GRAVITE Support for NOAA-20 VIIRS SDR Reprocessing

Wayne McCullough
wayne.mccullough@noaa.gov
DPES Sustainment Team
Overview

- Background
  - GRAVITE
  - PGEs
  - ADL

- Request
  - NOAA 20 VIIRS had unexpected conditions after launch
  - SDR team wanted to reprocess with new LUTs

- Process
  - Strong user interaction

- Results
  - First run
  - Second run

- Conclusion
GRAVITE IPS

- GRAVITE has a lot of data
  - 112 Million Files, 91 Million unique granules (SNPP and NOAA-20)
  - All RDRs since launch, 34 day rolling storage of other XDRs
  - 560 TB of data

- GRAVITE has available resources
  - Computer
    - Workstations (at GSFC L40)
    - ICF Servers: dedicated to remote access and compute tasks
    - PGE Servers: dedicated to automated processing
    - All servers have direct access to data
    - 1.1 PB of dedicated disk space for operational system
  - Tools
    - IDL, Matlab, Python, Redmine, PGEs, etc.
  - Support
    - Operators, Developers, Engineers, etc.

- It is there for the JPSS Cal-Val & Data Quality community to use
• What is a PGE?
  – Product Generation Executable
  – Any code we automatically run against data for time periods
    • E.g., Run this analysis every hour when the data is available, etc.

• Initially used heavily for ground comparison

• Broader use now:
  – Instrument DQ Checks
  – Static plot generation
  – Data Preview Tile sets
  – Granulated Ancillary generation (uses ADL)
  – Reprocessing
• Conditions for running a PGE:
  – Have rules defining the time periods (Execution Blocks)
    • Hourly, daily, etc.
    • Orbit
    • Custom lookup
  – Define input products
    • Input products may be optional or required.
    • A minimum number or maximum time gap may be set.
    • A geo-spatial area may be set.
    • E.g. “If I have full coverage for VIIRS M7, the cloud mask, and the GEOs, run xxx for this hour”
  – Automatically run PGE when conditions are met
    • Queue task
    • Execute code on available node
  – Selected Outputs are archived by GRAVITE
• Simplified reflection of IDPS architecture:
  – Processing Subsystem (PRO)
  – Data Management Subsystem (DMS)
  – No Ingest Subsystem (ING), No Data Delivery Subsystem (DDS), No Infrastructure Subsystem (INF)
  • Some functions replaced by ADL Toolkit
• STAR VIIRS SDR team needed to regenerate SDRs from RDRs
  – Unexpected conditions shortly after launch
  – New Lookup Tables needed
  – Wanted to reprocess all NOAA-20 VIIRS Science RDRs (from launch November 2017 to end of February 2018)

• February 2018 DPES and STAR VIIRS SDR team met
  – Various approaches considered
    • All centered on many runs of ADL
  – SDR team needed more time to finalize LUTs
  – DPES team needed more time to test and refine ADL calls

• Goal:

  Start Processing by August 2018
● DPES dedicated three computers in GRAVITE for run
  – Each machine 24 core, 256 GB RAM
    • Dell PowerEdge R430 servers, with two Intel Xeon E5-2680v3 CPUs at 2.5GHz, eight 32GB RDIMM with Advanced ECC, and two Intel Ethernet X540 DP 10GBASE-T
  – Each machine to run a max of 16 ADL processes

● Set up ADL to run in a PGE
  – STAR VIIRS SDR team provided significant support:
    • patch to ADL to turn on compression
    • Testing and reference data
    • Final LUT package for reprocessing

● Runs as part of operational PGE system
  – No impact to current PGEs, only minor configuration changes needed
  – Reprocessing PGE delivered in GRAVITE v4.4
● First Run start 2018-07-05
● Finish 2018-07-17
● VIIRS SDR team noted that about 1% of outputs were missing
  – ADL was not called properly to handle A2 Granules
  – A fix was developed
● Decided to re-run everything
  – Ensure all data was correctly processed.
  – Avoid duplicates.
• Second Run start 2018-08-16 19:22
• Finish 2018-08-27 19:01
• Performance: ~85 days of data processed in 11 days
  – Average time to run each hour of data: 6 hours
  – Run 48 simultaneously across three computers
  – Net: ~7.8x faster than real time
  – If needed, we could parallelize it more
• Outputs: ~42 TB of data
  – Segmented into the gvo domain. (i.e., Not ops, pop, or int)
  – DPES will keep for 1 year
  – Available for all GRAVITE users
Time to Run ADL
Conclusion

- **GRAVITE** can use ADL to reprocess large amounts of data
  - ADL is a complex utility
  - Requires a bit of trial and error

- **GRAVITE IPS PGE** system can support large reprocessing runs
  - This is the first run of this nature we have tried
  - Overall, the GRAVITE IPS system did what it is supposed to do

- **DPES** can support JPSS Reprocessing via ADL
  - ADL available on ICF machines
  - Talk with DPES for larger runs
Contacts

• To subscribe to DQA alerts, contact:
  – ops-gravite-dpes-jpss@lists.nasa.gov
  (Subscribers need to have a GRAVITE account)
• New GRAVITE account request, contact:
  – Erica Handleman: erica.handleman@nasa.gov
• System access issues, contact:
  – gravite.service@noaa.gov
• DQA functions, contact:
  – dqst-dpes-jpss@lists.nasa.gov
• All other issues, contact:
  – ops-gravite-dpes-jpss@lists.nasa.gov
CRTM and Data Assimilation activities at STAR supporting JPSS

Kevin Garrett$^1$

CRTM Federal Manager
NESDIS Associate Director for the Joint Center for Satellite Data Assimilation

Tong Zhu$^{1,2}$, Ming Chen$^{1,3}$, Biljana Orescanin$^{1,3}$, Yingtao Ma$^4$

Acknowledgments: Jean-Luc Moncet$^4$, Mark Liu$^1$, Benjamin Johnson$^5$, Hui Shao$^5$

1: NOAA/NESDIS/STAR  2: CIRA  3: CICS  4: AER, Inc.  5: UCAR/JCSDA
Outline

- **CRTM**
  - Status
  - Cal/Val and algorithm support
  - Current and future development

- **Data Assimilation**
  - Current activities
  - STAR plans/priorities
CRTM activities
## CRTM development history
### Impacting JPSS Applications

<table>
<thead>
<tr>
<th>CRTM Version</th>
<th>Date</th>
<th>Enhancements</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.0/2.0.5</td>
<td>12/2011</td>
<td>• New user interface</td>
</tr>
</tbody>
</table>
| 2.1/2.1.1    | 3/2012   | • SOI solver  
• Fastem5  
• MW Land Surface Emissivity Model  
• NLTE Correction |
| 2.1.3        | 6/2013   | • Implement reflection correction in Fastem (use clear-sky trx)  
• Enhanced absorption coefficients (6 absorbers)  
• Solar irradiance in spectral coefficient files (CrIS)  
• IRSSEM improvements |
| 2.2.1        | 4/2015   | • Enable reflection correction for non-scattering clouds  
• Fastem6  
• Revert to box car SRF for SNPP ATMS |
| 2.2.3        | 8/2015   | • IRRSEM improvements |
| 2.3 (current)| 11/2017  | • NOAA-20 coefficients  
• ATMS snow and sea-ice emissivity models  
• Cloud fraction capability  
• Reflection correction (use cloudy trx) |
## CRTM Cal/Val and algorithm support

### Applications applied to JPSS data

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>CRTM v.</th>
<th>Current use</th>
<th>Some desired enhancements?</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICVS</td>
<td>2.0.5-2.3</td>
<td>Forward operator, clear-sky, ocean</td>
<td>Ocean emissivity/reflectance modeling</td>
</tr>
<tr>
<td>MiRS</td>
<td>2.1.1</td>
<td>Forward operator, K-matrix, all-sky variational retrieval</td>
<td>Hydrometeor handling (scattering properties)</td>
</tr>
<tr>
<td>ACSPO</td>
<td>2.1.3</td>
<td>Forward operator, clear-sky, ocean</td>
<td>IRSSEM, reflectance enhancements, aerosol handling (species, scattering)</td>
</tr>
<tr>
<td>Enterprise Cloud Products</td>
<td>2.1.3</td>
<td>Long-wave IR clear-sky transmittance profiles</td>
<td>Shortwave IR transmittance, cloudy transmittances</td>
</tr>
<tr>
<td>Enterprise Volcanic Ash</td>
<td>2.1.3</td>
<td>Long-wave IR clear-sky transmittance profiles</td>
<td>Shortwave IR transmittance, cloudy transmittances</td>
</tr>
</tbody>
</table>

*All applications could benefit from improved efficiency*
• JCSDA partners collectively manage CRTM development (B. Johnson lead)

• STAR-led contributions to JCSDA CRTM project
  • Code management, new sensors, testing & maintenance, package/delivery of software
  • Surface emissivity modeling, BRDF improvements (CSEM)
  • Modernization of LBLRTM with through the Community Line-By-Line Model (CLBLM)
  • Extension to UV sensors

• Summary of other JCSDA projects
  • Fast solvers for scattering
  • Full Stokes polarization
  • Improvements to aerosol/hydrometeor scattering properties/LUTs
  • Improved code efficiency (vectorization/OpenMP)

• Next release is v3.0 ~Jan/Feb 2019
Objective: Release of the CSEM package and integration into CRTM

- CSEM is OOP-based system to compute emissivity and BRDF over all surfaces, in the MW, IR and Vis
  - Easy to integrate and test new emissivity models
  - Easy to interface with other tools (e.g. CRTM)
  - Includes tangent-linear and adjoint

Enhancements over existing CRTM surface emissivity models

<table>
<thead>
<tr>
<th>Microwave</th>
<th>Vis/IR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Improved NESDIS Land Phys. Model</td>
<td>UW IR Emissivity Atlas (SEEBOR)</td>
</tr>
<tr>
<td>Semi-physical ATMS Snow Model</td>
<td>UW Vis/NIR BRDF Atlas (Vidot &amp; Borbas)</td>
</tr>
<tr>
<td>Semi-Physical ATMS Sea-ice Model</td>
<td></td>
</tr>
<tr>
<td>TELSEM 1, 2 (climatology)</td>
<td></td>
</tr>
</tbody>
</table>

CSEM can be used as a stand-alone package or interface with other tools
CRTM current developments (CSEM)
The Community Surface Emissivity Model (CSEM)

O-B over land using NESDIS Land Physical Model
(TOP: CRTM 2.3) (Bottom: CSEM)

REL 2.3
23.8GHz

REL 2.3
50.3GHz

TELSEM

CSEM

O-B over sea-ice for 50.3 GHz
(TOP: TELSEM) (Bottom: CSEM)

CSEM

Demonstration of CSEM improvements
Objective: Development/release of the Community Line-By-Line Model (CLBLM)

- Monochromatic RTM to train CRTM fast model
- Modernization of heritage LBLRTM
  - Refactored/modular code
  - Improved I/O
  - Redesigned RT/Jacobian routines
  - Double line-shape convolution scheme for improved narrow-lines
- CLBLM Alpha released 1/2018
- CLBLM v1.0 released 8/2018
Data assimilation activities
Objective: Increase the number and quality of ATMS surface-sensitive (non-ocean) observations assimilated (NOAA GDAS/GFS)

• Requires accurate forward operator
• ...which requires accurate surface characterization (e.g. emissivity)

Implement in 2 phases

• Improve the background surface emissivity
• Implement surface emissivity as a control variable in the GSI

Compare Current Land Model in CRTM and TELSEM2 for background

<table>
<thead>
<tr>
<th></th>
<th>CRTM</th>
<th>TELSEM 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surface type</td>
<td>All</td>
<td>Land &amp; sea-ice only</td>
</tr>
<tr>
<td>Frequency</td>
<td>3 – 190 GHz</td>
<td>10 – 700 GHz</td>
</tr>
<tr>
<td>Polarization</td>
<td>H + V</td>
<td>H + V</td>
</tr>
<tr>
<td>Spatial Resolution</td>
<td>0.25°</td>
<td>0.25°</td>
</tr>
<tr>
<td>Temporal Resolution</td>
<td>Instantaneous</td>
<td>Monthly</td>
</tr>
<tr>
<td>Base</td>
<td>“Physical”</td>
<td>Empirical</td>
</tr>
</tbody>
</table>

CSEM improved Land Emissivity physical model will also be tested
Data assimilation activities

ATMS surface-sensitive radiance assimilation

Improving the background: 31 GHz Emissivity from 2 GDAS Cycles

Replacing the background to use TELSEM2 increases x2 the number of observations assimilated (from 2 GDAS cycles)
Data assimilation activities
ATMS surface-sensitive radiance assimilation

Improving the analysis: O-B, O-A, and Obs Count from 9 GDAS Cycles

While replacing the background (orange) improves the observation count, implementing emissivity as control variable improves the analysis.

Further improvement can be realized:
- Use off-diagonal elements of emissivity covariance matrix
- Improve bias correction over land
Data assimilation/CRTM activities

CrIS and IASI shortwave IR 4 µm band assimilation
(Boukabara, Ide, Garrett, Barnet)

- Assess CRTM capability
  - NLTE, shortwave reflectance
- Extend global DA
  - Dynamic CO2/N2O, obs errors, etc.
- Assess analysis and forecast impact vs longwave IR

Other efforts at STAR

IASI O-B without NLTE (top) and with NLTE (bottom)

Chen et al. 2013

Improve PMW all-sky radiance retrieval/assimilation

- Utilize datasets like GPM 2BCSATGPM
  - Quantify accuracy of CRTM in precip for ATMS
- ... or GPROF
  - Training set to improve a-priori of hydrometeor microphysical properties
Cal/Val Systems and Science Support from CRTM

- Address priorities and needs of STAR EDR, Cal/Val teams
  - Science needs, e.g. improvements to quality of output
  - Technical needs, e.g. supporting transitions to new versions

Science/Coordination Support for Data Assimilation

- Address priorities across STAR, NESDIS (program offices)
  - Assimilation of land EDRs (LST, GVF, soil moisture)
  - Assimilation of ocean EDRs (SST, color)
  - Assimilation of cryospheric products (IST, SIC, Snow Cover/SWE)
  - Assimilation of trace gases, aerosol (V8Tot/Pro, AOD)
CLASS Access and Future Trends for S-NPP and JPSS Data

Brent Hefner, CLASS Program Manager (Acting)
Alan Hall, CLASS System Owner & Operations Manager

2018 STAR JPSS Annual Science Team Meeting
August 28, 2018
Overview-
Comprehensive Large Array data Stewardship System (CLASS)

- CLASS provides long-term, secure storage of NOAA-approved data, information, and metadata and to enable access to these holdings through both human and machine-to-machine (M2M) interfaces
- CLASS is not intended to support near-real-time nor mission-critical product delivery
- CLASS follows the concepts defined in the Open Archival Information System Reference Model (OAIS-RM)
- CLASS Development phase completed on June 30, 2017
- CLASS has transitioned to Sustainment:
  - Minor problem resolution and enhancements are delivered through regularly scheduled Sustainment Software Releases
  - Software Releases are scheduled once per quarter

CLASS is fully operational and meeting all performance objectives
CLASS within the NESDIS Ground Enterprise
S-NPP and J1 Archive & Dissemination Metrics

• Recent Archive metrics

<table>
<thead>
<tr>
<th></th>
<th>May</th>
<th></th>
<th></th>
<th>June</th>
<th></th>
<th></th>
<th>July</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. of Files</td>
<td>Volume (TB)</td>
<td>No. of Files</td>
<td>Volume (TB)</td>
<td>No. of Files</td>
<td>Volume (TB)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S-NPP</td>
<td>965,298</td>
<td>40.5</td>
<td>775,771</td>
<td>40.14</td>
<td>805,603</td>
<td>41.33</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NOAA-20</td>
<td>877,245</td>
<td>39.1</td>
<td>724,863</td>
<td>38.17</td>
<td>743,578</td>
<td>39.1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

• July Dissemination metrics

<table>
<thead>
<tr>
<th></th>
<th>July</th>
</tr>
</thead>
<tbody>
<tr>
<td>S-NPP</td>
<td></td>
</tr>
<tr>
<td>Subscriptions</td>
<td>1,852,940</td>
</tr>
<tr>
<td>Ad-Hoc</td>
<td>1,172,499</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>July</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOAA-20</td>
<td></td>
</tr>
<tr>
<td>Subscriptions</td>
<td>1,047,550</td>
</tr>
<tr>
<td>Ad-Hoc</td>
<td>548,604</td>
</tr>
</tbody>
</table>
CLASS Website

- Provides access to CLASS information holdings
- Requires registration to order data
- www.class.noaa.gov
- Manage Subscription orders
- Place Ad-Hoc Orders
## CLASS Access Services

<table>
<thead>
<tr>
<th>Order types</th>
<th>Avg Time to Available</th>
<th>File Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Subscription</strong></td>
<td>&lt; 6-7 hours</td>
<td>No limit</td>
</tr>
<tr>
<td>(Standing orders)</td>
<td>(As little as 45m depending on data)</td>
<td></td>
</tr>
<tr>
<td><strong>NPP-FTP</strong></td>
<td>&lt; 6-7 hours</td>
<td>No limit</td>
</tr>
<tr>
<td>(Rolling ~90 days)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Ad-Hoc</strong></td>
<td>~24 hours</td>
<td>Up to 3,000 files</td>
</tr>
<tr>
<td>(Historical/older data)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Bulk</strong></td>
<td>24-48 hours</td>
<td>3000 to 6000</td>
</tr>
<tr>
<td>(Large/non-typical)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### NOTES:

Some data types are available as fast as can be processed, others are delayed by the Program (i.e. JPSS).

CLASS is a tape Library System, retrieval times can vary. Operators work with large data requests to facilitate best delivery.
Subscription Orders

• Subscriptions are standing orders for newly archived data which are fulfilled automatically
• Users can manage Subscription orders via the CLASS website
Ad-Hoc Orders

- Ad-Hoc orders can be placed through the CLASS website
- Data is grouped into product families which can be searched
NPP and J1 Rolling FTP Directory

- Located at ftp://ftp-npp.bou.class.noaa.gov/
- Easy to navigate directory of recently archived NPP and J1 data
- ~90 day rolling window

Index of /20180817/VIIRS-EDR/VIIRS-Near-Constant-Contrast-NCC-EDR-GTM-Geo/J01/

<table>
<thead>
<tr>
<th>Name</th>
<th>Size</th>
<th>Date Modified</th>
</tr>
</thead>
<tbody>
<tr>
<td>VIIRS-EDR_VIIRS-Near-Constant-Contrast-NCC-EDR-GTM-Geo_20180817_28935.tar</td>
<td>1.9 GB</td>
<td>8/17/18, 2:00:00 AM</td>
</tr>
<tr>
<td>VIIRS-EDR_VIIRS-Near-Constant-Contrast-NCC-EDR-GTM-Geo_20180817_28935.tar.manifest.xml</td>
<td>6.4 kB</td>
<td>8/17/18, 2:00:00 AM</td>
</tr>
<tr>
<td>VIIRS-EDR_VIIRS-Near-Constant-Contrast-NCC-EDR-GTM-Geo_20180817_29997.tar</td>
<td>1.9 GB</td>
<td>8/17/18, 3:33:00 AM</td>
</tr>
<tr>
<td>VIIRS-EDR_VIIRS-Near-Constant-Contrast-NCC-EDR-GTM-Geo_20180817_29997.tar.manifest.xml</td>
<td>6.4 kB</td>
<td>8/17/18, 3:33:00 AM</td>
</tr>
<tr>
<td>VIIRS-EDR_VIIRS-Near-Constant-Contrast-NCC-EDR-GTM-Geo_20180817_30415.tar</td>
<td>1.9 GB</td>
<td>8/17/18, 6:03:00 AM</td>
</tr>
<tr>
<td>VIIRS-EDR_VIIRS-Near-Constant-Contrast-NCC-EDR-GTM-Geo_20180817_30415.tar.manifest.xml</td>
<td>6.4 kB</td>
<td>8/17/18, 6:03:00 AM</td>
</tr>
<tr>
<td>VIIRS-EDR_VIIRS-Near-Constant-Contrast-NCC-EDR-GTM-Geo_20180817_31155.tar</td>
<td>1.9 GB</td>
<td>8/17/18, 7:01:00 AM</td>
</tr>
<tr>
<td>VIIRS-EDR_VIIRS-Near-Constant-Contrast-NCC-EDR-GTM-Geo_20180817_31155.tar.manifest.xml</td>
<td>6.4 kB</td>
<td>8/17/18, 7:01:00 AM</td>
</tr>
<tr>
<td>VIIRS-EDR_VIIRS-Near-Constant-Contrast-NCC-EDR-GTM-Geo_20180817_31555.tar</td>
<td>1.9 GB</td>
<td>8/17/18, 8:23:00 AM</td>
</tr>
<tr>
<td>VIIRS-EDR_VIIRS-Near-Constant-Contrast-NCC-EDR-GTM-Geo_20180817_31555.tar.manifest.xml</td>
<td>6.4 kB</td>
<td>8/17/18, 8:23:00 AM</td>
</tr>
<tr>
<td>VIIRS-EDR_VIIRS-Near-Constant-Contrast-NCC-EDR-GTM-Geo_20180817_32095.tar</td>
<td>1.9 GB</td>
<td>8/17/18, 9:31:00 AM</td>
</tr>
<tr>
<td>VIIRS-EDR_VIIRS-Near-Constant-Contrast-NCC-EDR-GTM-Geo_20180817_32095.tar.manifest.xml</td>
<td>6.4 kB</td>
<td>8/17/18, 9:31:00 AM</td>
</tr>
<tr>
<td>VIIRS-EDR_VIIRS-Near-Constant-Contrast-NCC-EDR-GTM-Geo_20180817_32595.tar</td>
<td>1.9 GB</td>
<td>8/17/18, 11:02:00 AM</td>
</tr>
<tr>
<td>VIIRS-EDR_VIIRS-Near-Constant-Contrast-NCC-EDR-GTM-Geo_20180817_32595.tar.manifest.xml</td>
<td>6.4 kB</td>
<td>8/17/18, 11:02:00 AM</td>
</tr>
</tbody>
</table>
Machine to Machine (M2M) Interface

- M2M API was designed for the purpose of enabling software developers to create access clients capable of searching for and ordering datasets held within CLASS
- NCEI has implemented M2M Clients
M2M Next Steps

• CLASS is planning an Engineering Assessment to determine the feasibility of extending the M2M interface to STAR
  – Performance impact analysis
  – Cost estimate
Summary

• CLASS is the archive for NPP and J1 data
• Multiple options exist for ordering data from CLASS
• CLASS is investigating making the M2M interface available to STAR