Multi-resolution Nested Dust Forecast System Feasibility Study

Karl Benedict (PI)
Earth Data Analysis Center, University of New Mexico
Chowei Yang (Co-I), Qunying Huang
Center for Intelligent Spatial Computing (CISC), George Mason University
Project Background

✦ Public Health Applications
  • Public Health Applications in Remote Sensing (PHAiRS - NASA REASoN): 2003-2008
  • Adding NASA Earth Science Results to EPHTN via the NM/EPHT System (ENPHASYS - NASA DECISIONS): 2008-2011

✦ Interoperability Development & Testing

Goals

✧ Work with existing modeling cores (DREAM ETA-8, DREAM NMM)
✧ Modify model pre- and post-processors to support OGC and REST data transfer
✧ Develop algorithm for automated generation of dust forecast area(s) of interest
✧ Evaluate and report on performance characteristics of the nested model system
Earth System Models
DREAM Eta 4-bin Model
DREAM Eta 8-bin Model
DREAM NMM

Earth Observations
Historic and Current
Terra/Aqua: MODIS
Land Cover, NDVI
SRTM C-Band Elevation Model

Planned Missions
NPP, NPOESS: VIIRS
Land Cover, NDVI

Observations, Parameters & Products

Decision Support System
SYRIS, NM EPHTS

Analyses
Linkage between PM$_{2.5}$ and PM$_{10}$ dust concentrations and measures of public health (i.e. hospital admissions, school absences, etc.)

Location and severity of recent dust events within the model domain

Decisions/Actions
Evaluation and treatment of patients in the context of additional dust concentration information available to clinicians
Alerts to sensitive populations in anticipation of important dust events

Inputs
NASA / UNM / U of A / GMU

Outputs

Outcomes
New Mexico DOH / CDC / Lubbock HD / Texas R1

Impacts

Value and Benefits to Society
Timely delivery of improved information about predicted dust events to public health officials, clinicians, and the general public

Enhanced understanding of the linkage between public health measures and modeled dust concentrations
Earth System Models
- DREAM Eta 4-bin Model
- DREAM Eta 8-bin Model
- DREAM NMM

Earth Observations
- Historic and Current
  - Terra/Aqua: MODIS
    - Land Cover, NDVI
  - SRTM C-Band Elevation Model
- Planned Missions
  - NPP, NPOESS: VIIRS
    - Land Cover, NDVI

Predictions/Forecasts
- Daily 72-hour low-resolution dust concentration forecasts for a large model domain (i.e. southwestern US)
- Daily 72-hour high-resolution dust forecasts for local regions (i.e. 1-degree blocks) for which the low-resolution model run indicates an important dust event

Decision Support System
SYRIS, NM EPHTS

Analyses
- Linkage between PM$_{2.5}$ and PM$_{10}$ dust concentrations and measures of public health (i.e. hospital admissions, school absences, etc.)
- Location and severity of recent dust events within the model domain

Decisions/Actions
- Evaluation and treatment of patients in the context of additional dust concentration information available to clinicians
- Alerts to sensitive populations in anticipation of important dust events

Value and Benefits to Society
- Timely delivery of improved information about predicted dust events to public health officials, clinicians, and the general public
- Enhanced understanding of the linkage between public health measures and modeled dust concentrations

Integrated System Solution Diagram

Project Focus Area

Inputs
- NASA / UNM / U of A / GMU

Outputs
- New Mexico DOH / CDC / Lubbock HD / Texas R1

Partnership Area

Outcomes

Impacts

NASA ROSES 2008 A.19: Multi-Resolution Nested Dust Forecast System Feasibility Study (NASA CAN NNX09AN53G)
Feasibility Testing

✦ Systems Integration
  ✓ Model pre- and post-processor implementation
  ✓ Data management and storage
  ✓ Appropriateness of implemented service standards

✧ Performance
  ✓ Comparison of performance (time-to-delivery) of nested model vs. dedicated large domain/high-resolution model runs
Systems Integration
What Do These Components Look Like?
OGC/HTTP/OPeNDAP – THREDDS (EDAC/GMU)

http://129.24.63.106:8080/thredds/catalog.html
OGC/HTTP/OPeNDAP – THREDDS (EDAC/GMU)

http://129.24.63.106:8080/thredds/catalog.html
OGC/HTTP/OPeNDAP – THREDDS (EDAC/GMU)
OGC/HTTP/OPeNDAP – THREDDS (EDAC/GMU)

Catalog /thredds/catalog.html

Dataset

Dataset: NCEP ETA DREAM Forecast Data

- Data type: GRID
- ID: ETADatasetScan/dust2d.20070701.nc

Access:

1. OPENDAP: /thredds/dodsC/dream_eta/dust2d.20070701.nc
2. HTTPServer: /thredds/fileServer/dream_eta/dust2d.20070701.nc
3. WCS: /thredds/wcs/dream_eta/dust2d.20070701.nc
4. WMS: /thredds/wms/dream_eta/dust2d.20070701.nc

Dates:

- 2010-09-21 23:49:19Z (modified)
OGC/HTTP/OPeNDAP – THREDDS (EDAC/GMU)
OGC/HTTP/OPeNDAP – THREDDS (EDAC/GMU)
OGC/HTTP/OPeNDAP – THREDDS (EDAC/GMU)
REST Data Upload Services (EDAC/GMU)
DREAM ETA-8 Model

2007-07-01
Identified Areas of Interest
Identified Areas of Interest

Matrix Processing of NetCDF Model Output
Identified Areas of Interest

Matrix Processing of NetCDF Model Output

Generation of AOI
Raster mask (GeoTiff)

2007-07-01, 72-hour forecast, 3-hour time step, dl (dust loading, gm/m²), $1 \times 10^{-7}$ base threshold, 8 aggregate threshold
Identified Areas of Interest

Matrix Processing of NetCDF Model Output

Generation of AOI
Raster mask (GeoTiff)

Vectorization of discrete AOIs from raster

2007-07-01, 72-hour forecast, 3-hour time step, dl (dust loading, gm/m²), 1*10^{-7} base threshold, 8 aggregate threshold
Identified Areas of Interest

Matrix Processing of NetCDF Model Output

Generation of AOI
Raster mask (GeoTiff)

Vectorization of discreet AOIs from raster

2007-07-01, 72-hour forecast, 3-hour time step, dl (dust loading, gm/m²), 1*10⁻⁷ base threshold, 8 aggregate threshold
Identified Areas of Interest

Matrix Processing of NetCDF Model Output

Generation of AOI
Raster mask (GeoTiff)

Vectorization of discreet AOIs from raster

Generation of AOI vector files (GeoJSON, KML)

2007-07-01, 72-hour forecast, 3-hour time step, dl (dust loading, gm/m^2), 1*10^-7 base threshold, 8 aggregate threshold
Identified Areas of Interest

Matrix Processing of NetCDF Model Output

Generation of AOI Raster mask (GeoTiff)

Vectorization of discreet AOIs from raster

Generation of AOI vector files (GeoJSON, KML)

2007-07-01, 72-hour forecast, 3-hour time step, dl (dust loading, gm/m²), $1 \times 10^{-7}$ base threshold, 8 aggregate threshold
Identified Areas of Interest

Matrix Processing of NetCDF Model Output

Generation of AOI
Raster mask (GeoTiff)

Vectorization of discreet AOIs from raster

Generation of AOI vector files (GeoJSON, KML)

Publication of vector files via HTTP
AOI Access Client (GeoJson/KML)
NMM Model

NMM-dust on AOI 10, 11, 12 & 13 at 00:00 UTC 02-July-2007
Performance

AOIs Distribution

Graph showing the distribution of AOIs with length and width dimensions.

Graph showing the execution time (hours) against AOI.
## Performance Assessment

<table>
<thead>
<tr>
<th>Data Transfer</th>
<th>Time (HH:MM:SS)</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retrieval of Global Forecast System (GFS) parameters from EDAC servers</td>
<td>00:01:32</td>
<td>Initialization and boundary conditions for ETA-8bin model run</td>
</tr>
<tr>
<td>Delivery of ETA-8bin model results to EDAC servers</td>
<td>00:04:37</td>
<td>REST-based transfer</td>
</tr>
<tr>
<td>Retrieval of ETA-8bin and AOI data from EDAC servers</td>
<td>00:03:50</td>
<td>Time required for the retrieval of the required initialization and boundary condition data for a single AOI NMM-dust model run</td>
</tr>
<tr>
<td>Delivery of NMM-dust model results to EDAC servers</td>
<td>00:01:30</td>
<td>REST-based transfer</td>
</tr>
<tr>
<td></td>
<td>Total: 00:11:29</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model Runs</th>
<th>Time (HH:MM:SS)</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>ETA-8bin</td>
<td>0:20:00</td>
<td>Full 37 x20 degree model domain at a ~50 km resolution</td>
</tr>
<tr>
<td>NMM-dust</td>
<td>min: 0:27:00</td>
<td>Summary statistics for 18 NMM-dust runs for the AOIs identified following the 1-July-2011 dust forecast at 3km x 3km resolution.</td>
</tr>
<tr>
<td></td>
<td>max: 1:24:00</td>
<td>Estimated full-domain execution time: 92.3 hours.</td>
</tr>
<tr>
<td></td>
<td>Mean: 0:39:00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>median: 0:32:00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1st Quartile: 0:30:36</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3rd Quartile: 0:39:36</td>
<td></td>
</tr>
<tr>
<td></td>
<td>90th Percentile: 1:04:55</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n: 18</td>
<td></td>
</tr>
</tbody>
</table>

Total Parallel Execution Time: 1:55:29 (transfer total + ETA-8-bin + MAX NMM-dust)
Feasibility Analysis

✦ Systems Integration

✓ Model pre- and post-processor implementation
  • Relatively straightforward process
  • Challenge posed by models that require re-compilation to change model domain or other execution parameters

✓ Data management and storage
  • Simple file system approach works well
  • Separation of model execution from file storage allows for optimization for modeling independent of storage capacity
  • Need to develop more structured data management system (i.e. data registry & management utilities) in move towards operationalization

✓ Appropriateness of implemented service standards
  • **WCS** is very effective in supporting data subsetting prior to network transfer (i.e. parameter extraction from large model products)
  • **WMS** remains useful for quick visualization of products in a variety of platforms, but for the modeling activity is not key
  • For this application, **WFS** seems like overkill, simple HTTP access to GeoJSON data files is sufficient for delivery of AOIs in a compact data model to remote systems
  • **REST** exchange remains useful for flexible transfer of data products between systems where the OGC services don’t have a standard request-response model
Feasibility Analysis

✦ Performance

✓ Different geographic resolution will require different time to simulate.
  • Doubling of resolution (e.g., 4x4->2x2) requires 8-10 times more computing time

✓ NMM model execution at high resolution (3km) remains computationally and time intensive, but smaller domains are more feasible
  • Different domain sizes require different execution time.
  • Most sub-domain sizes are within 2x2 degrees, which can be processed within one hour for 3X3 km²

✓ Given long execution times for NMM model, network latency for transfer of initialization parameters and outputs is a small fraction of total execution time.
  • About 1-2 minutes/transfer between two sites (UNM & GMU), <12 minutes total transfer time.

✓ More dynamic parameters, such as soil moisture, could be assimilated into the model and implementation of this additional modeling capacity would require additional computing power (potentially provided by cloud computing)
Potential Follow-on / Related Opportunities

✧ Operationalization
  • Develop AOI generation service
  • Develop processing queue that tracks which AOIs have been processed for publication of un-executed AOIs as a feed (e.g. ATOM or RSS)
  • Enable time-enabled WMS for more efficient access and use
  • Auto-mosaic/overlay of time-enabled WMS for low- and high-resolution model outputs within a single service

✧ Integration with other modeling systems.
  • Soil Moisture from Hydrologic Models for Model Initialization (NASA EPSCoR Proof of Concept Project)
  • Community Multi-Scale Air Quality (CMAQ) model (ENPHASyS Project)

✧ Extension of on-demand high-resolution model execution into public/private cloud

✧ Automated air-quality alerts based upon AOI system
Contact Information

Karl Benedict
kbene@edac.unm.edu
(505) 277-3622 x 234
MSC01 1110, 1 University of New Mexico
Albuquerque, NM 87131
http://edac.unm.edu

Chowei Yang
cyang3@gmu.edu
GMU/CISC, MS 6A2
4400 Univ. Dr., Fairfax, VA, 22030-4444
http://www.cisc.gmu.edu/