Earth Data Analysis Center, University of New Mexico
Center for Intelligent Spatial Computing, George Mason University

November 30, 2010

# Project Final Report & Feasibility Report

**NASA Grant/Cooperative Agreement Number**: NNX09AN53G

**Project Title:**

Feasibility of Development and Deployment into Public Health Decision Support systems of a Multi-resolution Nested Dust Forecast System Enabled by Open Standards Based Model and Data Interoperability

**Project PI:** Karl Benedict, Earth Data Analysis Center, University of New Mexico

**Project Co-I:** Phil (Chaowei) Yang, Center for Intelligent Spatial Computing, George Mason University

Report Authors: Karl Benedict, Phil Yang, and Qunying Huang

## Introduction

The timely and efficient production of model outputs, and their fusion with observational data in support of applications and research is an area of growing importance as the volume of data used in analysis and decision support is rapidly growing to scales that require intelligent information management and processing at the data source as a complement to traditional analyses performed by researchers on their own systems [1]. This project has focused on the problem of assessing the feasibility of addressing these needs in two ways:

- speeding the production of dust forecast products through the development of a nested modeling system, the components of which are interconnected through open standard interfaces
- enhancing the availability of model outputs and intermediate products through the automated exposure of those products through service interfaces that are broadly supported by client analytic and visualization applications.

The products of this project have successfully demonstrated the feasibility of the interoperability enabled, nested modeling system, with metrics that compare the performance of the nested system to comparable full-domain high-resolution model runs.

## Project Background

The current project builds upon individual dust forecasting model components that have been developed over the past seven years of work through a series of NASA-funded research and application development projects (Figure 1). In particular, two projects have directly contributed components that are being utilized in this project: the Public Health Applications of Remote Sensing (PHAiRS) project; and an interoperability and high-performance computing testbed project focused on expanding the capabilities of the model components developed by the PHAiRS project.

The PHAiRS project began in 2003 as a collaboration between the University of Arizona (U of A), researchers at the University of Malta (now at the World Meteorological Organization), and the University of New Mexico (UNM), resulting in the development (with funding from NASA's REASoN Program) of an enhanced regionalized version of the Dust Regional Atmospheric Model (DREAM) dust forecasting model for the greater Southwestern US, and a set of OGC and web services for the delivery of model outputs for use in public health decision support systems.

The interoperability and high-performance computing (HPC) testbed (funded by NASA's Applied Sciences Program) built on the products and services developed by the PHAiRS project through the implementation of an 8 particle size bin (instead of the 4-bin DREAM ETA model implemented for the PHAiRS project) ETA DREAM model and an HPC version of the DREAM model implemented within the NCEP NMM meteorological forecasting system. The 8-bin and NMM DREAM modeling components implemented by the interoperability and HPC project provide the modeling cores upon which the current project are based.

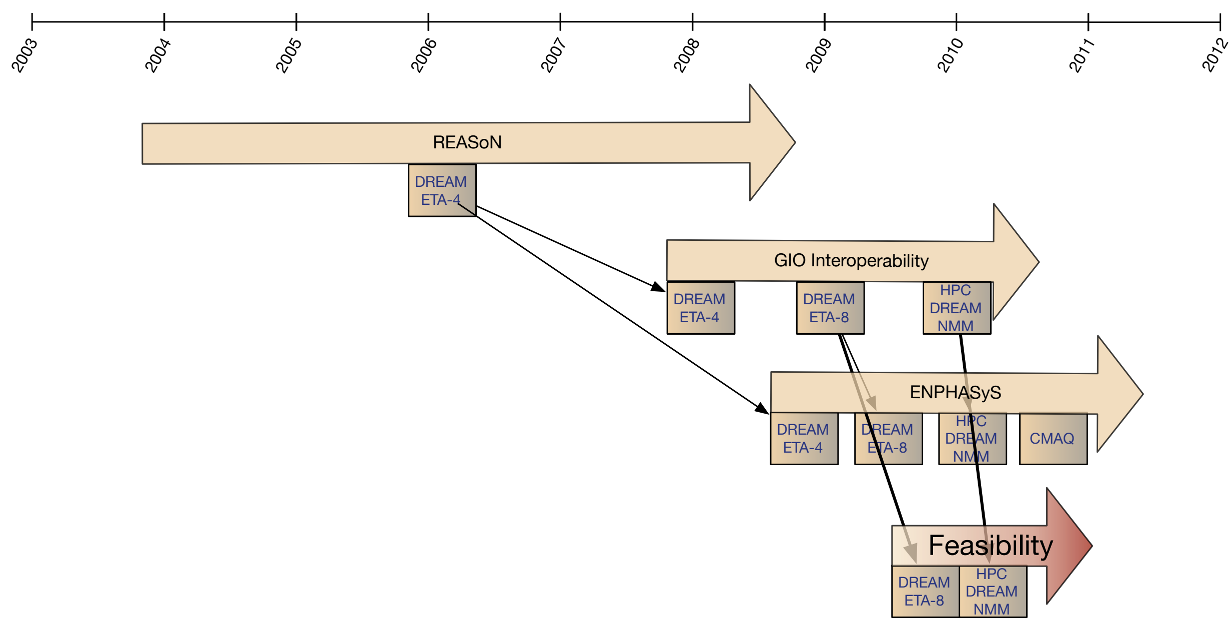



*Figure 1, Dust Modeling Projects Timeline Illustrating the Flow of Project Components Between Projects*

## Project Tasks

The overall goal of this project has been to evaluate the feasibility, through development of an initial set of capabilities and assessment of their performance, of developing a nested dust forecasting system that is based upon open service standards for communication between model components. The core set of project components and the corresponding set of standards-based services interconnecting the components are illustrated in Figure 2, with the timeline for the development of these components (as presented at the Public Health Applications Program Review held in Savannah, GA in September 2009) illustrated in Figure 3.

The high-level project tasks planned for implementation during the year 1 reporting period included the following:

- Publication of NOAA Global Forecast System (GFS) model outputs as OGC Web Coverage Services for initialization and boundary condition definition for DREAM ETA model runs
- Development of DREAM ETA 8-bin model pre- and post-processors for remote interaction with model output management services hosted by EDAC
- Establishment of DREAM ETA 8-bin model output repository and data ingest service
- Development of Area of Interest Processing (AOI) algorithms for definition of model domain sub-regions that will be subject for further high-resolution model execution

Additional project tasks completed in the time following the year-1 reporting period include:

- Publication of OGC and REST services for delivery of initialization and boundary condition parameters, and AOIs generated for a specific DREAM ETA model run

- Development of DREAM NMM pre- and post-processors for remote interaction with the model output management services hosted by EDAC
- Establishment of REST services for submission of DREAM NMM model products from GMU to EDAC
- Establishment of OGC services for delivery of DREAM NMM model outputs.
- Quantitative assessment (for the feasibility analysis) of relative performance characteristics of full-domain model execution when compared with the nested modeling system enabled by interoperable data services

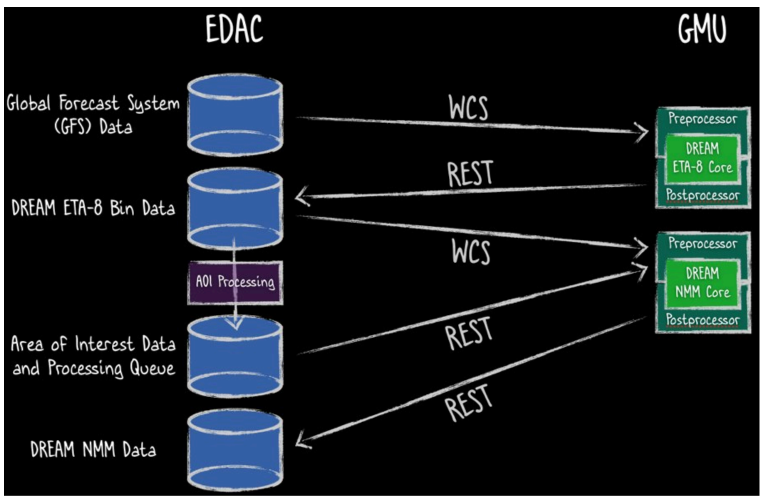



*Figure 2, System Integration Diagram*

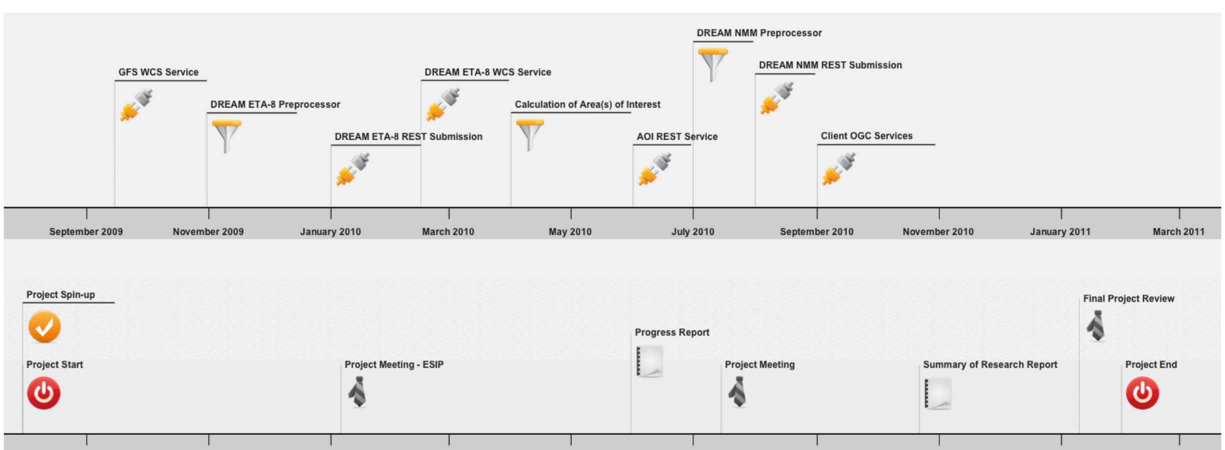



*Figure 3, Project Activity Timeline*

## Results of Research

### *Publication of GFS Data Via Web Coverage Services (WCS)*

The first step in the development of the interoperable standards based nested model workflow being developed and tested for this project consisted of web service based publication of the data required for DREAM model initialization and boundary conditions. This capability was accomplished through the development and hosting of an OGC Web Coverage Service (WCS) for the continuously growing collection of NOAA Global Forecast System (GFS) meteorological forecast products that EDAC has been acquiring since 2006 in support of the multiple dust modeling related projects illustrated in Figure 1 and described above. The developed service is based upon the creation of a dedicated virtual server running the Thematic Realtime Environmental Distributed Data Services (THREDDS) data server developed by Unidata[1]. This server provides access via WCS, WMS, OpenDAP, and HTTP for source GRIB2 formatted GFS data in addition to GRIB1 formatted data acquired directly from NOAA's NOMADS system, and converted GRIB2 data acquired from the operational National Weather Service GFS distribution site (http://129.24.63.106:8080/thredds/catalog.html, Figure 4). NetCDF versions (representing spatial and parameter subsets) of the GRIB data are also available through the hosted WCS service. It is these NetCDF files that are used by the DREAM ETA-8 model pro-processor for acquiring and processing model initialization parameters.

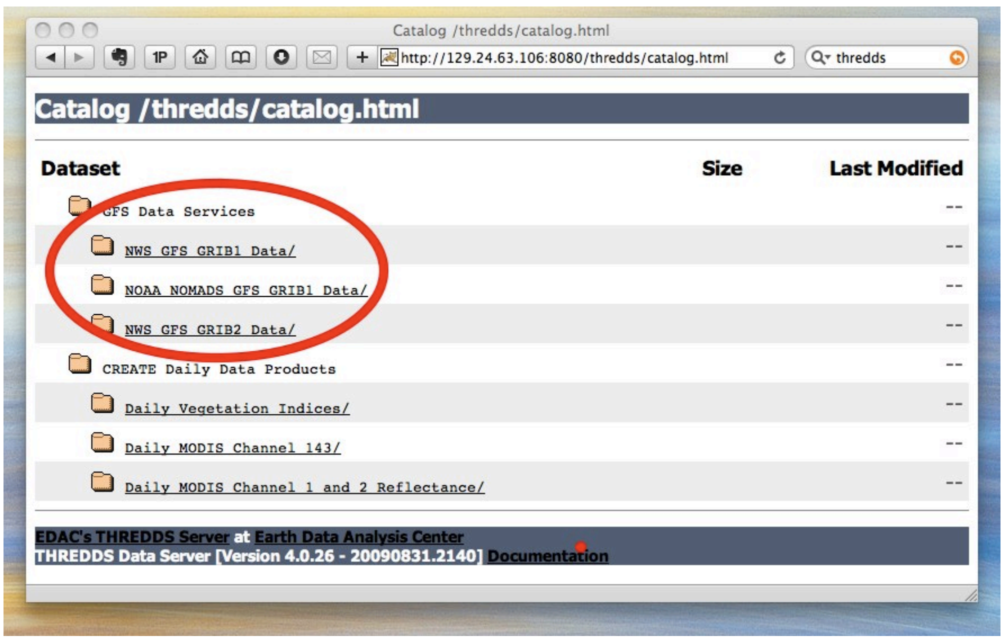



*Figure 4, GFS THREDDS Catalog Hosted by EDAC (GFS data collections circled in the figure)*

[1] http://www.unidata.ucar.edu/projects/THREDDS/

## *Development of DREAM ETA-8 Model Pre- and Post-Processors for Remote Interaction*

Modifications to the model pre- and post-processor components associated with the DREAM Eta-8 models are required to enable model interoperability. Specifically, the model pre-processor code was modified to support the retrieval of initialization parameters from remote systems via WCS requests that return data files which are then subject to further processing to match the required input formats (Fortran binary grids) expected by the model core.

The Post-Processor code required the addition of functionality for automatically pushing the model output to the directories where the THREDDS Data Servers at EDAC and GMU host the DREAM ETA-8 model outputs for dissemination via WCS.

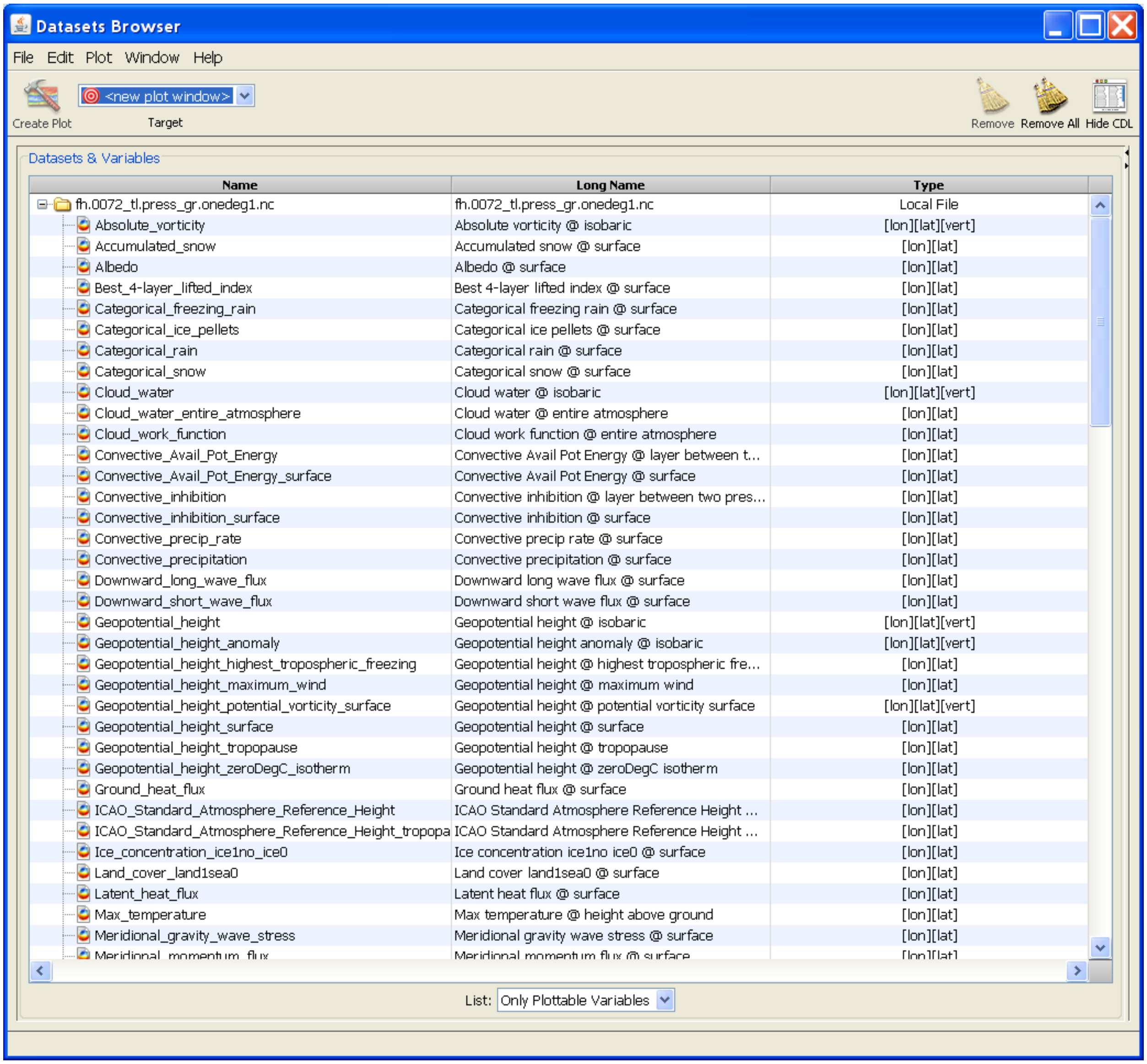
The inputs of DREAM ETA-8 model includes both static and dynamic data sets. The static data sets include of soil type, a SRTM-based digital elevation model for the ground surface, surface roughness, and dust source locations derived from MODIS MOD12 data. The dynamic datasets for model input of DREAM ETA-8 are GRIB1 data, which can be produced from NCEP's NAM[2] or GFS[3] model and contain time-varying meteorological fields. These GRIB files typically contain more fields than needed to initialize DREAM ETA-8. GRIB1 format uses various codes to identify the variables and levels in the GRIB1 file. The *degrib* program uses tables of these codes – called control file "*.ctl", for "variable tables" – to define which fields to extract from the GRIB file and write to an intermediate format.

Figure 5 shows the available variables of one GFS model output, viewed by *Panoply*[4], an application that enables the user to make plots of two-dimensional geo-gridded data from netCDF and HDF datasets. The listing includes more than 100 meteorological parameters that are included in a single GRIB file provided by NOAA for the GFS. However, after an analysis of control file of the *degrib* program and the required input parameters for the DREAM ETA-8 model, we found that there are only 5 dynamic parameters required to run the DREAM ETA-8 model, including geopotential height, relative humidity, temperature, u winds and v winds. These five parameters are provided by many data centers via different models or tools, suggesting that the DREAM ETA-8 model may be executed using meteorological parameters from a wide variety of sources, including the GFS data used in this project. In this way, model interoperability may be facilitated through preparing model input by developing the potential of accessing widely distributed data resources to enable the executions of Earth Science models.

In addition, it is time-consuming to transfer a complete GRIB1 file over the Internet as it includes much more meteorological information than is required to run the DREAM ETA-8 model. For example, a NetCDF file for a single meteorological parameter is about 20 KB (Kilobytes), while a GRIB1 file produced by the GFS model is about 30 MB (Megabytes) and there are 12 GRIB1 files required for a single 24-hour simulation. Therefore, the performance of transferring

[2]http://stu-in-flag.net/nam.php
[3]http://www.emc.ncep.noaa.gov/gmb/moorthi/gam.html
[4]http://www.giss.nasa.gov/tools/panoply/

model input over Internet can be significantly improved through parameter-specific WCS requests for NetCDF files as opposed to full model output GRIB requests.




| Name | Long Name | Type |
|---|---|---|
| fh.0072_tl.press_gr.onedeg1.nc | fh.0072_tl.press_gr.onedeg1.nc | Local File |
| Absolute_vorticity | Absolute vorticity @ isobaric | [lon][lat][vert] |
| Accumulated_snow | Accumulated snow @ surface | [lon][lat] |
| Albedo | Albedo @ surface | [lon][lat] |
| Best_4-layer_lifted_index | Best 4-layer lifted index @ surface | [lon][lat] |
| Categorical_freezing_rain | Categorical freezing rain @ surface | [lon][lat] |
| Categorical_ice_pellets | Categorical ice pellets @ surface | [lon][lat] |
| Categorical_rain | Categorical rain @ surface | [lon][lat] |
| Categorical_snow | Categorical snow @ surface | [lon][lat] |
| Cloud_water | Cloud water @ isobaric | [lon][lat][vert] |
| Cloud_water_entire_atmosphere | Cloud water @ entire atmosphere | [lon][lat] |
| Cloud_work_function | Cloud work function @ entire atmosphere | [lon][lat] |
| Convective_Avail_Pot_Energy | Convective Avail Pot Energy @ layer between t... | [lon][lat] |
| Convective_Avail_Pot_Energy_surface | Convective Avail Pot Energy @ surface | [lon][lat] |
| Convective_inhibition | Convective inhibition @ layer between two pres... | [lon][lat] |
| Convective_inhibition_surface | Convective inhibition @ surface | [lon][lat] |
| Convective_precip_rate | Convective precip rate @ surface | [lon][lat] |
| Convective_precipitation | Convective precipitation @ surface | [lon][lat] |
| Downward_long_wave_flux | Downward long wave flux @ surface | [lon][lat] |
| Downward_short_wave_flux | Downward short wave flux @ surface | [lon][lat] |
| Geopotential_height | Geopotential height @ isobaric | [lon][lat][vert] |
| Geopotential_height_anomaly | Geopotential height anomaly @ isobaric | [lon][lat][vert] |
| Geopotential_height_highest_tropospheric_freezing | Geopotential height @ highest tropospheric fre... | [lon][lat] |
| Geopotential_height_maximum_wind | Geopotential height @ maximum wind | [lon][lat] |
| Geopotential_height_potential_vorticity_surface | Geopotential height @ potential vorticity surface | [lon][lat][vert] |
| Geopotential_height_surface | Geopotential height @ surface | [lon][lat] |
| Geopotential_height_tropopause | Geopotential height @ tropopause | [lon][lat] |
| Geopotential_height_zeroDegC_isotherm | Geopotential height @ zeroDegC isotherm | [lon][lat] |
| Ground_heat_flux | Ground heat flux @ surface | [lon][lat] |
| ICAO_Standard_Atmosphere_Reference_Height | ICAO Standard Atmosphere Reference Height ... | [lon][lat] |
| ICAO_Standard_Atmosphere_Reference_Height_tropopa | ICAO Standard Atmosphere Reference Height ... | [lon][lat] |
| Ice_concentration_ice1no_ice0 | Ice concentration ice1no ice0 @ surface | [lon][lat] |
| Land_cover_land1sea0 | Land cover land1sea0 @ surface | [lon][lat] |
| Latent_heat_flux | Latent heat flux @ surface | [lon][lat] |
| Max_temperature | Max temperature @ height above ground | [lon][lat] |
| Meridional_gravity_wave_stress | Meridional gravity wave stress @ surface | [lon][lat] |
| Meridional_momentum_flux | Meridional momentum flux @ surface | [lon][lat] |

*Figure 5, Screenshot of Panoply Interface Displaying the Meteorological Parameters Contained within a Single GRIB1 GFS file*

Figure 6 shows the work routine of the DREAM ETA-8 Pre-processor to prepare model input for the dust model program for dust storm simulations by DREAM ETA-8 model as provided by a GRIB file. As mentioned above, the *degrib* program (g2g) is used to extract meteorological variables from the GRIB file. The ll.avn2eta program horizontally interpolates the intermediate-format meteorological data that are extracted by the g2g program onto the simulation domains defined by the model configuration file. The interpolated output by the ll.avn2eta can then be ingested by the climmasst, const, anecAVNsoil and ptetaOCT programs. The ptetaOCT program is used to vertically interpolate lateral boundary conditions. Then, the dboco program can be used to produce boundary condition files.

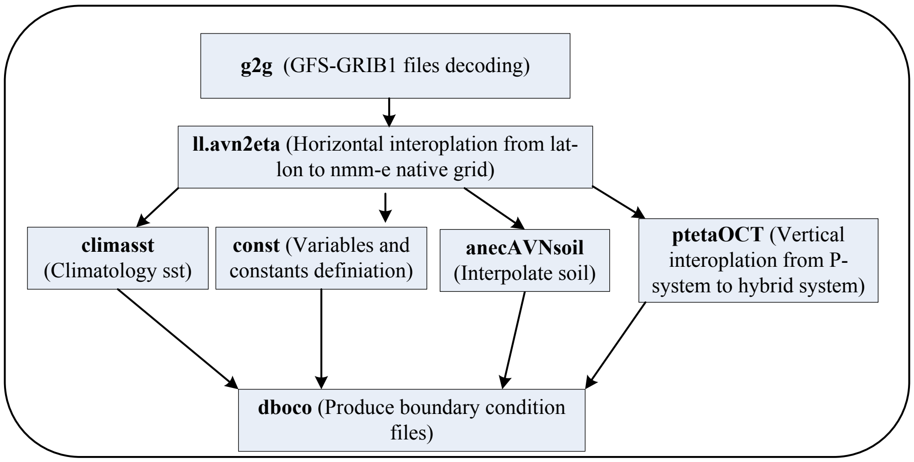



*Figure 6 Workflow of preprocess of DREAM ETA-8 Model*

As an alternative to this GRIB-based workflow, during the implementation, we first obtained each of the five parameters from the THREDDS Data Server hosted by EDAC through WCS, which returns a netCDF file for each parameter at each time interval. Sample shell script code to obtain the parameter of geopotential height (gpm) through a WCS request to EDAC's service is:

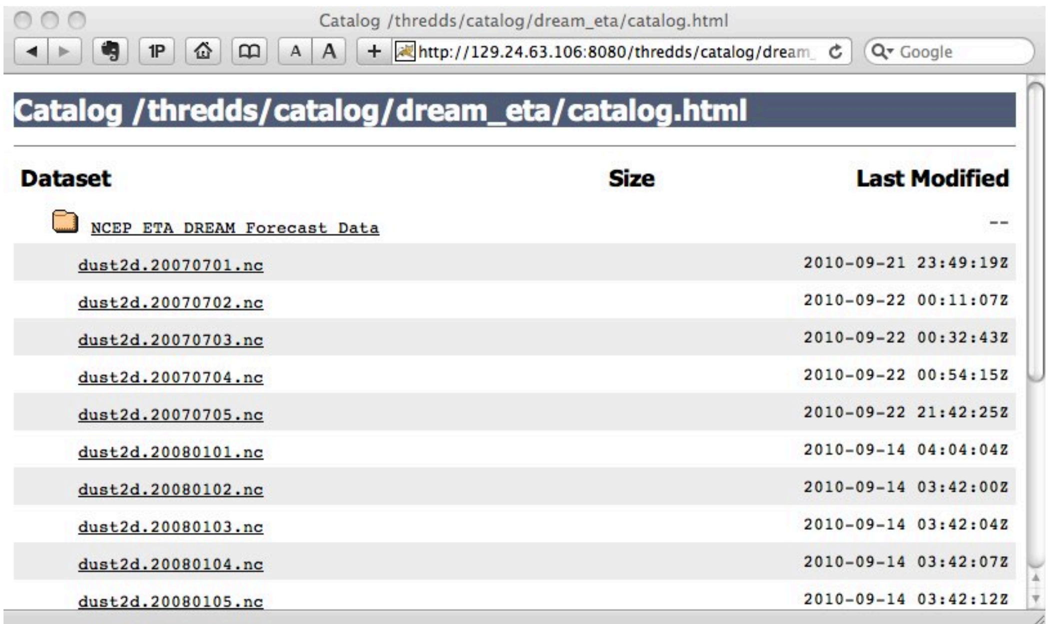
```
#!/bin/csh

set date='20070701'

foreach HH (00 06 12 18 24 30 36 42 48 54 60 66 72)

wget
"http://129.24.63.106:8080/thredds/wcs/nws_grib1//"$date"/fh.00"$HH"_tl
.press_gr.onedeg?request=GetCoverage&version=1.0.0&service=WCS&format=N
etCDF3&coverage=Geopotential_height&bbox=0,-90,359,90" -O
"H_"$date"_"$HH".nc"

end
```

The degrib module (g2g program) of the original preprocess component was then modified to directly extract the meteorological parameters from separate netCDF files instead of the GRIB1 files and write the resulting data to the intermediate format required for final model execution.

Similar to DREAM ETA-8, DREAM NMM model also required the five meteorological parameters (geopotential height, relative humidity, temperature, u winds and v winds) as dynamic model input. In addition to producing the concentration of eight particle size classes, the post-processor of DREAM was also modified to produce the five meteorological parameters that can be assimilated by the high resolution DREAM NMM model.

## *Establishment of DREAM ETA-8 Data Repository and Data Ingest Service*

To facilitate data exchange between the data storage servers at EDAC and the model execution systems at GMU, a THREDDS Data Server was developed at both EDAC and GMU[5] (http://cischpc.gmu.edu/thredds/catalog.html) for DREAM ETA-8 dust model output delivery. Figures 7 and 8 show the available DREAM ETA-8 outputs which are disseminated through EDAC's THREDDS Data Server via multiple protocols, including WMS, WCS, HTTP and OpenDAP.




**Catalog /thredds/catalog/dream_eta/catalog.html**

| Dataset | Size | Last Modified |
|---|---|---|
| NCEP ETA DREAM Forecast Data | | -- |
| dust2d.20070701.nc | | 2010-09-21 23:49:19Z |
| dust2d.20070702.nc | | 2010-09-22 00:11:07Z |
| dust2d.20070703.nc | | 2010-09-22 00:32:43Z |
| dust2d.20070704.nc | | 2010-09-22 00:54:15Z |
| dust2d.20070705.nc | | 2010-09-22 21:42:25Z |
| dust2d.20080101.nc | | 2010-09-14 04:04:04Z |
| dust2d.20080102.nc | | 2010-09-14 03:42:00Z |
| dust2d.20080103.nc | | 2010-09-14 03:42:04Z |
| dust2d.20080104.nc | | 2010-09-14 03:42:07Z |
| dust2d.20080105.nc | | 2010-09-14 03:42:12Z |

*Figure 7 THREDDS data server hosted by EDAC for DREAM ETA-8 model ouputs (http://129.24.63.106:8080/thredds/catalog.html)*

---

[5]http://cischpc.gmu.edu/thredds/catalog.html

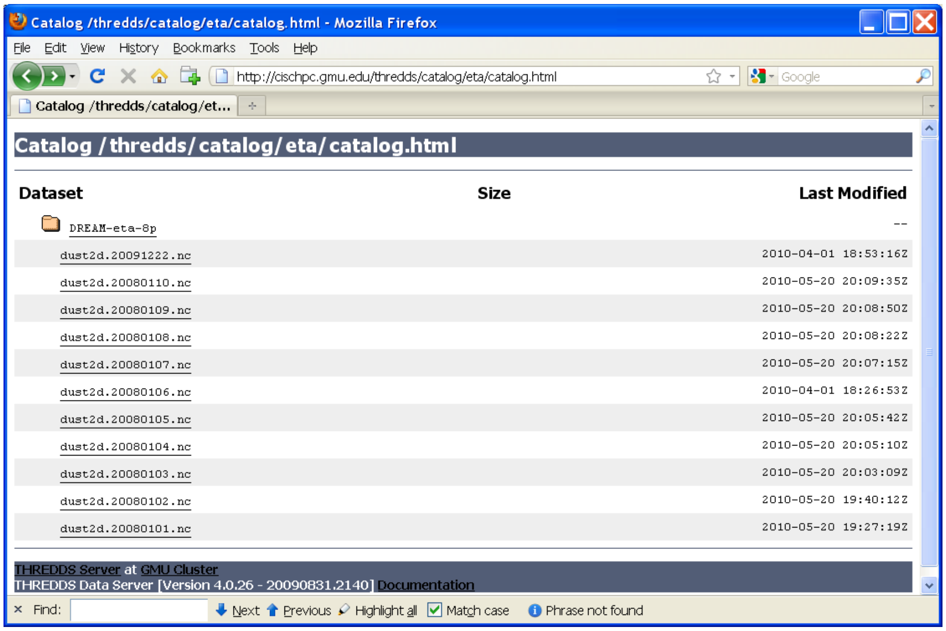



*Figure 8 THREDDSdata server hosted by GMU for DREAM ETA-8 model outputs(http://cischpc.gmu.edu/thredds/catalog.html)*

Model output is named as dust2d.**.nc, where ** is the date of the model output(e.g., dust2d.20091222.nc). For each output file and related WCS service, there are 9 layers and the following four layers are important in the context of the dust forecasting focus of this project: 1) sc, which represents dust concentration (ug/kg), 2) dd, which describes dust dry decomposition load (g/m$^2$), 3) dw, which represents dust Wet Deposition load (g/m$^2$), and 4) dl, which indicates the total dust load (dust wet deposition load plus dust dry deposition load). An example of a WCS request to GMU's service is:

```
http://cischpc.gmu.edu/thredds/wcs/eta/dust2d.20091222.nc?Request=GetCo
verage&version=1.0.0&service=WCS&format=NetCDF3&coverage=sc&bbox=-128,
24, -93, 50&time=2009-12-22T03:00:00Z
```

which enables acquisition of dust load data for the time of 2009/12/22, 03 AM. Figure 9 shows the coarse results simulated by DREAM-ETA 8p dust model for the southwest U.S. region and plotted by parameter "dl" (total dust load).

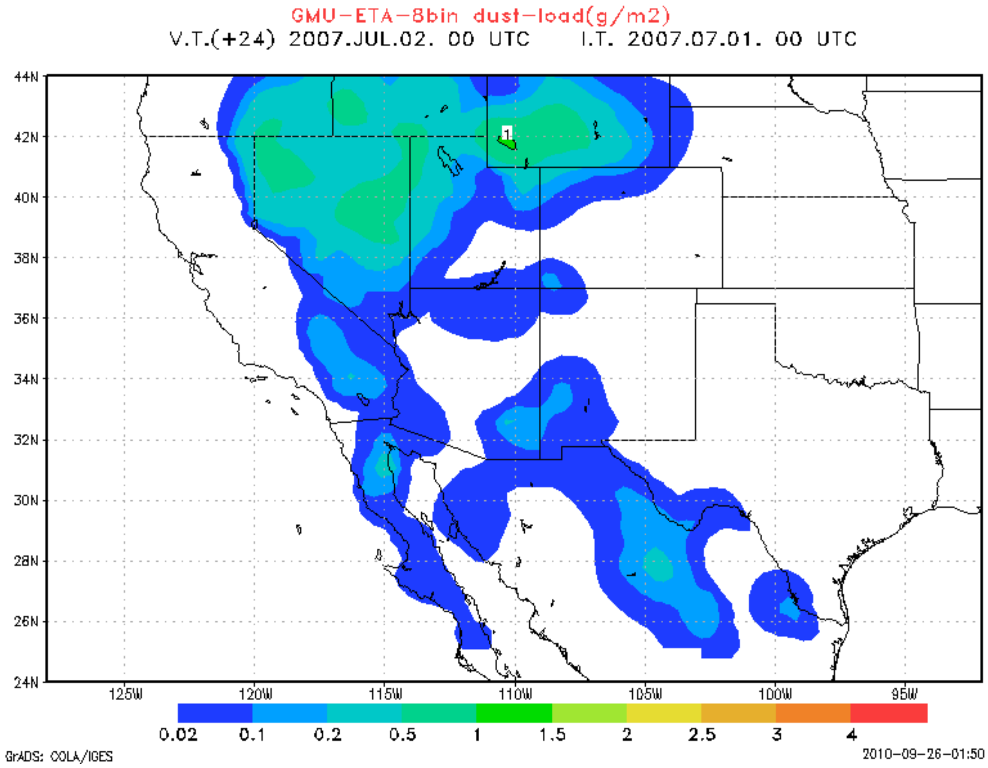



*Figure 9 Southwest U.S. dust storm simulation result from the DREAM ETA-8 model: illustrating a deadly Arizona dust storm event from Jul, 2007.*

As an extension of the interoperability model for the exchange of model outputs between remote systems, an additional REST-style web service has been developed and deployed on the servers hosted at EDAC. Figure  shows a simple client interface to a REST-style Web Service for Dust Output (with the illustrated interface representing the service hosted on EDAC's server). REST (Representational state transfer) is a web service model based upon the HTTP [2] protocol, and defines a model for remote access to and interaction with remote *resources* [3] and as such is based upon the concept of a *resource oriented architecture* [4]. In the case of the REST services developed in support of model interoperability for this project, the standard HTTP requests of *GET, POST, PUT,* and *DELETE* were developed, each of which initiates a different function within the service.

Through the simple web client illustrated in Figure 10 (or through any other client that supports the HTTP protocol), users can obtain the XML file that describes the available DREAM ETA-8 model output datasets (Figure 71). This XML file includes model information, date of model output, and access URLs (Uniform Resource Locator) to data sets via different types of protocols, including WMS, WCS, HTTP, and OpenDAP. The root element of the XML file is "outputs", and it contains multiple child elements with each child element including the information about one DREAM ETA-8 model output. Users can also use the "POST" method to push the model output to the EDAC data center by providing the information to construct a record for a

DREAM ETA-8 model output in the XML file, e.g., model type, date of model output and accessing URL (Figure 8).


*Figure 10 A Simple Client Interface for the REST-style Web Service for Dust Model Outputs Generated at GMU*

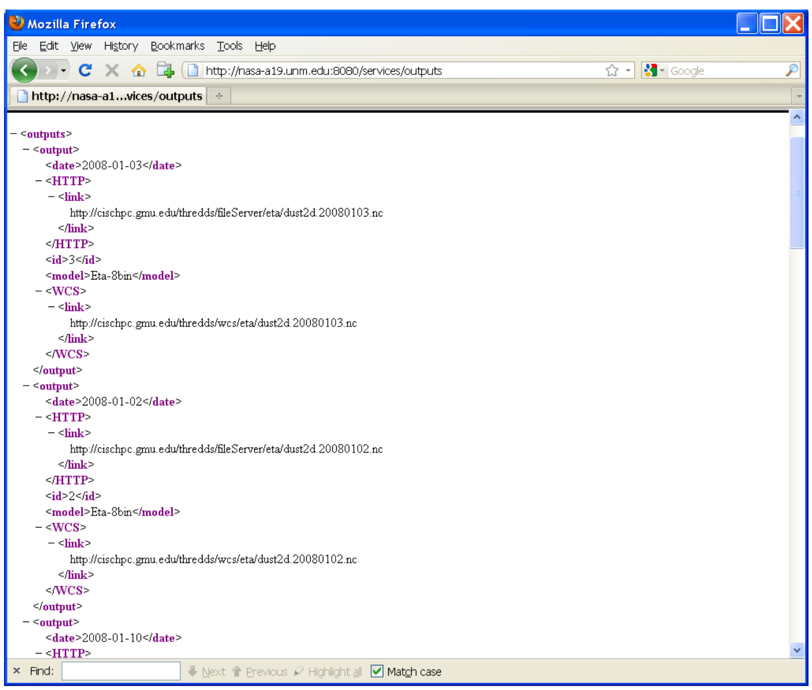



*Figure 71, DREAM ETA-8 Model Output Information Obtained through an HTTP GET Request to the Developed REST Services*

Unlike Simple Objects Access Protocol (SOAP) protocol [5] which has proved to be a powerful standard for web services but requires a client program and dedicated XML processing to encapsulate messages and extract responses from the XML envelopes that are a core of SOAP, users of REST-style web service only need to know the XML file URL, read it with a Web browser, interpret the content data using the XML information, and reformat and use it appropriately. For example, users only need to use the URL "http://nasa-a19.unm.edu:8080/services/outputs" to obtain the available DREAM ETA-8 model output datasets information from the REST web service hosted in the EDAC data center, and use the URL "http://nasa-a19.unm.edu:8080/services/outputs/push" to push data back to the EDAC data center. A REST data delivery component was developed using JAVA to invoke the "POST" method of REST web service hosted at EDAC data center to harvest the model output from the GMU through WCS. This component is integrated to the Post-Processors of DREAM ETA-8.

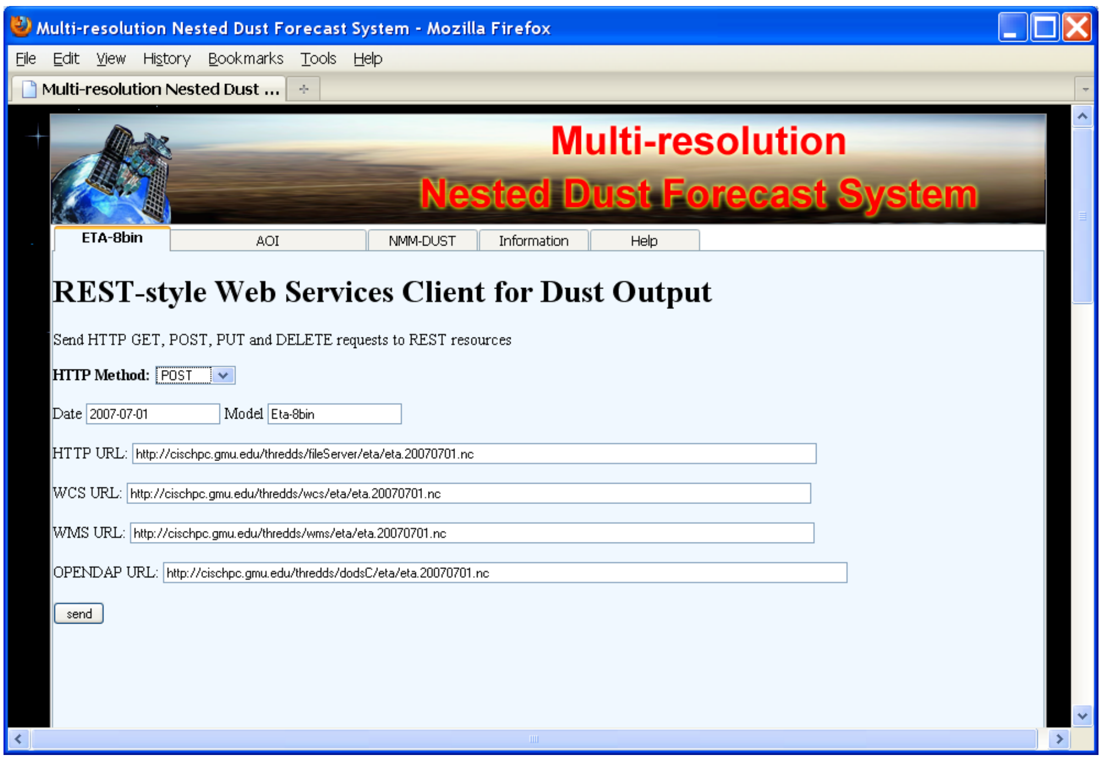



*Figure 82. DREAM ETA-8 REST Service Interface for Submitting Information Required to Create a New XML Element in the Collection XML Returned by the Service*

### *AOI Processing Algorithm and Processing into the Queue*

The problem of defining areas of interest for further, high-resolution model execution using the DREAM NMM model has been approached through the development of a general processing algorithm that has been implemented in the Python programming language as a series of matrix processing procedures that apply specified thresholds to collections of individual DREAM ETA-

8 output rasters. In combination these thresholds define the area ‘mask’ that will be subject to further execution in the DREAM NMM model (Figure 9).

The processing algorithm proceeds through the following steps:

1. Retrieval of NetCDF output file generated by GMU and placed on the data server at EDAC
2. Extraction from the NetCDF file the subdataset (e.g. “Total Dust Load”: parameter “dl”) of interest
3. Processing of the 200 bands associated with the subdataset: 25 3-hour forecast products for a 72 hour forecast X 8 particle sizes
4. Identification of the pixels within each band of interest (e.g. the bands associated with particle size 1) that exceed a specified threshold value
5. Counting of the exceedence pixels from step 4 to develop a single raster that represents the number of forecast time steps that exceeded the specified threshold.
6. Identification of the pixels where the exceedence count exceeds a specified value

With the areas of contiguous pixels identified the raster produced by step 6 representing the areas of interest for subsequent model execution for the DREAM NMM model, output data formats representing those AOIs may be generated for publication via HTTP.

Given the rapidity with which the AOI analysis may be performed (a full analysis may be performed in under 30-seconds), and the flexibility of the developed code for identifying AOIs for further processing (i.e. the threshold values may be defined on demand when the processing script is executed, producing custom outputs depending upon the needs of a specific analysis or modeling effort) further development of static processing queue products is on hold until a need for stored AOIs is defined through further work.

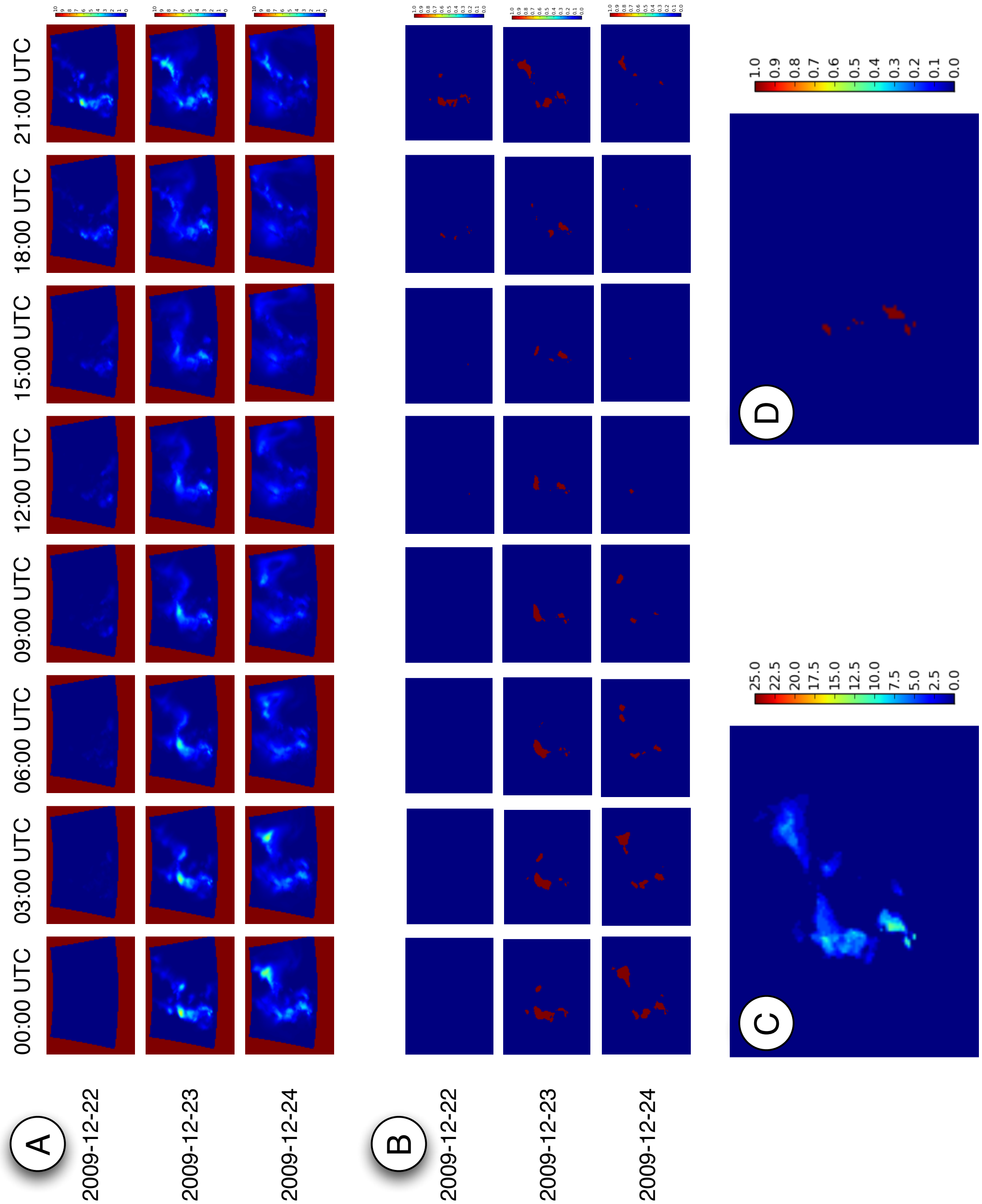
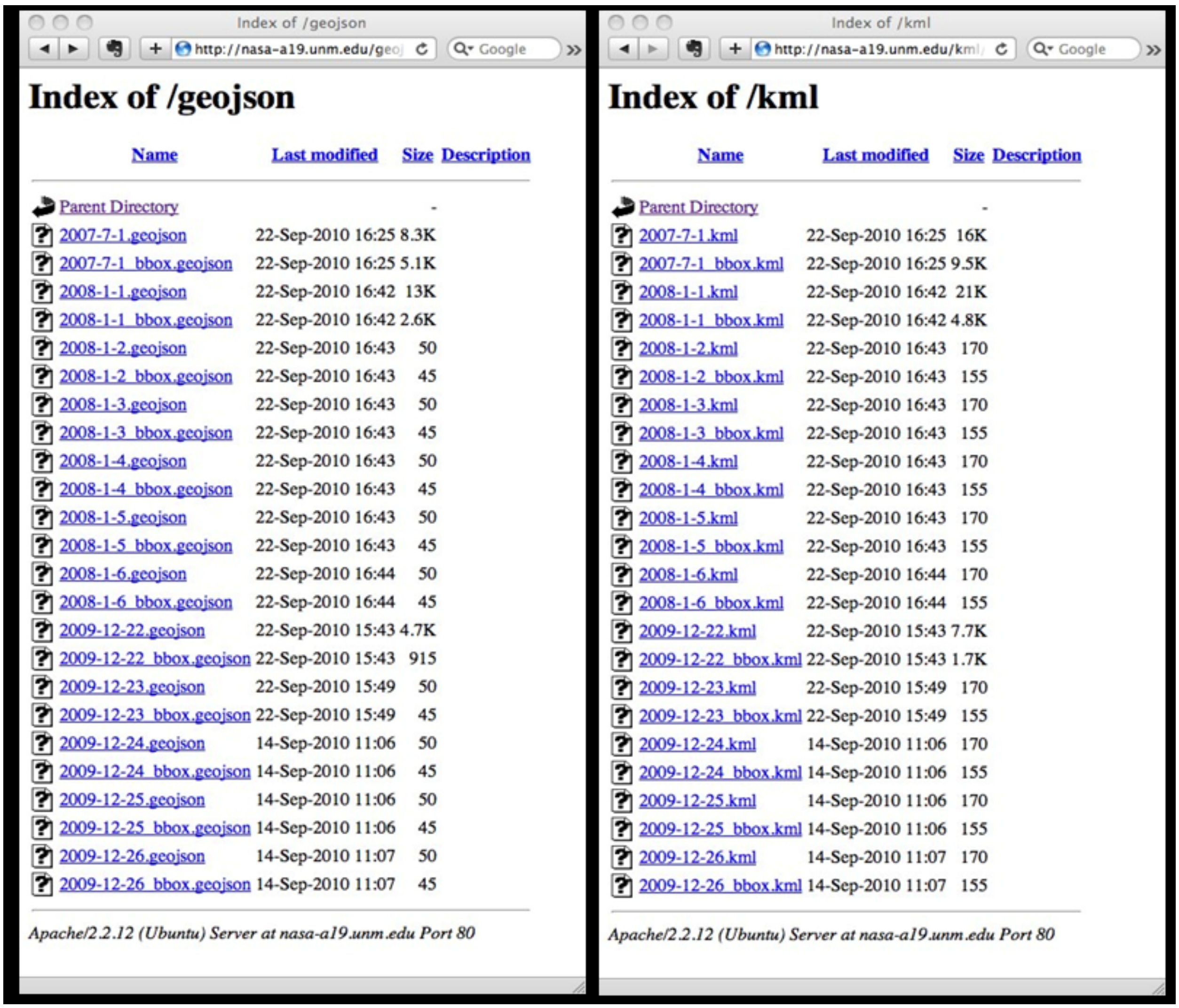



*Figure 93, AOI Processing Illustration. A) Twenty-four 3-hour total deposition (parameter "dl") values for DREAM ETA-8 particle size 1(ug/m2); B) Regions within each 3-hour forecast product that exceed a value of 2.0 ug/m2; C) Count of 3-hour forecast products that exceed 2.0 ug/m2 out of the 25 rasters generated by a single model run; D) the regions in the DREAM ETA-8 model run where more than 8 3-hour forecast products exceed the defined 2 ug/m2 threshold*

For example, 18 AOIs (Area of Interest) are found through the AOI processing on the output of DREAM ETA-8 for July 1st , 2007. The AOIs are provided in both GeoJSON and KML formats (Figure 14). Figure 15 shows the AOIs transmitted through GeoJSON[6]. GeoJSON is a format that can be used to encode a variety of geometries and their associated properties (attributes). GeoJSON allows geographic data to be stored in a human-readable way that is generally more compact than XML.



Index of /geojson

http://nasa-a19.unm.edu/geo

# Index of /geojson

| Name | Last modified | Size | Description |
|---|---|---|---|
| Parent Directory | | - | |
| 2007-7-1.geojson | 22-Sep-2010 16:25 | 8.3K | |
| 2007-7-1_bbox.geojson | 22-Sep-2010 16:25 | 5.1K | |
| 2008-1-1.geojson | 22-Sep-2010 16:42 | 13K | |
| 2008-1-1_bbox.geojson | 22-Sep-2010 16:42 | 2.6K | |
| 2008-1-2.geojson | 22-Sep-2010 16:43 | 50 | |
| 2008-1-2_bbox.geojson | 22-Sep-2010 16:43 | 45 | |
| 2008-1-3.geojson | 22-Sep-2010 16:43 | 50 | |
| 2008-1-3_bbox.geojson | 22-Sep-2010 16:43 | 45 | |
| 2008-1-4.geojson | 22-Sep-2010 16:43 | 50 | |
| 2008-1-4_bbox.geojson | 22-Sep-2010 16:43 | 45 | |
| 2008-1-5.geojson | 22-Sep-2010 16:43 | 50 | |
| 2008-1-5_bbox.geojson | 22-Sep-2010 16:43 | 45 | |
| 2008-1-6.geojson | 22-Sep-2010 16:44 | 50 | |
| 2008-1-6_bbox.geojson | 22-Sep-2010 16:44 | 45 | |
| 2009-12-22.geojson | 22-Sep-2010 15:43 | 4.7K | |
| 2009-12-22_bbox.geojson | 22-Sep-2010 15:43 | 915 | |
| 2009-12-23.geojson | 22-Sep-2010 15:49 | 50 | |
| 2009-12-23_bbox.geojson | 22-Sep-2010 15:49 | 45 | |
| 2009-12-24.geojson | 14-Sep-2010 11:06 | 50 | |
| 2009-12-24_bbox.geojson | 14-Sep-2010 11:06 | 45 | |
| 2009-12-25.geojson | 14-Sep-2010 11:06 | 50 | |
| 2009-12-25_bbox.geojson | 14-Sep-2010 11:06 | 45 | |
| 2009-12-26.geojson | 14-Sep-2010 11:07 | 50 | |
| 2009-12-26_bbox.geojson | 14-Sep-2010 11:07 | 45 | |

*Apache/2.2.12 (Ubuntu) Server at nasa-a19.unm.edu Port 80*

Index of /kml

http://nasa-a19.unm.edu/kml

# Index of /kml

| Name | Last modified | Size | Description |
|---|---|---|---|
| Parent Directory | | - | |
| 2007-7-1.kml | 22-Sep-2010 16:25 | 16K | |
| 2007-7-1_bbox.kml | 22-Sep-2010 16:25 | 9.5K | |
| 2008-1-1.kml | 22-Sep-2010 16:42 | 21K | |
| 2008-1-1_bbox.kml | 22-Sep-2010 16:42 | 4.8K | |
| 2008-1-2.kml | 22-Sep-2010 16:43 | 170 | |
| 2008-1-2_bbox.kml | 22-Sep-2010 16:43 | 155 | |
| 2008-1-3.kml | 22-Sep-2010 16:43 | 170 | |
| 2008-1-3_bbox.kml | 22-Sep-2010 16:43 | 155 | |
| 2008-1-4.kml | 22-Sep-2010 16:43 | 170 | |
| 2008-1-4_bbox.kml | 22-Sep-2010 16:43 | 155 | |
| 2008-1-5.kml | 22-Sep-2010 16:43 | 170 | |
| 2008-1-5_bbox.kml | 22-Sep-2010 16:43 | 155 | |
| 2008-1-6.kml | 22-Sep-2010 16:44 | 170 | |
| 2008-1-6_bbox.kml | 22-Sep-2010 16:44 | 155 | |
| 2009-12-22.kml | 22-Sep-2010 15:43 | 7.7K | |
| 2009-12-22_bbox.kml | 22-Sep-2010 15:43 | 1.7K | |
| 2009-12-23.kml | 22-Sep-2010 15:49 | 170 | |
| 2009-12-23_bbox.kml | 22-Sep-2010 15:49 | 155 | |
| 2009-12-24.kml | 14-Sep-2010 11:06 | 170 | |
| 2009-12-24_bbox.kml | 14-Sep-2010 11:06 | 155 | |
| 2009-12-25.kml | 14-Sep-2010 11:06 | 170 | |
| 2009-12-25_bbox.kml | 14-Sep-2010 11:06 | 155 | |
| 2009-12-26.kml | 14-Sep-2010 11:07 | 170 | |
| 2009-12-26_bbox.kml | 14-Sep-2010 11:07 | 155 | |

*Apache/2.2.12 (Ubuntu) Server at nasa-a19.unm.edu Port 80*

*Figure 14 AOI in GeoJSON and kml formats*

[6] http://geojson.org/

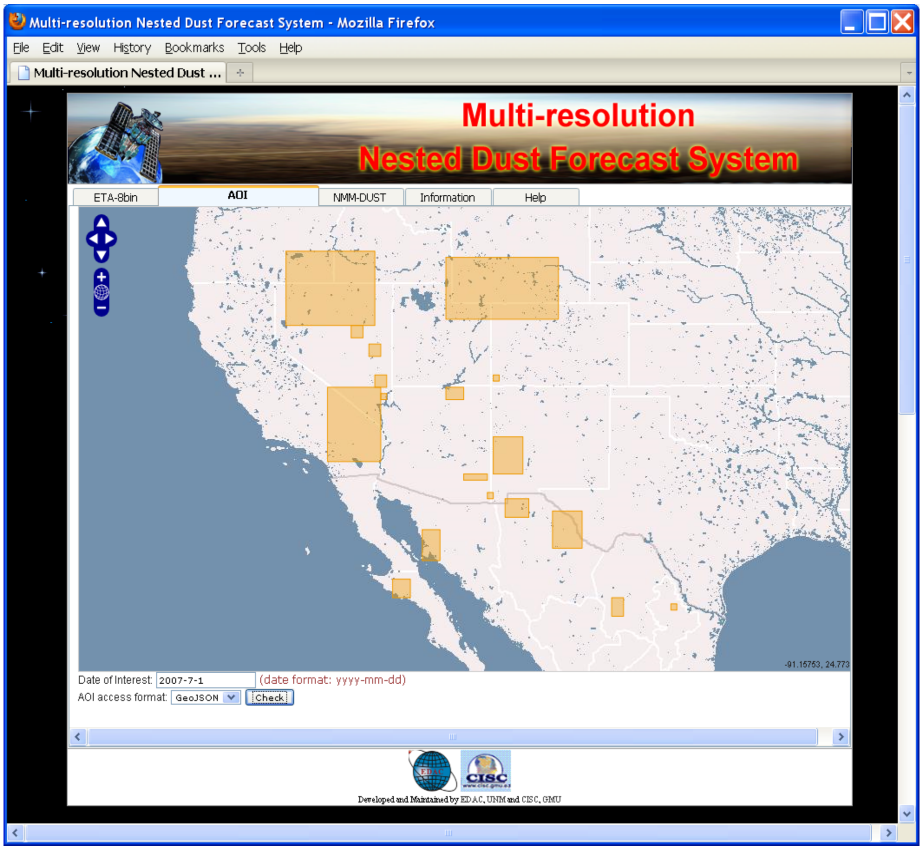



*Figure 15. The AOIs for July 1st, 2007, accessed through GeoJSON*

## *Development of DREAM NMM Model Pre- and Post-Processors for Remote Interaction*

Modifications to the model pre- and post-processor components associated with the DREAM NMM model are required to enable it assimilate the output of DREAM ETA-8 for the "hot" start run, which incorporates the dust information obtained from the output of DREAM ETA-8 during initialization.

Similar to DREAM ETA-8, the DREAM NMM's pre-process module was modified to enable the direct assimilation of the model output from DREAM ETA-8 through WCS. After ingesting the DREAM-eta 8p output, the DREAM NMM can start simulation and run each hotspot area in parallel within GMU HPC environment.

DREAM-NMM output is in GrADS[7] binary format and includes the parameters of mean sea level pressure (pa), pressure reduced to mean sea level (pa), pressure (pa), geopotential height (gpm), u wind (m/s), v wind (m/s), pressure vertical velocity (Pa/s), turbulent kinetic energy

[7] http://www.iges.org/grads/

(J/kg), for 45 levels. It also includes one layer of relative humidity (%), dust concentration (ug/$m^3$), dust load (g/$m^2$), temperature, and other meteorological parameters. The post-processor component of DREAM NMM was modified to extract the parameters of dust load and dust concentration into netCDF format. Additionally, the post-processor of the DREAM NMM model was modified to integrate the REST data delivery component enabling the transfer of model output back to EDAC data center.

### *Develop DREAM NMM REST Web Service Submission Capability and Publication of Model Outputs as OGC Services*

The REST service already developed and deployed (as described above) on the servers both at EDAC and GMU already provides the foundation for exchange of DREAM NMM model outputs between the servers at GMU (where the models are being run) and the data management system at EDAC (the repository for storage and distribution of all project intermediate and final data products). As DREAM NMM model outputs are generated, the existing REST service was modified as needed to also manage and support the new collection of DREAM NMM model outputs. After the DREAM NMM finishes the simulation for each each AOI, the modified post-processor produces dust concentration and dust load data in netCDF format. Another module is invoked to call the REST services to push these netCDF files to EDAC. The post-processor is triggered automatically once results are generated through model simulation. The netCDF data are put in the directory used to publish DREAM NMM model results through THREDDS data server at the EDAC. The configuration of EDAC's THREDDS server was extended to enable publication of the DREAM NMM model outputs using the same services (WCS, WMS, HTTP, and OPeNDAP). Figure 16 shows 18 AOI simulation results identified from DREAM ETA-8 outputs, produced by DREAM NMM for July 1$^{st}$ , 2007, and published by the THREDDS server at EDAC.

Catalog /thredds/catalog/dream_nmm/20070701/catalog.html - Mo...

File Edit View History Bookmarks Tools Help

Catalog /thredds/catalog/dr...

**Catalog /thredds/catalog/dream_nmm/20070701/catalog.html**

| Dataset | Size | Last Modified |
|---|---|---|
| 20070701 | | -- |
| nmm.20070701.aoi-001.nc | | 2010-11-22 22:25:40Z |
| nmm.20070701.aoi-002.nc | | 2010-11-22 22:25:45Z |
| nmm.20070701.aoi-003.nc | | 2010-11-22 22:26:07Z |
| nmm.20070701.aoi-004.nc | | 2010-11-22 22:26:11Z |
| nmm.20070701.aoi-005.nc | | 2010-11-22 22:26:26Z |
| nmm.20070701.aoi-006.nc | | 2010-11-22 22:26:41Z |
| nmm.20070701.aoi-007.nc | | 2010-09-27 05:10:09Z |
| nmm.20070701.aoi-008.nc | | 2010-11-22 23:02:47Z |
| nmm.20070701.aoi-009.nc | | 2010-11-22 23:02:58Z |
| nmm.20070701.aoi-010.nc | | 2010-09-27 05:15:33Z |
| nmm.20070701.aoi-011.nc | | 2010-11-22 22:19:52Z |
| nmm.20070701.aoi-012.nc | | 2010-09-27 05:11:18Z |
| nmm.20070701.aoi-013.nc | | 2010-09-27 05:13:15Z |
| nmm.20070701.aoi-014.nc | | 2010-11-22 23:03:40Z |
| nmm.20070701.aoi-015.nc | | 2010-09-27 05:12:28Z |
| nmm.20070701.aoi-016.nc | | 2010-11-22 23:04:19Z |
| nmm.20070701.aoi-017.nc | | 2010-11-22 23:04:36Z |
| nmm.20070701.aoi-018.nc | | 2010-11-22 23:04:41Z |

EDAC's THREDDS Server at Earth Data Analysis Center
THREDDS Data Server [Version 4.0.26 - 20090831.2140] Documentation

*Figure 16. THREDDSdata server hosted by EDAC for DREAM NMM model outputs*

## Feasibility Analysis

The project results demonstrate that it is feasible to utilize interoperability and distributed HPC to support the dust storm forecasting and other data intensive modeling efforts. Specially, the dust event on July 1$^{st}$, 2007 was used to test the feasibility of the multi-resolution nested dust storm framework. Figures 17, 18 and 19 show different time-frame results produced by DREAM ETA-8 and DREAM NMM models at different spatial resolutions. The model results shows similar pattern for areas of elevated dust concentrations and the DREAM NMM with 3 km resolution displays much more detailed information about the dust concentrations than the two lower resolution model products.

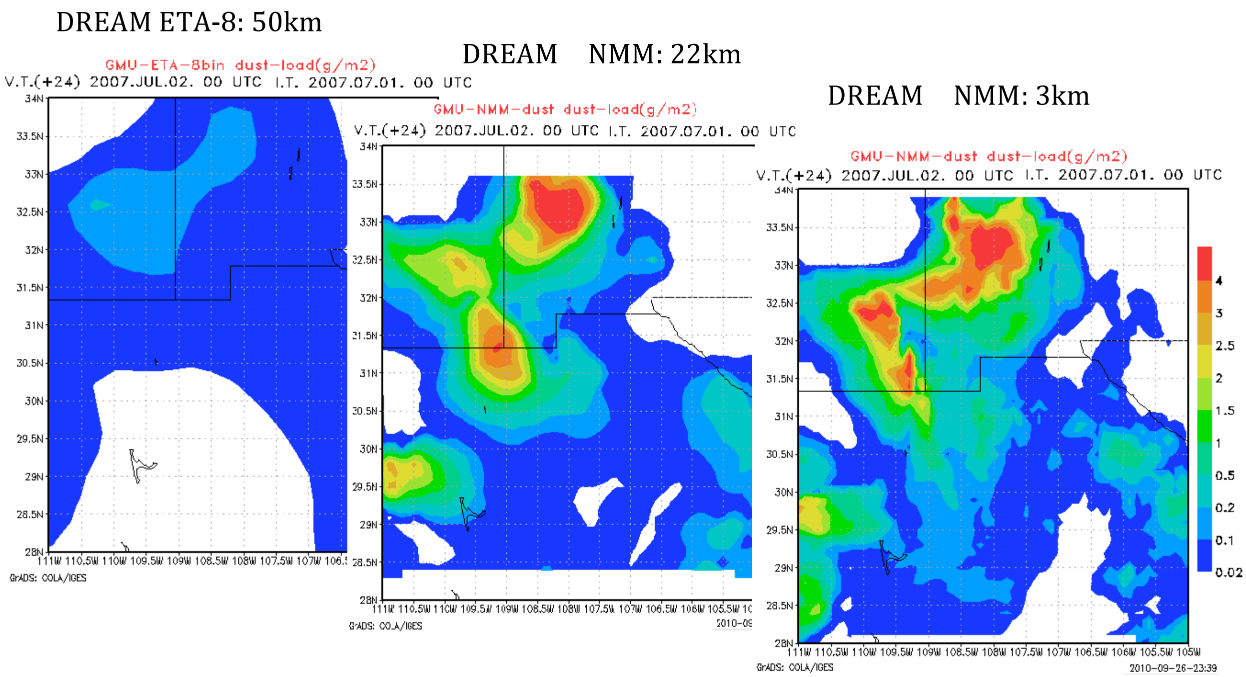



*Figure 16. Comparison of the simulation results by DREAM ETA-8 and DREAM NMM for AOIs 10, 11, 12 and 13 at 00 AM UTC , July 02, 2007 for a set of model runs initiated for July 1, 2007 at 00 AM UTC*

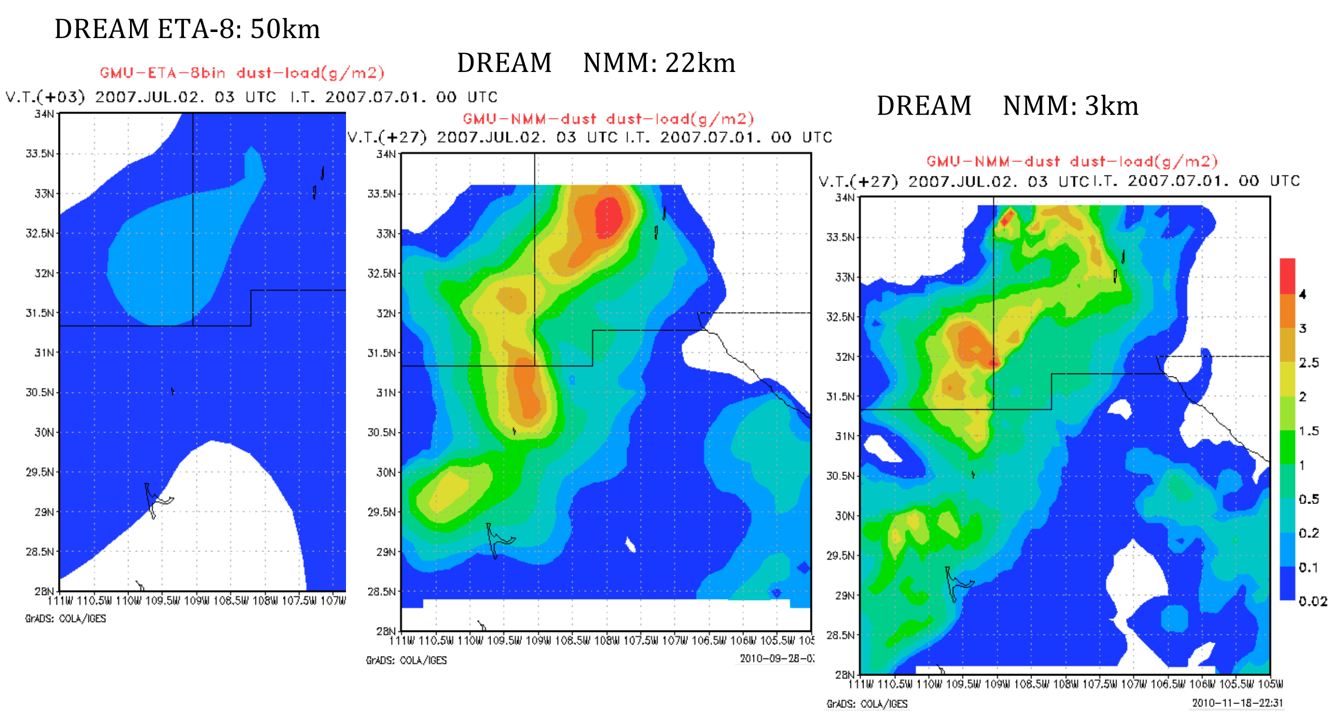



*Figure 17. Comparison of the simulation results by DREAM ETA-8 and DREAM NMM on AOIs 10, 11, 12 and 13 at 03 AM , July 02, 20077 for a set of model runs initiated for July 1, 2007 at 00 AM UTC*

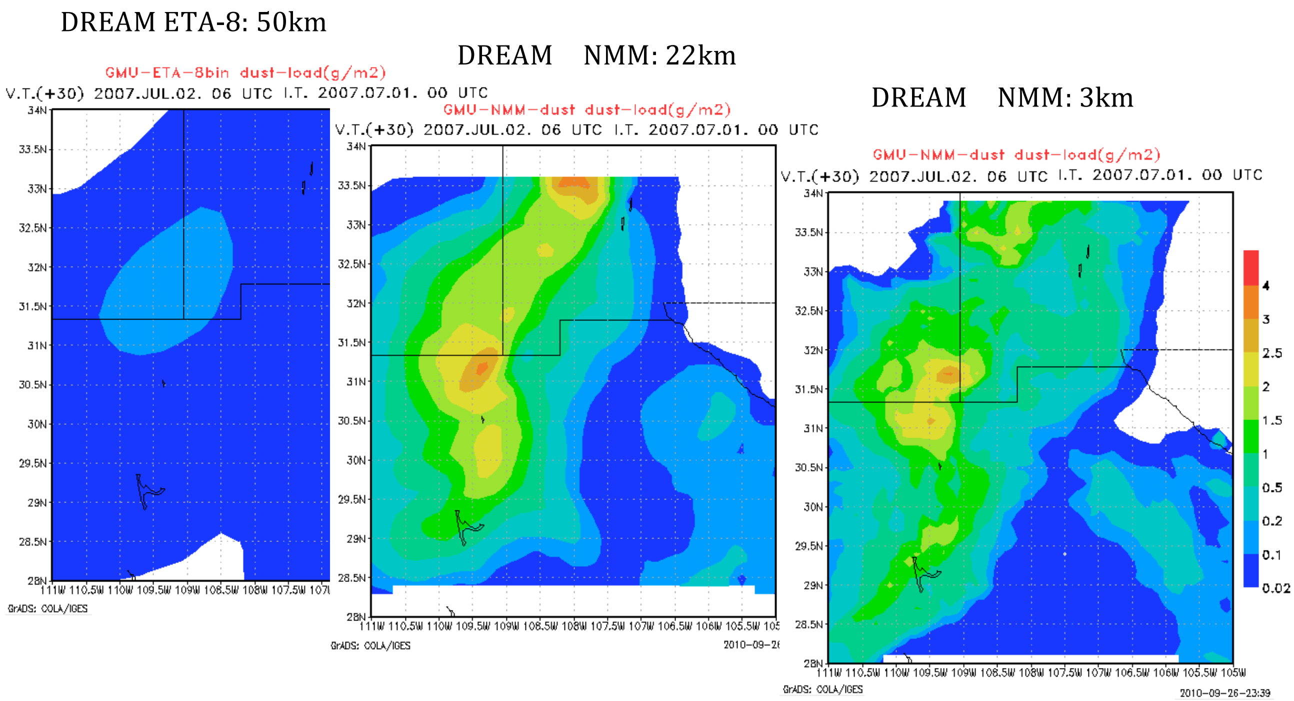



*Figure 18. Comparison of the simulation results by DREAM ETA-8 and DREAM NMM on AOIs 10, 11, 12 and 13 at 06 AM, July 02, 2007 for a set of model runs initiated for July 1, 2007 at 00 AM UTC*

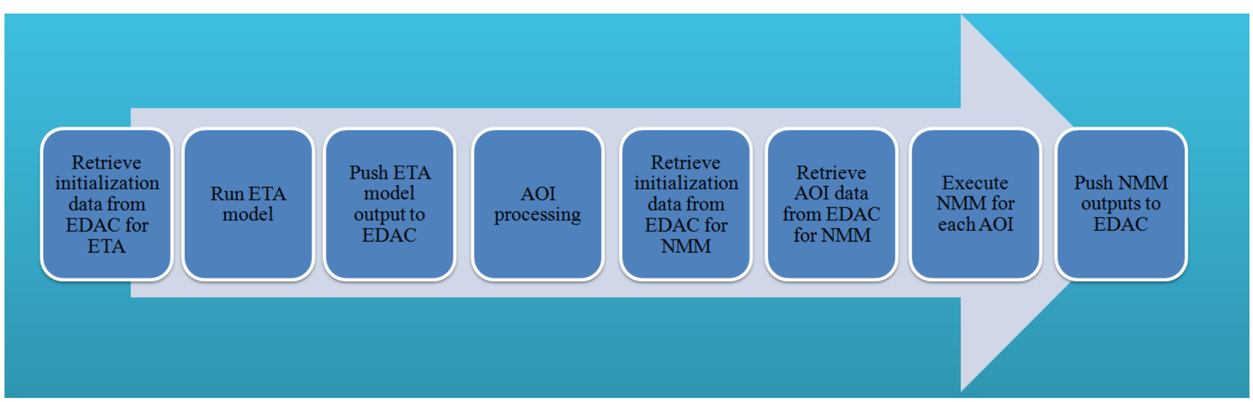



*Figure 19. Execution steps for nested running of DREAM ETA-8 and DREAM NMM*

When we utilize DREAM NMM model only to simulate the entire domain (36*27 degree), DREAM NMM model cannot support such a large domain size running with 3KM as spatial resolution due to high computing and memory consumption [6]. To support the full-domain runs, we would need to either a) redesign the existing algorithms, codes, and data structures, or b) significantly increase the speed of the CPU and, to a lesser extent, the network connection throughput. However, retrieving initialization data from EDAC would require about 2.5 minutes while the time required to run the NMM model for a domain size of 10 * 10 degrees would take about 12.7 hours with 8 CPU cores, suggesting that remote access of model initialization and boundary condition data is only a small fraction of the total time required to execute a model run. Theoretically, the computational cost of and explicit three-dimensional hydrodynamics weather forecasting model behaves like a function of $n^3$ for a given domain size, where n is a grid dimension [7]. If the domain size is doubled, 8-10 times more computing time [6] would be required to complete the model run. Therefore, a model run for the entire DREAM domain by the DREAM NMM model would be expected to finish more than 101 hours.

During the feasibility study, the comparison between the nested running of DREAM ETA and DREAM NMM and only utilizing DREAM NMM model to obtain the results for entire domain size (36*27 degree) have been conducted. Figure 19 shows all required steps to implement the proposed framework and enable the nested running of DREAM ETA-8 and DREAM NMM. To calculate the possibility of forecasting based on this approach, the time for the data transferring between the EDAC and GMU and the running of DREAM ETA have also been sampled and calculated as follows.

- Retrieval of initialization data from EDAC: 1.3 mins
- Execution of the DREAM ETA model (including local processing of WCS provided NetCDF files): 20 mins
- Transfer of DREAM ETA model outputs to EDAC: 4.5 mins
- Retrieval of initialization data from EDAC for DREAM NMM : 3.8 mins
- Retrieval of AOI data from EDAC for DREAM NMM: 0.5 mins
- Transfer of DREAM NMM outputs to EDAC: 1.5 mins per AOI

The execution time of DREAM NMM for a specific AOI depends on the AOI domain size. Figure 20 shows the AOI width and length distribution for the July 1, 2007 test case. It can be seen that most AOIs are within a 2x2 degrees sub-domain size for use in the execution of the DREAM NMM for each AOI. The largest domain size is 5.7x3 degrees. Figure 21 shows the execution time required for the different domain sizes defined for the test case. DREAM NMM can process the results within one hour for a model domain of 2 x 2 degrees, around 2 hours for a 4x4 degree model domain and about 4 hours for a 6 x 6 degree domain size. Therefore, it is expected that the entire collection of test case AOIs could be processed in 2-3 hours if all of the AOIs are simulated by the DREAM NMM in parallel, which would make the AOI with the largest domain size (5.7x3 degrees) determining the total execution time. Therefore, the total time for simulation including data transferring and execution time for both DREAM ETA-8 and DREAM NMM is 3 hours. In the test case reported here, the simulation was performed on a 1 Gbps network connection using Quad Core 16 GB computing servers. With more powerful computers with 10Gbps connections and six core 64 GB HPC computing nodes, the simulation can be performed in under one hour, therefore, the study found that the nested-model simulation is feasible for rapidly producing dust forecasts for a very large region (the Southwest U.S. in this case), using computational resources that are much more modest than those traditionally employed in generating large-scale forecasts such as those developed here.

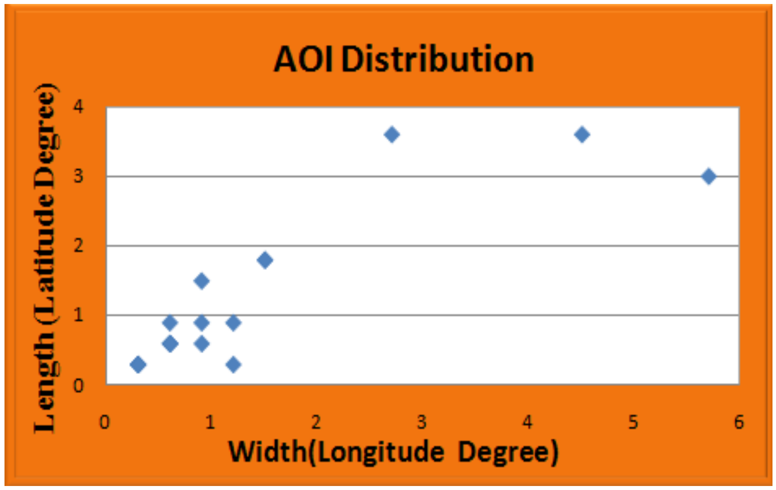



*Figure 20. AOI width and length distribution*

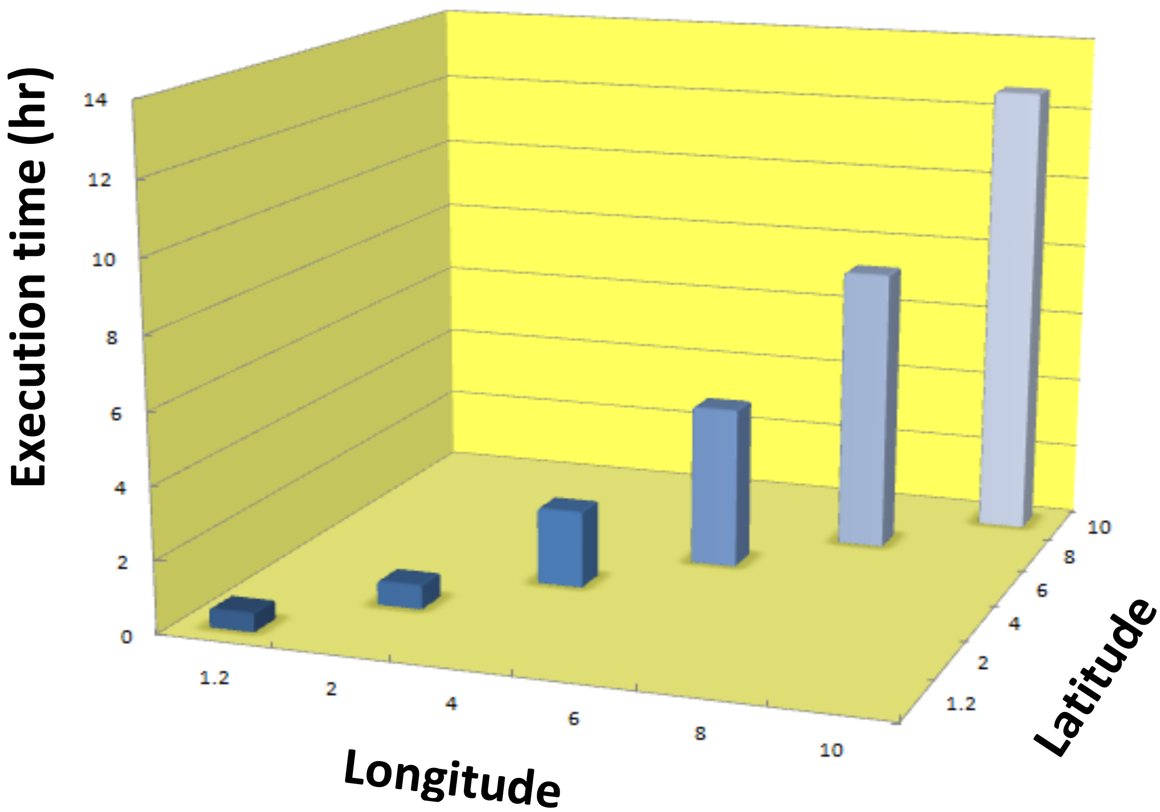



*Figure 21. DREAM NMM execution time*

Both the model results and performance comparison demonstrate that utilizing model interoperability technologies and high performance computing technologies are capable of enhancing dust storm forecasting by facilitating model integration, data discovery, data access, and data utilization for a) integrating widely distributed, high-volume, geospatial datasets as model inputs to enable model execution, b) reducing the computational time through focused execution of high-resolution (computationally expensive) models only for sub-regions for which there are indications of high dust concentrations, and c) strategically increasing spatial resolution and lengthening the period of forecast. Given long execution times for the DREAM NMM model, network latency for transfer of initialization parameters and outputs is a small fraction of total execution time, which is about 1-2 minutes for transfer data between two sites (UNM & GMU).

## Publications and presentations supported by and produced through this project

Xie, J., Yang, C., Zhou, B., and Huang, Q., 2010. High performance computing for the simulation of dust storms. Computers, Environment, and Urban Systems, 34(4):278-290.

Yang, C.,H.Wu, Q.Huang, Z.Li and J.Li. 2010, Spatial Computing for Supporting Physical Sciences, PNAS. (in review).

Benedict, Karl Kent, 2010. 'Delivery of Air Quality Model Data into New Mexico's Environmental Public Health Tracking System: A Services Oriented Architecture for Rapid Information Delivery'. Paper presented at the Annual Meeting of the American Thoracic Society. New Orleans, LA. May 19, 2010.

Benedict, Karl Kent, Phil Yang, and Qunying Huang, 2010. 'Environmental Model Interoperability Enabled by Open Geospatial Standards - Results of a Feasibility Study'. Paper presented at the American Geophysical Union Fall Meeting. San Francisco, CA. December 13-17, 2010.

Benedict, Karl Kent, Phil Yang, and Qunying Huang, 2010. ' 'OGC Web Coverage and Web Feature Services (WCS & WFS) for Enabling Nested Environmental Modeling'. Paper presented at the International Congress on Environmental Modelling and Software. Ottawa, Ontario, Canada. July 5, 2010.

Qunying Huang, Chaowei Yang, Jibo Xie, Huayi Wu, Jing Li,2009. Utilizing Model Interoperability and High Performance Computing to Enhance Dust Storm Simulation (AGU 2009 Fall Meeting, San Francisco, CA, Dec 14-18, 2009-12-30)

## References Cited


1. Gray, K., Jim Gray on eScience: A Transformed Scientific Method, in The Fourth Paradigm, Data-Intensive Scientific Discovery, T. Hey, S. Tansley, and K. Tolle, Editors. 2009, Microsoft Research: Redmond, Washington. p. xvii-xxxi.
2. Gettys, J., et al., eds. RFC 2616 – Hypertext Transfer Protocol – HTTP/1.1. 1999, Internet Engineering Task Force (IETF).
3. Fielding, R., Architectural Styles and the Design of Network-based Software Architectures, in Information and Computer Science. 2000, University of California, Irvine: Irvine, CA. p. 162.
4. Richardson, L. and S. Ruby, RESTful Web Services. 2007, Sebastopol, CA: O'Reilly Media, Inc. 446.
5. World Wide Web Consortium (W3C). SOAP Version 1.2 Part 1: Messaging Framework (Second Edition). 2007 [accessed 2008-06-27]; Available from: http://www.w3.org/TR/soap12-part1/.
6. Jibo Xie, Chaowei Yang, High Performance Computing Support for Environmental Models (AAG 2009 Annual Meeting, Las Vegas, NV, 3/25/09 at 8:00 AM)
7. Yang, C.,H.Wu, Q.Huang, Z.Li and J.Li. 2010, Spatial Computing for Supporting Physical Sciences, PNAS. (in review). 7. Baillie C, Michalakes J, Skilin R (1997) Regional weather modeling on parallel computers. Parallel Computing 23(14):2135-2142.