REPLACING MODEL PARAMETERS WITH EARTH OBSERVATION DATA TO IMPROVE ATMOSPHERIC DUST FORECASTS AND PUBLIC HEALTH RESPONSES

S.A. Morain

Earth Data Analysis Center, University of New Mexico, Albuquerque, NM USA, smorain@edac.unm.edu

Commission VIII, WG VIII/2

KEY WORDS: Modelling, Public Health, Environmental Health, Dust Forecasting, Earth observations

ABSTRACT:

A challenge for assimilating Earth observation data in public health applications is to demonstrate that such data actually improve model predictions of environmental conditions known to trigger health responses. The body of medical and epidemiological knowledge linking environmental health to human health responses is growing rapidly. Through these known linkages, it has become increasingly clear to science and government that satellite observations will play a significant role in forecasting short term weather episodes, as well as mid- to long-term environmental changes that cycle over human generations. Among the approaches to demonstrating the value of Earth observations in environmental modelling is data assimilation. This approach is fundamentally different from other approaches in that it seeks to replace parameters in Earth science models with comparable parameters measured by experimental and operational satellite sensors. Thus, assimilation is a process that incorporates satellite data as part of the model; not as an adjunct to the model. Assimilation is often perceived to be straight forward, but in fact is an exacting process demanding several stages of verification and validation to benchmark model improvements. These stages include: (a) Validating the performance of the model using its baseline design parameters; (b) Assessing sensor products to find candidate satellite measurements to replace baseline parameters; (c) Assimilating candidate data sets and measuring their impact on model performance while retaining the model's original design integrity; (d) Iterating each of the above stages with subsequent candidates, first as a single dataset replacement, then with multiple replacements; (e) Performing statistical analyses that measure and validate step-wise improvements; and, (f) Devising ways to visualize model outputs in ways that are compelling to environmental and public health authorities. This paper presents preliminary results from a project that is replacing parameters in a dust generation model that, itself, is driven by ground-based and satellite weather measurements.

1. INTRODUCTION

The *Public Health Applications in Remote Sensing (PHAiRS*) project (Morain and Sprigg, 2005) has three parallel thrusts. The first focuses on assimilating satellite observations into the Dust Regional Atmospheric Model (DREAM). This model, in turn, is driven by the National Centers for Environmental Prediction (NCEP)/Eta weather forecasting model. The aim is to: (a) verify that advanced satellite image data from current research sensors can replace model parameters from traditional non-satellite sources, or from earlier (coarser resolution) satellite sources; and, (b) validate that parameter replacements lead to more reliable model forecasts of dust episodes.

The second 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.

This paper concentrates on activities and results of the first two thrusts.

2. VALIDATE MODEL PERFORMANCE

2.1 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 parameterization includes land surface processes, turbulent mixing, convection, large-scale precipitation, lateral diffusion and radiation.

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: soil types converted into texture classes at 2'x2' resolution; 10' resolution vegetation cover; and 1x1 km resolution topography. 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.

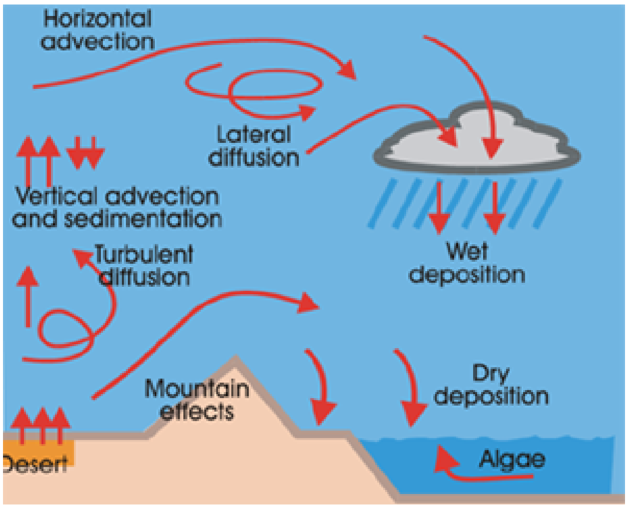

Figure 1. Governing Concept for Dust Entrainment, Diffusion, and Deposition.

2.2 Baseline Performance

The baseline version was run for two dust-storm events. One occurred on December 8-10, 2003; the other occurred on December 15-17, 2003. Both cases 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[1]. 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.

3. ASSESS EO DATA

DREAM, was not designed originally to use 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 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.

[1] In June 2006, the Eta version of NCEP was superseded by a non-hydrostatic version, NCEP/NMM

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.

3.1 Topography / Elevation

A basic parameter for DREAM is an accurate representation of topography. Elevation gives the model a realistic representation of the air-land interface. Terrain induced systems include land-see breeze, mountain valley winds, and forced airflow over and around rough terrain. Data from the Shuttle Radar Topography Mission (SRTM) are used. The most recent version of this data set (released in May 2006) is called SRTM30, the global 30 arcsec [1km] product.

3.2 Land Cover

Land Cover is an important variable in DREAM, mainly as a data source for identifying dust source areas. Having an accurate portrayal of where the dust originates is vital in obtaining accurate model results. For the model to be temporally accurate, it requires up to date land cover data. Currently, DREAM is using the Moderate Resolution Imaging Spectroradiometer (MODIS) Land Cover product, or MOD-12. MOD-12 identifies 17 classes of land cover derived by the International Geosphere-Biosphere Programme (IGBP). Only five of these categories are prone to significant dust generation: Open Shrublands; Grasslands; Croplands; Urban and Built-Up; and, Barren or Sparsely Vegetated land.

3.3 LAI & FPAR

MODIS Leaf Area Index (LAI) and Fraction of Photosynthetically Active Radiation (FPAR), MOD-15, has also been investigated as a potential identifier of dust source areas. LAI measures the one-sided leaf area per unit ground area of vegetation. FPAR measures absorbed wavelengths between 0.4 – 0.7 microns, which are the photosynthetically active wavelengths (Knyazikhin et al., 1999). Dust source areas should have low, or no, LAI or FPAR response. However, the categories having the lowest responses represent six MOD-12 cover types, of which only three are found in semi-arid and arid environments: Unclassified; Urban and built-up areas; and, Barren, desert, or very sparsely vegetated. These three categories seem to match potential dust source areas more accurately than the corresponding MOD-12 Land Cover categories.

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 MOD-12 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. fill values are not updated routinely along with non-fill classes.

Visual comparisons of MOD-12 and MOD-15 with commercial satellite products having ≤ 1m resolution over sites in southeastern CA, AZ, NM, and west TX suggest that MOD-12 overestimates, and MOD-15 underestimates, the area of possible dust generation. Moreover the MOD-12 product seems to identify small (~1km) dust source areas where there may be 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 2003.

3.4 Soil Moisture

The less water held in a soil the more prone it is to wind erosion and dust entrainment. According to Fecan et al. (1999), retention of soil water consists of two factors: (a) molecular adsorption on the surface of the soil grain; and (b) inter-particle capillary forces. The latter of these determines whether dust will be lifted from a surface at a given wind speed. As soil moisture is increased, the threshold wind velocity is also increased, thus reducing the amount of dust injected into the atmosphere (Nickovic et al., 2001).

The Advanced Microwave Scanning Radiometer (AMSR-E) measures passive microwave emissions as a surrogate for soil moisture in the surface few centimetres. However, there are several challenges to overcome before data can be assimilated into DREAM; (a) the effective data footprint is almost 70km, while the model outputs are aiming toward finer resolution; (b) AMSR-E data are available in Hierarchical Data Format (HDF), while DREAM uses an ASCII GRID format; (c) AMSR-E data are formatted to an Equal-Area Scalable Earth Grid or EASE-Grid; (d) there are serious data voids in areas of dense vegetation (high LAI) and under snow cover; and, (e) there are measurement errors associated with sampling depth and vegetation density.

The HDF format stores much information in a single data file. Specific data may be extracted from the file, but special software is required that is not available through open sources. HDF-EOS and GeoTIFF (HEG) tools are available from various data distribution websites, and there are commercial off-the-shelf (COTS) applications such as ENVI and ERDAS that will also read this file format. In sum, AMSR-E data processing requires several steps to make the data format compatible with DREAM. Following these steps, the data must then be re-projected. There are several tools available from the National Snow and Ice Data Center (NSIDC), but the processes are not straight forward.

The project team is addressing these assimilation issues. It believes AMSR-E data may be useful as a DREAM input despite its relatively coarse spatial resolution, data gaps, and accuracy. It is, after all, the best contiguous data product currently available for this most important parameter. 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 at the interface between land and atmosphere, including snow and vegetated areas. When precipitation occurs below 0° C, 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 modelling improvements continue, assimilate and evaluate the most promising products for improved model performance.

3.5 Aerodynamic Roughness

As air flows over a surface, it is disrupted by the topography of that surface. This disruption is related to the aerodynamic roughness length, or z_0, of the surface. This length is different for various surfaces or objects. For vegetated surfaces, z_0 has a 1:1 relationship with the RMS height of vegetation at the canopy top (Saatchi et al., 2001).

In DREAM, surface roughness has been pre-defined using 12 SSiB (Simplified Simple Biosphere Model) land cover types and topography. Measurements made by EO sensors would be a more accurate way to determine this factor. There are experimental ways of determining z_0 using Synthetic Aperture Radar (SAR) technology. Multiple studies and papers are devoted to this topic. Unfortunately, there are no directly usable SAR roughness length products at this time.

For PHAiRS, a look-up table (LUT) was developed as the way to determine surface roughness. A routine developed at NASA's Stennis Space Center merges aerodynamic roughness length values to MOD-12 Land Cover categories. Table 1 shows the values for categories being most prone to dust entrainment.

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 Vegetation	0.00 - 0.01	0.01
253	Fill	0.00	0.00

TABLE I. AERODYNAMIC ROUGHNESS LENGTHS LINKED TO MODIS LAND COVER CATEGORIES (*BLONSKI ET AL., 2005*)

4. ASSIMILATE CANDIDATE EO DATA

The above products were 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 MODs-12, and -15); (b) SRTM 3 arcsec (90m) resampled to 30 arcsec (1km); (c) surface roughness length, z_o, from Mod-12 land cover; and (d) soil moisture from AMSR-E. The orginal baseline parameters and the EO replacement parameters are shown in Table 2.

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

TABLE II. DATA SETS ASSIMILATED INTO DREAM : * SRTM 3 ARCSEC WILL BE REPLACED BY SRTM 30 ARCSEC IN FUTURE MODEL RUNS

The sequence of model runs with assimilated EO data is given in Table 3. Mod-12 land cover has been a consistent parameter replacement, followed by resampled SRTM-90, z_0, soil moisture, and FPAR. To date, DREAM(v.1) has been configured to simulate the two dust storms for December 2003. The team is now developing a 2-3 year running history of meteorological parameters (2003/4-2006 and beyond) to perform statistical analyses on DREAM's performance under day-to-day conditions (v.1A). Yet a third version of the model (v.2) will be configured to run with NCEP/NMM, a new non-hydrostatic version replacing the hydrostatic NCEP/Eta.

Run #	MOD-12	SRTM-30	z_0 (m)	FPAR	AMSR-E
Baseline	No assimilated EO data				
Run 2c	✓				
Run 4a	✓	✓			
Run 5a	✓	✓	✓		
Run 5b	✓	✓	✓		
Run 6a	✓			✓	
Run 10a	✓	✓	✓		✓
Run 15a	✓				✓

TABLE III. MODEL RUN SEQUENCE AS OF MAY 2006

5. ITERATE CANDIDATES

Given the transition to NCEP/NMM, the re-adaptation of DREAM to this parent model, and the versioning of candidate replacement parameters within DREAM, the team has created an experimental design and naming convention for comparing the model results.

Conceptually, the design is analogous to a rack of digital layers each one representing a different EO measurement needed by DREAM. The task is to systematically remove trays from the baseline design of the model and replace them with "fresher" trays of presumed higher value to create an ultimate rack of highest value. To do this for six parameter replacements requires 41-model runs for parameters taken (1, 2, 3…6) at a time, excluding duplicates. Statistical and analytical procedures will be performed to assess which of the model iterations provides the greatest improvements as validated by independent ground-based observation networks.

6. VALIDATE IMPROVEMENTS

If one compares model results before and after MOD-12 data were assimilation, it appears that surface weather patterns (sea level pressure, 500 hPa potential height, and temperature) match well with the observed weather patterns. As hoped for tThis suggests that finer resolution land cover data had little noticeable affect on the performance of the atmospheric simulator. 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.

Similarly, the upper-air fields were *not* affected by the model 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 for the "before" model run had much coarser spatial resolution ($10'$x$10'$) than the run after assimilating MOD-12 data (1x1km). Even though both data sets result in good model performance, finer resolution land cover, combined with topography's influence on surface wind speeds should have an effect on z_0, soil moisture status, and the ability of wind to entrain dust.

The performance statistics of the modeled surface meteorological variables using MOD-12 data showed that model performance in 2m (height above surface) temperature improved by comparison to OWE results. The model performance for 10m wind speed and direction showed slight improvement using assimilated data.

Table 4 lists the performance statistics for modelled surface temperature and wind. The biggest differences between results from before and after MOD-12 data assimilation are for 2m temperature. The agreement index after data assimilation was 0.95, in comparison with 0.71 obtained using the original DREAM parameters This is a significant model improvement. The mean bias and mean error after parameter replacement are less than those for the baseline parameters.

TABLE IV. 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	$\frac{1}{N}\sum_{i=1}^{N}\lvert M_i - O_i\rvert$
Agreement Index	*0.74* 0.75	*0.74* 0.76	*0.71* 0.95	$1-\frac{\sum_{i=1}^{N}(M_i - O_i)^2}{\sum_{i=1}^{N}\left(\lvert M_i-\overline{O}\rvert+\lvert O_i-\overline{O}\rvert\right)}$

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

The agreement index for 10m wind direction and speed was slightly better after MOD-12 data set replacement, but the mean bias and mean error were actually slightly higher than those obtained using the original DREAM parameters.

Figure 2 visualizes the observed and modelled patterns of dust for the storm of December 15-16, 2003. The observed pattern is extrapolated from data obtained by a sparse network of ground reporting air quality stations. The modelled pattern shows more detail but still compares favourably with those same ground observations.

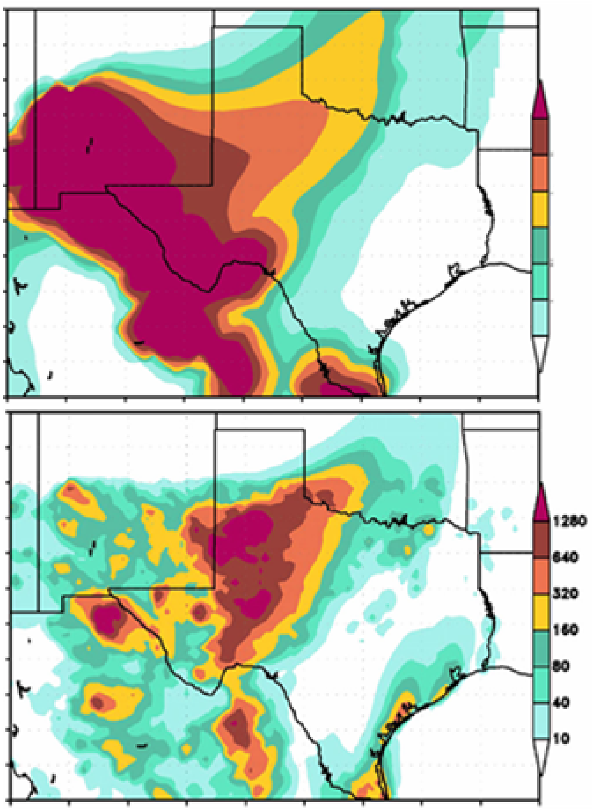

Figure 2. Comparison of DREAM Output: (above) without EO data assimilated, and (below) with EO data assimilated

7. REFERENCES

Arakawa, A., and V.R. Lamb, 1977. Computational design of the basic dynamical processes of the UCLA general circulation model. *Methods in Computational Physics,* 17, pp. 173-265.

Blonski, S., D. Holland, and V. Henderson, 2005. NASA soil moisture data products and their incorporation in DREAM. Contractor Report to UNM Earth Data Analysis Center.

Fecan, F., B., Marticorena, and G. Bergametti, 1999. Parametrization of the increase of the aeolian erosion threshold wind friction velocity due to soil moisture for arid and semi-arid areas. *Annals Geophysicae*, 17, pp. 149-157.

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

Cosby, B.J., G.M. Hornberger, R.B. Clapp, and T.R. Ginn, 1984. A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils. *Water Resources Research*, 20, pp. 682-690.

Janjic, Z.I., 1984. Non-linear advection schemes and energy cascade on semi-staggered grids. *Monthly Weather Review*, 118, pp. 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, pp. 927-945.

Knyazikhin, Y., J. Glassy, J.L. Privette, Y. Tian, A. Lotsch, Y. Zhang, Y. Wang, J.T. Morisette, P. Votava, R.B. Myneni, R.R. Nemani, and S.W. Running, 1999. MODIS leaf area index (LAI) and fraction of photosynthetically active radiation absorbed by vegetation (FPAR) product (MOD15) algorithm theoretical basis document. Version 4.0. http://eospso.gsfc.nasa.gov/atbd/modistables.html (accessed May 23, 2006)

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 alpine lee cyclogenesis and for a case of an Appalachian redevelopment, *Monthly Weather Review*, 116, pp. 1493-1518.

Morain, S. and W. Sprigg, 2005. Initial benchmark report for public health. NASA Cooperative Agreement NNS04AA19A. Sep. 30. 36 pages.

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), pp. 18,113-18,129.

Saatchi, S., E. Rodriguez, S. Denning, and R. Dubayah, 2001. Estimation of aerodynamic roughness from synergistic use of satellite imagery, *Proceeding of IGARSS*, Sydney, Australia.

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, pp. 12719-12726.

8. ACKNOWLEDGMENTS

This research was funded by the National Aeronautics and Space Administration under Agreement NNSO4AA19A. The author wishes to thank the PHAiRS Project Team for their many contribution, in particular Brian Barberis, and Dazhong Yin (University of Arizona Department of Atmospheric Sciences), and Gary Sanchez (University of New Mexico Earth Data Analysis Center).