Investigator (who also have a list of syndromes based solely on information from unexpected death reports)

Laboratory technicians (who can report all lab tests for infectious agents in less than 1 minute per week)

Animal control and environmental health officials (who report on captured stray animals or wildlife and the number requiring euthanasia due to severe illness)

Wild-life rehabilitators

Enhanced mapping features based on the "open source" Minnesota Mapping Server that provides for near instantaneous map updating and query to any region where SYRIS is in use.

Full NEDSS compliance

Extremely rapid data entry: less than 15 seconds for physicians and veterinarians

Automated and manual alarm features so that public health officials can be notified by digital paging and e-mail when cases that meet specifically defined criteria (at the discretion of local public health officials) are met.

Easy statistical analysis of all current and historical SYRIS data

Easy training: SYRIS is intuitive to use and a full manual is available on¬line tailored to each of the 8 user communities defined above

Low cost: approximately 7 – 8 cents per capita. So, in our catchment area of 250,000 people, SYRIS will cost less than $18,000. This licensing fee includes 24/7 support, all database maintenance and storage and automatic updates to the software each time a user starts SYRIS

We believe that SYRIS will solve the vast majority of our disease surveillance and response needs (including emergency response in the case of bioterrorism) with a low false alarm rate and high sensitivity.

Summary

Our experience with properly designed active, clinician-driven SBDSS demonstrates that physicians and other busy health professionals will report cases of suspected infectious disease if the system is fast (less than 15 – 30seconds), provides immediate feedback to clinicians on local infectious disease outbreaks, permits selective interaction between public health officials and clinicians on a real-time basis as warranted, and which is inexpensive. SYRIS meets all of these criteria. In addition, unlike the "passive" or "data-mining" approaches, SYRIS has a low false-positive rate (thus mitigating the investigation of a large number of false alarms and squandering limited public health resources) while at the same time facilitating enhanced relationships between local public health officials and all health care providers.

SYRIS makes public health part of daily human and veterinary medical practice and medicine part of daily public health operations.

Appendix 3: Publications

Benedict, K.K. and W. Hudspeth. 2005. Technology Products of the PHAiRS REASON Project – Year 1. In: Proceedings of the 2005 Sun-Earth System Technology Conference. College Park, MD. CD-ROM.

Budge, A.M., K.K. Benedict, and W. Hudspeth. 2006. Developing Web-Based Mapping Services for Public Health Applications. Proceedings, ISPRS Commission IV Symposium. 27-30 September, Goa, India. 1060 pgs. GITC, Lemmer, The Netherlands. ISS 1682-1750.

Hudspeth, W., S. Nickovic, D.Yin, B. Chandy, B. Barbaris, A.M. Budge, T.K. Budge, S. Baros, K.K. Benedict, C. Bales, C. Catrall, S.A. Morain, G.M. Sanchez, W.A. Sprigg, and K. Thome. 2005. PHAiRS – A Public Health Decision Support System: Initial Results. In: Proceedings 31st International Symposium on Remote Sensing of Environment: Global Monitoring for Sustainability and Security. 20-24 June, St. Petersburg, Russia. CD-ROM.

Hudspeth, W., K.K. Benedict, and A. Budge. 2007. Architecture and Functionality of the PHAiRS (Public Health Applications in Remote Sensing) Web Services. In: Proceedings 32nd International Symposium of Remote Sensing of Environment: Sustainable Development Through Global Earth Observations. 25-29 June, San Jose, Costa Rica.

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: Chu, A., Szykman, J. and S. Kondragunta (eds) Remote Sensing of Aerosol and Chemical Gases, Model Simulation/Assimilation, and Applications to Air Quality. Proc. SPIE Vol. 6299, 62990L. DOI: 10.1117/12.679868.

Morain, S.A., and A.M. Budge. 2004. Remote Surveillance Technologies for Assessing Biological Threats. In: BTR 2004: Unified Science and Technology for Reducing Biological Threats and Countering Terrorism – Proceedings. 18-19 March. Albuquerque, NM.

Morain, S.A., A.M. Budge, T.K. Budge, S. Baros, K.K. Benedict, W. Hudspeth, C. Bales, G.M. Sanchez, W.A. Sprigg, D. Yin, B. Barbaris, B. Chandy, S. Nickovic, S. Caskey, J. Speer, and J. Bradbury. 2005. Modeling Atmospheric Dust for a Public Health Decision Support System. In Proceedings 31st International Symposium on Remote Sensing of Environment: Global Monitoring for Sustainability and Security. 20-24 June, St. Petersburg, Russia. CD-ROM.

Morain, S.A., and A.M. Budge. 2005. Engineering Satellite Data for Environmental Health Issues. In: Proceedings of ISPRS WGI/4 International Conference on Advanced Remote Sensing for Earth Observation: Systems, Techniques, and Applications. Riyadh, Saudi Arabia.

Morain, S.A., and A.M. Budge. 2006. Integrating Earth Observations Data into Geospatial Databases that Support Public Health Decisions. Proceedings, ISPRS Commission IV Symposium. 27-30 September, Goa, India. 1060 pgs. GITC, Lemmer, The Netherlands. ISS 1682-1750.

Morain, S.A. and A.M. Budge. 2006. Science Data Products for Public Health Decision Support. In: Proceedings of Geoscience and Remote Sensing Symposium, IEEE International Conference. 31 July – 04 August. Pages 421-424. DOI: 10.1109/IGARSS.2006.112.

Morain, S.A. and A.M. Budge. 2008. Extending Environmental Surveillance to Useful Public Health Information. In: Proceedings of ASPRS 2008 Annual Conference – Bridging the Horizons: New Frontiers in Geospatial Collaboration. 28 April-2 May, Portland, OR. CD-ROM.

Morain, S.A. and A.M. Budge. 2008. Environmental Sensing and Human Health. Chapter 29 in: Zhilin Li, Jun Chen, and Emmanuel Baltsavias (eds.). *Advances in Photogrammetry, Remote Sensing and Spatial Information Sciences*. London: Taylor and Francis Group, CRC Press. A Balkema Book.

Morain, S.A., and A.M. Budge. 2008. Verification and Validation of Desert Dust Forecasts and Their Impact on Respiratory Health Applications in the Southwestern United States. In Proceedings: International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII, Part B8. ISSN 1682-1750. Beijing 2008.

Morain, S.A. In press. Improving Public Health Services Through Space Technology and Spatial Information Systems. In: Proceedings of Erice International Seminars on Planetary Emergencies, 40th Session. 19-24 Aug 2008. Erice, Sicily, Italy.

Shaw, P. 2008. Application of Aerosol Speciation Data as an In-situ Proxy for Validation of the Dust Regional Atmospheric Model (DREAM). Atmospheric Environment, 42(31):7303-7309.

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, DOI:10.1016/j.atmosenv.2005.07.009.

Yin, D., S. Nickovic, and W.A. Sprigg. 2006. The Impact of Using Different Land Cover Data on Wind-blown Desert Dust Modeling Results in the Southwestern United States. Atmospheric Environment, DOI:10.1016/ j.atmosenv.2006.10.061.

Appendix 4: Oral presentations and posters

Benedict, K.K., and W. Hudspeth. 2005. Technology Products of the PHAiRS REASON Project - Year 1. 2005 Sun-Earth System Technology Conference. College Park, MD.

Benedict, K.K. Menke, W. Hudspeth, and J. Cavner. 2005. EDAC's Web-based Geospatial Applications and the Open Source Technologies Behind Them. Fall Meeting of the New Mexico Geographic Information Council, Albuquerque, NM

Benedict, K. 2006. Workshop. Implementation of OGC Web Services with MapServer. Summer 2006 Earth Science Information Partnership (ESIP) Federation Meeting. Palisades, New York.

Benedict, K.K. and W. Hudspeth. 2006. Technology Products of the PHAiRS REASoN Project – Year 2 Web Services and Demonstration Interfaces Development. ESTC 2006 Conference. College Park, MD.

Benedict, K. 2007. "Development of a 'Slippy-map' Interface for OGC WMS Services" Demonstration at the Winter ESIP Federation Meeting, Portland, OR.

Benedict, K. 2007. From Routine to Extreme: Flexible Services Oriented Architectures for Rapid Public Health Information Delivery. Research Association of Medical and Biological Organizations. Sevilleta Long Term Ecological Research Facility, NM.

Benedict, K. 2008. Poster: The ESIP Federation's Long Experience in the Development and Deployment of Web Services. NOAA Climate Prediction Application Science Workshop, Chapel Hill, NC. March 4-7, 2008.

Budge, A.M., K.K. Benedict, and W. Hudspeth. 2006. Developing Web-Based Mapping Services for Public Health Applications. ISPRS Commission IV Symposium. Goa, India.

Budge, A.M. 2006. Poster: Using Satellite Data to Improve Dust Storm Forecasts. New Mexico State Legislature. Santa Fe, NM.

Budge, A.M. and K.K. Benedict. 2007. Poster: Public Health Applications in Remote Sensing. 32nd International Symposium on Remote Sensing of Environment. San Jose, Costa Rica.

Budge, A.M. 2008. Poster. Human Health Societal Benefit Area. New Mexico State Legislature. Santa Fe, NM.

Hudspeth, W., S. Nickovic, D. Yin, B. Chandy, B. Barbaris, A. Budge, T.K. Budge, S. Baros, K.K. Benedict, C. Bales, C. Catrall, S.A. Morain, G. Sanchez, W.A. Sprigg, and K. Thome. 2005. PHAiRS – A Public Health Decision Support System: Initial Results. 31st International Symposium for Remote Sensing of the Environment. St. Petersburg, Russia.

Hudspeth, W., K.K. Benedict, and A.M. Budge. 2007. Architecture and Functionality of the PHAiRS (Public Health Applications in Remote Sensing) Web Services. 32nd International Symposium of Remote Sensing of Environment. San Jose, Costa Rica.

Mahler, A-B., K. Thome, D. Yin, W.A. Sprigg. 2006. Dust Transport Model Validation Using Satellite- and Ground-based Methods in the Southwestern United States. The International Society for Optical Engineering (SPIE) Annual Meeting. San Diego, CA.

Morain, S.A., and A.M. Budge. 2004. Remote Surveillance Technologies for Assessing Biological Threats. BTR 2004: Unified Science and Technology for Reducing Biological Threats & Countering Terrorism "Homeland Security: Toward Converging Partnerships." Albuquerque, NM.

Morain, S.A., and A.M. Budge. 2004. NASA/EDAC Public Health Initiative: Satellite Technology for Public Health Information and Decision Support. Research Association of Medical and Biological Organizations. Sevilleta Long Term Ecological Research Facility, NM.

Morain, S.A.. 2004. Remote Sensing and Public Health. Mississippi Gulf Coast Geospatial Conference. Biloxi, MS.

Morain, S.A. 2005. Integrating Satellite Image Data with Respiratory Reporting Systems. NM Quarterly Epidemiology Meeting. Albuquerque, NM.

Morain, S.A., and A.M. Budge. 2005. Engineering Satellite Data for Environmental Health Issues. Remote Sensing Arabia. Riyadh, Saudi Arabia.

Morain, S.A. 2005. Dimensions of Remote Sensing for Environmental and Public Health. Plenary Session. 31st International Symposium for Remote Sensing of the Environment. St. Petersburg, Russia.

Morain, S.A. 2005. Modelling Atmospheric Dust for a Public Health Decision Support System. 31st International Symposium for Remote Sensing of the Environment. St. Petersburg, Russia.

Morain, S.A. 2005. Climate Modeling Using Earth Observation Data to Improve Public Health Decisions. Climate Change Science Program Workshop: Climate Science in Support of Decision Making. Arlington, VA.

Morain, S.A. 2006. Public Health Applications in Remote Sensing (PHAiRS). New Mexico Geographic Information Council Annual Spring Meeting. Albuquerque, NM.

Morain, S.A. and A.M. Budge. 2006. Science Data Products for Public Health Decision Support. 2006 IEEE International Geoscience and Remote Sensing Symposium & 27th Canadian Symposium on Remote Sensing. Denver, CO.

Morain, S.A. and A.M. Budge. 2006. Integrating Earth Observations Data into Geospatial Databases that Support Public Health Decisions. ISPRS Commission IV Symposium. Goa, India.

Morain, S.A. 2006. Adding GIS and Earth Observations to Syndromic Surveillance. International Society for Disease Surveillance Annual Meeting. Baltimore, MD.

Morain, S.A., W.A. Sprigg, and A.M. Budge. 2007. NASA Public Health/Air Quality Workshop. Applications of Environmental Remote Sensing to Air Quality and Public Health Workshop. Potomac, MD.

Morain, S.A. and A.M. Budge. 2007. PHAiRS Project Overview. Stakeholder Training Workshop. Albuquerque, NM.

Morain, S.A. 2007. Respiratory Health Applications Using New Satellite Air Quality Sensors. SCANEX Annual Meeting. Moscow, Russia.

Morain, S.A. and A.M. Budge. 2008. Overview of PHAiRS. Presented to University of Mississippi Medical Center. Jackson, MS.

Morain, S.A. 2008. Environmental Sensing: An Evolving Program for Air Quality and Human Health. Earth Observation for Hazard and Risk Research. Durham University. Durham, England.

Morain, S.A. and A.M. Budge. 2008. Environmental Surveillance for Public Health. ASPRS Annual Conference. Portland, OR.

Morain, S.A. 2008. Environmental Sensing: Air Quality and Human Health. Xiangshan Conference. Beijing, China.

Morain, S.A. and A.M. Budge. 2008. Verifying and Validating Dust Forecasts for Health Risks. XXI ISPRS Congress. Beijing, China.

Morain, S.A. 2008. Improving Public Health Services Through Space Technology and Spatial Information Systems. Erice International Seminars on Planetary Emergencies, 40th Session. Erice, Sicily, Italy.

Shaw, P. 2007. Applications of Soil Component Proxy to Dust Model Validation. UA Graduate Seminar, Tucson, AZ.

Sprigg, W.A. 2006. Public Health Applications for Remote Sensing and Atmospheric Modeling. Presented to the Institute of Atmospheric Physics, Chinese Academy of Sciences and to the CAS Institutes of Hydrology & Water Resources and Geographic Sciences & Natural Resources. Beijing, China.

Sprigg, W.A. 2006. Public Health Applications for Remote Sensing and Atmospheric Modeling. Presented to the Beijing Normal University and Tsinghua University. Beijing, China.

Sprigg, W.A. 2006. High Resolution Modeling of Weather, Dust and Surface Hydrology. Presented to the Haihe River Conservancy Commission, Ministry of Water Resources. Tianjin, China.

Sprigg, W.A. and D. Yin. 2007. Airborne Dust Simulations and Forecasts. Air Quality Trends in the Southwest Forum. Westward Look Resort, Tucson, AZ.

Sprigg, W.A. and D. Yin. 2007. New Developments on Particulate Matter Forecasting for Public Health Applications. 32nd International Symposium on Remote Sensing of Environment. San Jose, Costa Rica.

Sprigg, W.A. 2008. Public Health Applications in Remote Sensing. China Meteorological Administration Seminar. Beijing, China.

Yin, D. 2006. A Wind-blown Dust Simulation and Forecast System for Public Health Services. NCEP/EMC Seminar Series. Camp Springs, MD.

Yin, D., W.A. Sprigg, B. Barbaris, and P. Shaw. 2006. Poster. Dust Modeling and its Applications to the Border Region. Building Environmental Security Through Binational Cooperation, SCERP. Tucson, AZ.

Appendix 5: Terminology

Animation: Process of giving the illusion of movement to drawings, models, or inanimate objects. Computer animation is a form of animated graphics that has replaced “stop motion” of scale-model puppets or drawings.

Chronic Obstructive Pulmonary Disease (aka CO[Lung]D): A disease of the lungs in which the airways become narrowed, leading to limited air flow and shortness of breath. In contrast to asthma, the limitation to air flow is usually irreversible and gets worse over time. Though usually caused by noxious particles or gases from smoking COPD can be exacerbated by chronic exposures to atmospheric contaminants.

Data Mining: Type of database analysis that attempts to discover useful patterns or relationships in a group of data. Analysis of data in a database using tools which look for trends or anomalies without knowledge of the meaning of the data. Data mining was invented by IBM who hold some related patents. Data processing using sophisticated data search capabilities and statistical algorithms to discover patterns and correlations in large preexisting databases; a way to discover new meaning in data (from WordNet ® 2.0, © 2003 Princeton University).

Decision Support System: An information system that accumulates input from a variety of sources such that authorities are able to make informed decisions about evolving situations. Such systems may include numerous subsystems for specific kinds of required information, and that prescribe the flow of data into models and the flow of outputs from those models into higher levels of abstraction for making decisions (see Kaupp et al., 2004 for a more complete tutorial).

Earth Science Results: In this context refers to NASA satellite and aerial measurements of environmental parameters, specifically those that become inputs to models that simulate Earth system processes.

Geopotential Height (hPa): is the vertical coordinate referenced to Earth's mean sea level--an adjustment to geometric height (elevation above mean sea level) using the variation of gravity with latitude and elevation. It is considered to be a "gravity-adjusted height." Geopotential height of a certain pressure level would correspond to the geopotential height necessary to reach the given pressure.

Grand Challenge: A call for a specific scientific or technological innovation that would remove a critical barrier to solving an important health problem in the developing world with a high likelihood of global impact and feasibility (Varmus, H. et al., 2003).

Model: There are many types of models (scaled physical models, conceptual models, mathematical models, and others). PHAiRS is using a mathematical model that provides forecasting capabilities of atmospheric dust episodes in the Southwest. Outputs from this model are used as inputs to a conceptual model for facilitating health reporting and consequent public health alerts electronically.

Surveillance: The ongoing, systematic collection, analysis, interpretation, and dissemination of health data (Binder et al., 1999); also, for ethical and Institutional Review Board (IRB) purposes, “Public health surveillance is essentially descriptive in nature. It describes the occurrence of injury or disease and its determinants in the population. It also leads to public health action…, if we confuse surveillance with research, we may be motivated to collect large amounts of detailed data on each case. The burden of this approach is too great for the resources available…” quoted by Fairchild A.L. and R. Bayer, 2004 from World Bank Group “Public health surveillance toolkit (2002).

Syndrome/Syndromic: A number of symptoms occurring together that characterize a specific disease or a group of diseases, each having separate causes and health outcomes.

Visualization: Process of graphically displaying real or simulated scientific data (Concise Encyclopedia, Encyclopedia Britannica Online).

Appendix 6: Acronyms

AE – Angström Exponent

ACRIMSAT – Active Cavity Radiometer Irradiance Monitor Satellite

AIRS – Atmospheric Infrared Sounder

AMSR-E – Advanced Microwave Scanning Radiometer for EOS

AMSU – Advanced Microwave Sounding Unit

AQI.--.Air Quality Index

AQS – Air Quality System AOD – Aerosol Optical Depth (aka AOT – Aerosol Optical Thickness

ASTER – Advanced Spaceborne Thermal Emission and Reflection Radiometer

BGC – BioGeochemical Cycles

BLM – Bureau of Land Management

CAMS – Continuous Air Monitoring Stations

CDC – Centers for Disease Control and Prevention

CERES – Clouds and the Earth's Radiant Energy System

CGI – Common Gateway Interface

COPD -- Chronic Obstructive Pulmonary Disease (aka Chronic Obstructive Lung Disease

CSV – Comma-separated-value

DAAC – Distributed Active Archive Center

DOH – Department of Health

DREAM – Dust Regional Atmospheric Model

DREAM/MED -- DREAM/Mediterranean

DREAM/SW -- DREAM/Southwest

DSS – Decision Support System

EASE-Grid – Equal-Area Scalable Earth Grid

ECMWF – European Center for Medium-Range Weather Forecast

eD/SW -- Enhanced DREAM/Southwest

EID – Emerging Infectious Disease

EO – Earth Observation

EOS – Earth Observation System

EPA – Environmental Protection Agency

ER – Emergency Room

ESMF – Earth System Modeling Framework

ESR – Earth Science Results

ESSENCE – Electronic Surveillance System for the Early Notification of Community-Based Epidemics

EVI – Enhanced Vegetation Index

FAO – Food and Agriculture Organization

FPAR – Fraction of Photosynthetically Active Radiation

G8.-- Group of 8 (international economic summit)

GAP.--.Gap Analysis Program

GEO – Group on Earth Observations

GEOSS -- Global Earth Observing System of Systems

GIS – Geographic Information System

GLAS – Geoscience Laser Altimeter System

GRASS – Geographic Resources Analysis Support System

GSFC – Goddard Space Flight Center

HDF – Hierarchical Data Format

HSB – Humidity Sounder for Brazil

HTTP – Hypertext Transfer Protocol

ICESAT – Ice, Cloud, and land Elevation Satellite

IEEE -- Electronic and Electrical Engineers

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

LSM – Land Surface Model

MEDSIS – Medical Electronic Disease Surveillance and Intelligence System

METAR – Meteorological Aerodrome Report

MI – Myocardial Infarction

MISR – Multi-angle Imaging SpectroRadiometer

MODIS – Moderate resolution Imaging Spectroradiometer

MOPITT – Measurements of Pollution in the Troposphere

MRLC – Multi-Resolution Land Characteristics Consortium

NAAQS – National Ambient Air Quality Standards

NASA – National Aeronautics and Space Administration

NCAR – National Center for Atmospheric Research

NCEP – National Centers for Environmental Prediction

NCRS – Natural Resources Conservation Service

NEDSS – National Electronic Disease Surveillance System

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

OGC – Open Geospatial Consortium

OPeNDAP – Open-Source Project for a Network Data Access Protocol

OWE – Olson World Ecosystems

PDEQ – Pima County Department of Environmental Quality

PHAiRS – Public Health Applications in Remote Sensing

PHO(s) -- Public Health Officials

POI – Plan of Implementation

PR – Precipitation Radar

QA/QC – Quality Assurance / Quality Control

REASoN – Research, Education, and Applications Solution Network

REGAP – Regional Gap Analysis Project

RMSE – Root Mean Square Error

RODS – Real-Time Outbreak and Disease Surveillance

RSVP – Rapid Syndrome Validation Project

SARS – Severe Acute Respiratory Syndrome

SBA -- Societal Benefit Area

SBDSS – Syndrome-Based Disease Surveillance System

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

TEOM – Tapered-Element Oscillation Microbalance

TM – Thematic Mapper

TMI – TRMM Microwave Imager

TRMM – Tropical Rainfall Measuring Mission

UIC -- User Interface Committee (a committee of GEO)

UMD – University of Maryland

UN – United Nations

UNESCO – United Nations Education, Scientific and Cultural Organization

URL – Uniform Resource Locator

USAMRIID – United States Army Medical Research Institute of Infectious Diseases

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

VIRS – Visible Infrared Scanner

WCS – Web Coverage Services

WFS – Web Feature Services

WHO – World Health Organization

WMS – Web Mapping Services

WSSD – World Summit on Sustainable Development

ZIP – Zero-inflated Poisson regression