

# Science Data Products for Public Health Decision Support

S.A. Morain

Earth Data Analysis Center  
University of New Mexico  
Albuquerque, NM, USA  
[smorain@edac.unm.edu](mailto:smorain@edac.unm.edu)

A.M Budge

Earth Data Analysis Center  
University of New Mexico  
Albuquerque, NM, USA  
[abudge@edac.unm.edu](mailto:abudge@edac.unm.edu)

**Abstract**— The *Public Health Applications in Remote Sensing (PHAiRS)* project is engineering an enhanced syndromic surveillance system for dust-related respiratory diseases in the southwestern United States based on assimilating Earth observation (EO) data from NASA experimental satellites. There is a rich literature describing the roles and benefits of using EO data in public health, but most of the documentation is based on anecdotal inferences derived from traditional image interpretation. For several reasons, public health communities cannot rely on evidence of this type because: (1) they need science results that verify, validate, and benchmark the statistical and economic benefits from these exotic inputs; and, (2) they lack the systems that can deliver such reliable information economically and swiftly.

In PHAiRS, several data sets are being assimilated as replacement parameters in the Dust Regional Atmospheric Model (DREAM) to improve simulations of particulate matter entrainment, timing of entrainment, concentrations, and subsequent movement as governed by hourly weather variables available in a regional version of the National Centers for Environmental Prediction (NCEP/Eta) model. On-going simulations from DREAM measure hourly, daily and weekly model improvements from individual EO data replacements that are refreshed on a weekly, seasonal, or inter-annual basis. The overall aims are to: (a) combine the measured improvements from several EO data series that optimize dust forecast scenarios for public health authorities; (b) benchmark each step in the process to document the benefits of EO data inputs into respiratory health care; and (c) develop retrospective and forecast statistics from model runs that boost system reliability and user confidence. Ultimately, the goal is to develop a reliable respiratory public health syndromic surveillance system that can be translated into routine uses of EO data from future NPOESS sensors.

**Keywords**—*modeling, dust, data assimilation, health, decision support*

## I. INTRODUCTION

One of the challenges for Earth system science is to model complex natural and physical processes that couple biogeochemical phenomena and that integrate human health responses into solutions that benefit populations at risk.

---

Sponsored by the National Aeronautics and Space Administration (NASA) under Cooperative Agreement NNS04AA19A

Coupling biogeochemical and dynamical processes that lift dust into the atmosphere with the ecology of airborne pathogens should allow epidemiologists to better understand the medical consequences of dust transport across regions and continents. The role dust plays in human health is an important part of Earth system science that has fundamental socioeconomic and political importance.

## II. PHAiRS

The *Public Health Applications in Remote Sensing (PHAiRS)* project has three parallel thrusts. The first focuses on assimilating satellite observations from MODIS Terra and other sources into the Dust Regional Atmospheric Model (DREAM). This model, in turn, is driven by the National Centers for Environmental Prediction's regional weather forecasting model (NCEP/Eta). The aim of this effort is to: (a) verify that advanced satellite image data from current research sensors can replace traditional model parameters from non-satellite sources, or from earlier (coarser resolution) satellite sources; and, (b) validate that parameter replacements lead to more reliable model forecasts of dust episodes.

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

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

Ultimately, the goal of PHAiRS is to improve public health decision support systems that can evolve toward operational

status for the next generation of space-based sensing. The National Polar-orbiting Environmental Satellite System (NPOESS) is scheduled for launch around 2010. It will consist of several platforms carrying operational versions of NASA's current experimental sensors. It is time now to build the scientific and technological underpinnings of these near-future capabilities, and to test them with appropriate public health user communities.

### III. SYNDROMIC SURVEILLANCE

Respiratory diseases and syndromes are widely recognized as important indicators of population health. To study public health, one needs medical training and an appreciation of those processes that impact not only environments but that impose further impacts on populations at risk. Using satellite-acquired data and imagery to study environmental health has many immediate attractions; however, the extension of these studies for better understanding public health patterns and outcomes presently does not include medical communities.

The public health system in the United States consists 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. This is a powerful motivation for health departments to adopt better, faster, and cheaper ways of making decisions. Almost all public health decisions are made at the local level by departments that are strapped for resources. Nevertheless, they deliver essential public health services through surveillance, health education, and prevention. Furthermore, electronic syndromic reporting systems integrated with geospatial analytical tools must serve many purposes because of the widespread nature of some health conditions, like asthma and (potentially) H5N1 bird flu, or the quite restricted localities for others, like plague. The challenge is to develop syndromic reporting systems that start with a few well-known syndromes and that can be expanded to include not only zoonotic and chronic respiratory diseases, but other health conditions as well.

PHAiRS is focusing on the coarse fraction of airborne thoracic particles ranging in size from  $PM_{2.5}$ - $PM_{10}$ . These particles can carry lethal concentrations of finer biological material having sizes  $\leq 0.01\mu m$ . The framework for coupling atmospheric dust processes with human health responses begins with experimental Earth science satellite products and modifies them for assimilation into DREAM. The output from DREAM becomes input to syndromic reporting systems. These systems are queried by doctors and clinicians who desire additional corroborating information about similar cases being reported by their local or regional colleagues. The ultimate goal is that the output from SYRIS will be delivered to public health decision makers to decide appropriate health alerts.

### IV. BASE-LINING DREAM

#### A. Model Design

DREAM (Nickovic et al., 2001) has been adapted for use in the southwestern United States, and its performance has been tested and validated using observed weather patterns and dust events. It is a desert dust cycle model developed under the

NCEP/Eta framework (Janjic, 1984; Mesinger et al., 1988; Janjic, 1994) consisting of two modules: an atmospheric simulator, and a dust cycle simulator. The atmospheric simulator parameters include land surface processes, turbulent mixing, convection, large-scale precipitation, lateral diffusion, and radiation. With the Eta, vertical coordinate topography is represented by step-like elements.

The dust cycle module simulates dust production, advection and turbulent diffusion, and dry and wet deposition (Nickovic et al., 2001; Shao et al., 1993; Georgi, 1986). The module consists of three static surface parameters: (1) soil types converted into texture classes at  $2 \times 2'$  resolution; (2)  $10'$  resolution vegetation cover; and (3)  $1 \times 1$  km resolution elevation. Texture categories for sand, silt and clay, which determine the physical properties of wind-blown dust, are assigned according to Cosby et al., (1984). Land cover is from the Olson World Ecosystems (OWE) classification scheme, which contains 59 categories.

#### B. Baseline Performance

The baseline version was run for a domain encompassing most of the western U.S. for two dust-storm events. One occurred on December 8-10, 2003; the other occurred on December 15-17, 2003. Both events were modeled to see how well critical meteorological variables were predicted. A comparison between the observed and model-generated patterns was made to assess: (1) whether the high resolution dust model embedded in NCEP/Eta could forecast Southwest meteorology accurately; and, (2) whether the dust forecasts matched the observed dust measurements.

The DREAM-modeled meteorological fields were compared with measurements and analysis products from 95 surface synoptic sites, 663 surface Meteorological Aerodrome Report (METAR) sites, and 77 upper-air radiosonde sites. The modeled dust field patterns and dust concentrations were compared with satellite images, measured visibility distributions, and surface  $PM_{2.5}$  and  $PM_{10}$  observations made by the Texas Commission on Environmental Quality and the Environmental Protection Agency's (EPA) Air Quality System (AQS). Graphical measures, such as pattern comparison, site against site time series, vertical profile comparison, and statistical metrics, were used.

NCEP/Eta predicts meteorological patterns quite well<sup>1</sup>. Performance of the baseline DREAM model in the American southwest, however, is mixed (Morain and Sprigg., 2005). This suggests that DREAM can be improved by assimilating EO data that replace selected baseline parameters.

### V. ASSIMILATING DATA INTO DREAM

DREAM, like most models used for Earth system science, was not designed to assimilate EO data. Compatibility issues therefore arise, among which are: (a) measurement units, (b)  $x,y,z,t$  resolution, (c) map projection and ease of re-projection to fit model requirements, (d) file formats, (e) error and error propagation, and (f) validity of the data set as a replacement

---

<sup>1</sup> In June 2006, the Eta version of NCEP was superseded by a non-hydrostatic version, NCEP/NMM

input. Assuming that these issues can be overcome, the next steps are to iterate the replacement process with different products and resolutions, and to measure the incremental improvements in model outputs.

Assimilation processes are multifaceted and hampered by a general absence of metadata. DREAM, for example, was designed to use a semi-staggered Arakawa E-grid (Arakawa and Lamb, 1977). The E-grid spacing between neighboring mass (h) and wind (v) points is 0.33 degree. To assimilate higher resolution MODIS land cover data, this spacing had to be reduced to 0.11 degree. Vertically, DREAM uses the Eta coordinate with step-mountain representation (Mesinger et al., 1988). The Eta surfaces are quasi-horizontal in both mountain and non-mountain areas. From sea level to 100 hPa there are 24 half-Eta levels.

A list of candidate products was prepared for assimilation into DREAM. These were intended to replace equivalent surface parameters in the baseline version to achieve finer landscape resolution and more dynamic temporal resolution. They include: (a) land cover from the Moderate Resolution Imaging Spectrometer (MODIS MODs-12,-13, and -15); (b) soil texture from the USDA Natural Resources Conservation Service's (NRCS) Digital General Soil Map of the United States; (c) surface roughness length, "z<sub>0</sub>," from Mod-12 land cover; and (d) soil moisture from the Advanced Microwave Scanning Radiometer (AMSR-E).

MODIS was chosen because an aerosol optical depth product is available, and the aerosol models have been validated (Kaufman and Tanre, 1998; Ichoku et al., 2005; Remer, 2005).

#### A. MOD-15 FPAR

To pinpoint dust source areas, Leaf Area Index (LAI), Enhanced Vegetation Index (EVI), and Fraction of Photosynthetically Active Radiation (FPAR) were examined. The MOD-15 FPAR product may be the most useful since there is an FPAR "class" (value 253) defined as "barren, desert, or very sparsely vegetated." In the FPAR algorithm, this value is known as a "fill" class that includes surface features like water, urban, and permanent snow and ice.

Since the FPAR algorithm requires MOD-12 as an input, it may be possible to use fill class 253 to seasonally update MOD-12 in the DREAM model. This idea has been tested at White Sands National Monument (WSNM), NM. It was hypothesized that wherever value 253 occurred, it could be substituted for equivalent MOD12 pixels to help DREAM identify potential dust sources. The FPAR fill value recognized the slightly vegetated transitional areas and only classified the barren areas as "desert." Errors of omission and commission in MOD-12 over WSNM suggest that the relationship is much more complicated and must be further assessed. Another complicating issue is the effect of winter snow. In an assessment of value 253 over New Mexico for December 2003 and July 2005 (seasonal opposites), it is obvious that fill values are not updated routinely along with non-fill classes.

Visual comparisons of MOD-12 and MOD-15 with commercial satellite products having  $\leq 1\text{m}$  resolution over

sites in southeastern CA, AZ, NM, and west TX suggest that MOD-12 overestimates, and MOD15 underestimates, the area of possible dust generation. Moreover the MOD-12 product seems to identify small (~1km) dust source areas where there are none, especially in eastern NM and west TX. Both products seem to show credible patterns, especially in the larger dust source areas. Another advantage to considering MOD-15 instead of MOD-12 is its more frequent refresh (every 8 days, if the fill values are also updated). MOD-12 was last updated in 2001.

#### B. AMSR-E Soil Moisture

The AMSR-E soil moisture product may be useful as a DREAM input despite a few major issues regarding spatial resolution, data gaps, and reliability. The data footprint is nearly 70km for DREAM outputs aiming toward higher resolution. There is also an absence of data under snow-covered and densely-vegetated areas, so as LAI increases, measurement error increases. Retrieval is not possible under dense vegetation. Lastly, under many conditions, soil moisture can be retrieved only from the surface centimeter (actual sampling depth varies with the amount of surface moisture present). Soil moisture below ~1cm, therefore, may not be sensed. These are all sources of error in the data set, but they do not affect data processing. Moreover, one could argue that recent 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. In DREAM, there is a module called the land surface model (LSM) that treats interactions among soil, vegetation, and atmosphere. LSM 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.

In terms of assimilating AMSR-E soil moisture data, there may be several alternatives. Ultimately the decision will be based on data availability and the quality of those data. The project's strategy is to retrieve the best data available and to develop ways to: (a) augment with other data sources in areas where there are no good measurements; (b) expand with additional satellites and data products; and, (c) as modeling improvements continue, assimilate and evaluate the most promising products for improved model performance.

#### C. Aerosol Optical Thickness

Over land, the MODIS aerosol optical thickness (AOT) is derived by using the dark target approach, but is limited to humid regions (Kaufman and Tanre, 1998). AOT is 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 AOT 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 were ill-defined by the MODIS AOT product. Although AOT products for the desert SW

appear to be geographically incoherent, and are interspersed with many pixels of no data, they do show dust patterns over some parts of the reported dust areas. This could be significant and lead to further improvements in DREAM performance. Table I is a list of the model runs with assimilated EO data. The greatest improvements occurred with MOD-12 (land cover) and SRTM (elevation). Only very minor improvements have been seen with the addition of  $z_0$  (roughness length), FPAR (category 253) and AMSR-E (soil moisture)

TABLE I. MODEL RUNS WITH ASSIMILATED EO DATA

Run #	MOD-12	SRTM	NASA $z_0$	FPAR	AMSR-E
Baseline	No assimilated EO data				
Run 2c	✓				
Run 4a	✓	✓			
Run 5a	✓	✓	✓		
Run 5b	✓	✓	✓		
Run 6a	✓			✓	
Run 10a	✓	✓	✓		✓
Run 15a	✓				✓

Table II compares three parameters in DREAM before and after EO data assimilation. The agreement indices in the bottom row indicate only a slight improvement is achieved for wind speed and wind direction by assimilating MOD-12 data, but that a significant improvement is achieved in the surface temperature parameter. Overall, the higher index values improve the ability of the model to forecast dust entrainment.

TABLE II. DREAM PERFORMANCE BEFORE AND AFTER EO DATA ASSIMILATION

Metrics	Wind Speed (m/s)	Wind Direction (°)	Temp (K)	Definition
Mean Obs.	5.53	231.40	276.74	$\frac{1}{N} \sum_{i=1}^N O_i$
Mean Mod.	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  M_i - O_i $
Agreement Index	0.74 0.75	0.74 0.76	0.71 0.95	$1 - \frac{\sum_{i=1}^N (M_i - O_i)^2}{\sum_{i=1}^N ( M_i - \bar{O}  +  O_i - \bar{O} )}$

*Italic values are before EO data assimilation; other values are after assimilation. For the equations  $M = modeled$ ;  $O = observed$*

#### D. Future Candidate Assimilations

For future data assimilations a list of relevant imagery products has been compiled. Currently these products are documented in a variety of hardcopy books and manuals as well as on many websites. Since they are so dispersed, it is difficult to view them simultaneously for comparison purposes. The EOS Data Products Handbooks, Volumes 2 (GSFC, 2003)

and 1 (GSFC, 2000), have been reviewed for all relevant candidate products. These have been included in an Access Database, resulting in roughly 120 candidates that might have value for the project. The satellites of most immediate interest are Aqua, Terra, TRMM, and Acrimsat. A wide variety of sensors are carried by these platforms including: MODIS, MISR, CERES, ASTER, MOPITT, AMSU-A, AMSR-E, AIRS, HSB, Acrimsat III, TMI, PR, and VIRS.

#### ACKNOWLEDGMENT

The Authors thank all PHAiRS team members for contributing materials that made this paper possible. These team members, in alphabetical order by affiliation, are Karl Benedict, Amelia Budge, Thomas Budge, Gary Sanchez, and William Hudspeth, all of the University of New Mexico; and Brian Barbaris, Christopher Catrall, Beena Chandy, William Sprigg, and Dazhong Yin, all of the University of Arizona.

#### REFERENCES

- [1] 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.
- [2] 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.
- [3] Georgi, F. 1986, 'A particle dry-deposition parameterization scheme for use in tracer transport models,' Journal of Geophysical Research, 91: 9794-9806.
- [4] GSFC. 2000. EOS Data Products Handbook. Volume 1: TRMM, Terra, Data Assimilation System. 258 pages.
- [5] GSFC. 2003. EOS Data Products Handbook. Volume 2. ACRIMSAT, Aqua, Jason-1, Landsat 7, Meteor 3M/SAGE III, Quiksat, QuikTOMS, VCL. 253 pages.
- [6] 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.
- [7] Janjic, Z. I. 1984. "Non-linear advection schemes and energy cascade on semi-staggered grids," Monthly Weather Review, 118: 1234-1245.
- [8] 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.
- [9] Kaufman, Y. and D. Tanre, D., 1998. Algorithm for Remote Sensing of Tropospheric Aerosol from MODIS," MODIS ATDB Product ID: MOD04.
- [10] Mesinger, F., Z.I. Janjic, S. Nickovic, D. Gavrilo 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.
- [11] Morain, S. and W. Sprigg. 2005. Initial Benchmark Report for Public Health. NASA Cooperative Agreement NNS04AA19A. Sep. 30. 36 pages.
- [12] Nickovic, S., G. Kallos, A. Papadopoulos, and O. Kakaliagou. 2001. "A Model for Prediction of Desert Dust Cycle in the Atmosphere," Journ. Geophys. Res., 106(D16): 18,113-18,119.
- [13] Remer, L., 2005. Journal of Atmospheric Sciences. April issue.
- [14] 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.