



Roadmap to the 2017 GHRC UWG Meeting

2016 Recommendations Response

Helen Conover

2017 GHRC User Working Group Meeting
Sept 26-27, 2017



Recommendations Recap – Highlights

Recommendation 1: Establishing relationships with PIs and Program Managers to get access to invited *program-level meetings and relevant data workshops.*

Recommendation 2: Further develop 5-10 year plan. The GHRC is in a unique position to be a data provider for “Storm Induced Hazards”, which already includes the following major categories: lightning, hurricanes and floods.

- a) Become *known within the community as the go-to place for datasets related to these subjects.*
- b) Clearly and succinctly *demonstrate the linkages or traceability of current/proposed future data holdings to the GHRC mission components.*
- c) Determine a set of *useful user metrics*, with feedback obtained from the UWG, which can be routinely updated and made available to the NASA sponsor, UWG and broader community as a page on the GHRC site that displays user metrics, hot datasets, etc.

Recommendation 3: Lifecycle plans and workflows should be moved up a level (e.g., Resources instead of Resources>Documents), to *provide clarity on the process for investigators interested in providing data.*

Recommendation 4: *Publicize ISS LIS datasets*, and make technical specifications of the data products and latency available to potential users. Also provide *virtual lightning data holdings*, such as GLM data at CLASS, or other future lightning datasets from other countries.

Recommendations – Highlights

Recommendation 5: Develop a single tool that can provide broad use to multiple field campaigns and data types. For example, *apply FCX to other field programs* serviced by the GHRC (e.g., the numerous GPM field campaign datasets it archives).

Recommendation 6: Create *data bundles* for scientists who want to study processes. Use the UWG as a sounding board for *micro articles* that the GHRC feels would be interesting for publication on the website. Research the idea of using an automatic program to do data bundling.

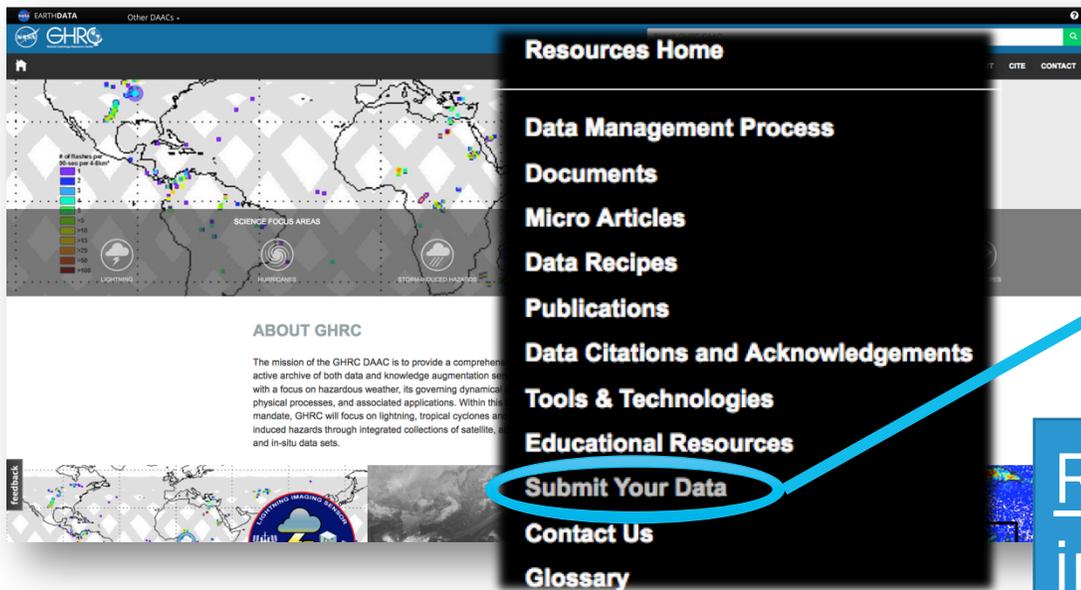
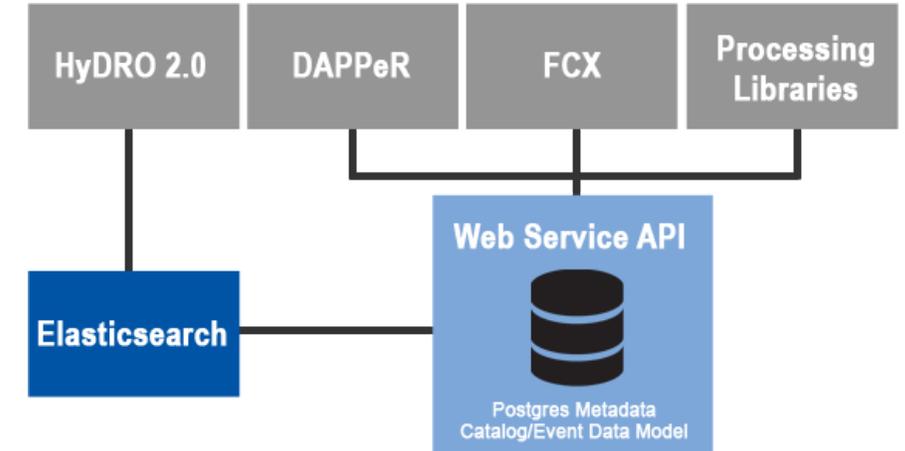
Recommendation 7: Develop a plan to *obtain the targeted SWOT datasets* to complement hazardous weather related to *floods caused by excess precipitation*.

Recommendation 8: *Include GOES GLM data and reference data* collected in the GOES-R Field Campaign in the GHRC portfolio of accessible data whether stored in house or as a virtual data set since NOAA is funded for stewardship of these datasets. Functionality should be seamless with other holdings. GHRC should continue planning to be the first stop for accessing global lightning data. GHRC should *coordinate with CLASS so the data are viewable at GHRC* and point to CLASS, perhaps providing readers for GHRC user communities.

Session 1: New Data Publication Workflow

Infrastructure updates to support more efficient DAAC operations and better service to our users

- Database migration and metadata review
- Improved HyDRO search tool
- New data publication workflow tool - DAPPeR



Recommendation 3 - clarity for investigators interested in providing data

Session 2: ISS LIS / Lightning

Lightning Imaging Sensor (LIS) on the International Space Station (ISS)

- February 19, 2017 - SpaceX 10 with STP-H5 (Space Test Program-Houston 5) launched
- February 28, 2017 - LIS instrument power-up and checkout

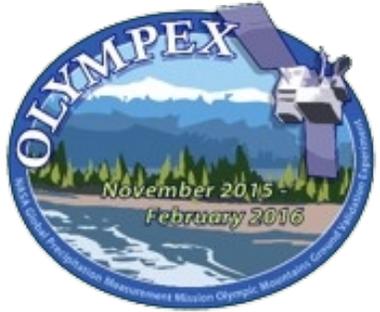


Recommendations 2a, 4, 8

- “Go to” place for lightning data
- Publicize ISS LIS datasets
- Include GOES GLM and reference data in portfolio

Related Work: UAH ITSC (some GHRC staff) are working with the GOES-R and Geostationary Lightning Mapper (GLM) Science Teams to provide data management and access

- *GLM Cal/Val Data Portal*
 - Provide lightning team access to GLM Level 0, Level 1 / Level 2, and many reference data sets from a single web site
 - Target audience: GLM Cal/Val community (includes EUMETSAT, InPE, Environment Canada, and University partners)
- *GOES-R Post Launch Test Field Campaign Portal*
 - Field Campaign coordination and information management (plans of the day, weather forecasts, mission reports)
 - Post-campaign data acquisition, documentation, and management
 - Field campaign archive to be transferred to NOAA NCEI (~2019)
 - Target audience: Field Campaign investigators



GPM Ground Validation Olympic Mountain Experiment (OLYMPEX)
OLYMPEX Data Workshop
Seattle, WA
March 21-23

Deborah Smith delivered a presentation on GHRC and how we will archive and deliver OLYMPEX data to a broader audience, showing tools and information. She also worked with scientists during the meeting to improve metadata information.



NASA Time-Resolved Observations of Precipitation structure and storm Intensity with a Constellation of Smallsats (TROPICS)
TROPICS Applications Workshop
Miami, FL
MAY 8-10

Amanda Weigel participated in discussions to determine the data latency and other user needs of TROPICS applications data users. During these discussions she communicated the keen interest GHRC has in curating and archiving TROPICS data.



Geostationary Lightning Mapper (GLM)
2017 GLM Annual Science Team Meeting
Huntsville, AL
September 12-14

Helen Conover discussed the GHRC DAAC and its lightning data holdings in the context of local data and information management for GLM data, GOES-R Post Launch Test Field Campaign and other reference lightning data

Recommendation 2a: Become known within the community as the go-to place for weather hazards data



Stronger focus on **science presence** at AGU and AMS, in addition to informatics. A primary goal was to interact with current and potential users of GHRC data, as well as data providers.

Deborah Smith, Kaylin Bugbee, Amanda Weigel, Rahul Ramachandran, Manil Maskey, Sara Graves

- Multiple sessions, presentations and posters highlighting GHRC data, tools and services, and related research
- Flash talk at NASA Hyperwall showcasing Field Campaign Explorer

Deborah Smith, Amanda Weigel

- Posters highlighting GHRC data, tools and services
- Sessions related to GHRC data such as lightning, severe storms and field campaigns



AGU FALL MEETING

New Orleans
11-15 Dec. 2017 What will *you* discover?

Recommendation 2a: Become known within the community...

- Convening sessions on Natural Hazards; Innovative Data Tools and Services
- Submitted abstracts on
 - Metadata curation using DAPPeR
 - Broadening the use of OLYMPEX data
 - Outreach and education resources
 - VISAGE project – extending FCX to GPM GV data
 - Hands-on outreach – lightning color-by-number



- Submitted abstracts on
 - Data recipes
 - Science communication
 - GIS to study tornadogenesis
 - VISAGE project – extending FCX to GPM GV data
 - Hands-on outreach – lightning color-by-number

Session 3: Science & Outreach

- Updated Web Features and Information provide more ways to learn about our data and get what you need
 - Virtual Collection
 - Micro Articles
 - Data Recipes

Recommendation 6: data bundles, micro articles

<https://ghrc.nsstc.nasa.gov/home/micro-articles>

<https://ghrc.nsstc.nasa.gov/home/data-recipes>

- User characterization

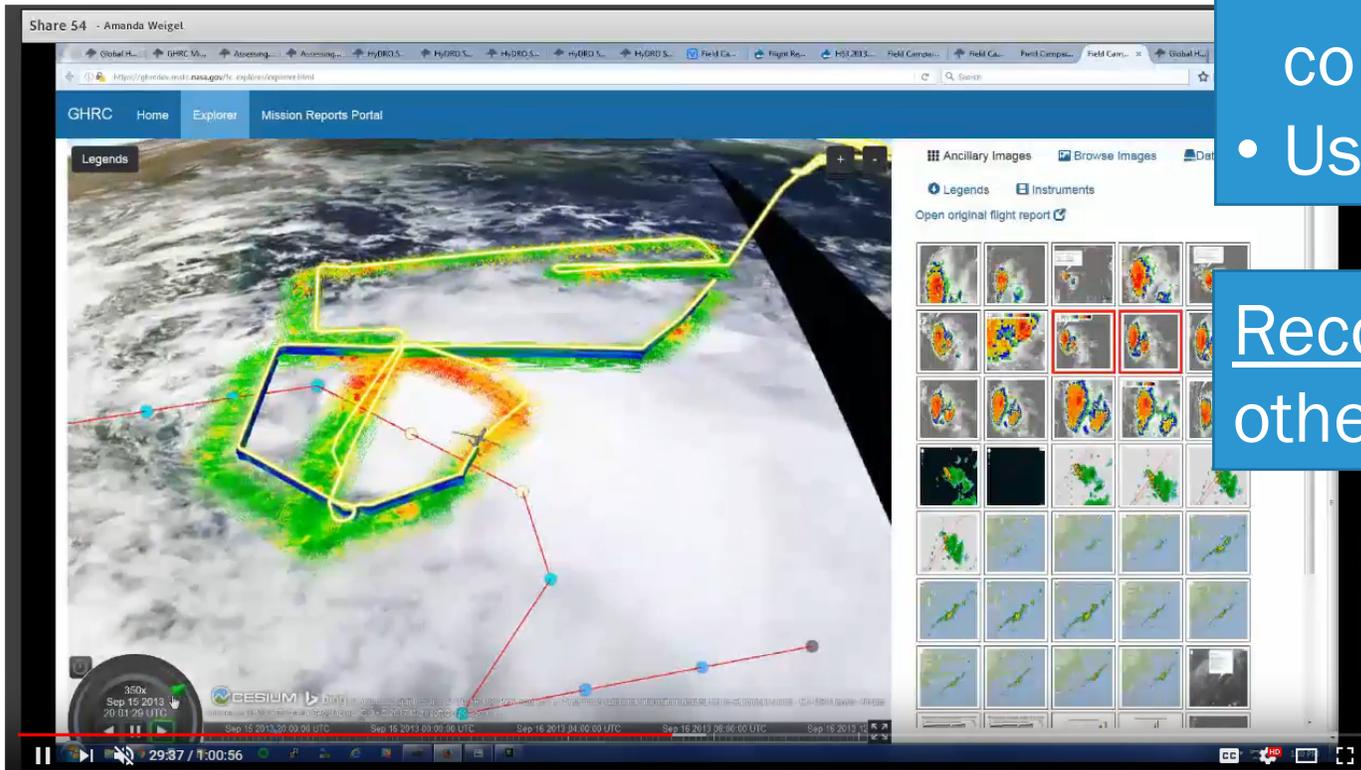
Recommendation 2c: user metrics

Session 4: Future Plans

- Data and Tool Assessment
- FCX Improvements

Recommendation 2b, 2c:

- Link current and proposed data holdings to GHRC mission components
- User metrics



Recommendation 5: extend FCX to other field programs

Side Meeting on SWOT Products

GHRC staff will meet tomorrow with Jessica Hausman, from PO.DAAC's Surface Water and Ocean Topography (SWOT) mission team, to discuss how some of the SWOT products might fit at the GHRC DAAC



Recommendation 7: plan to obtain targeted SWOT datasets to complement hazardous weather related to floods caused by excess precipitation

Artist's rendition of the deployed SWOT spacecraft (image credit: NASA/JPL)



Back-up Slides

Details of AGU 2016 and AMS 2017 Participation
Full Text of Recommendations

2017 GHRC User Working Group Meeting
Sept 26-27, 2017



AGU 2016 & AMS 2017 Details

AGU Fall Meeting 2016

GHRC posters and presentations include:

IN14A-06 **Riding the Hype Wave: Evaluating new AI Techniques for their Applicability in Earth Science** (Rahul Ramachandran)

IN22A-06 **Public-Private Partnership: Joint Recommendations to Improve Downloads of Large Earth Observation Data** (Ramachandran)

IN52A-04 **Deep Learning-Powered Insight from Dark Resources** (Manil Maskey)

ED13F-07: **Earth Science Data Education through Cooking Up Recipes** (Amanda Weigel)

IN53E-07: **Field Campaign Explorer: Simultaneous Data Exploration, Discovery, and Visualization** (Weigel)

Lightning Talk at NASA Booth about Field Campaign Explorer (Weigel)

IN33B: **Near Real-Time Data for Earth Science and Space Weather Applications II, Land, Atmosphere Near Real-time Capability for EOS (LANCE) AMSR2 Data System** (Deborah Smith)

PA51A: **Art and Design of Visualizations in the Geosciences: Inspiration, Interpretation, and Communication I, Art-inspired Presentation of Earth Science Research** (Kaylin Bugbee)

IN53C: **Innovative Tools and Services to Enable Data Use across Broad User Communities I, Virtual Collections: An Earth Science Data Curation Service** (Bugbee)

ED21B: **Amazing Technologies and Capabilities that Contribute to STEM I, Making connections: Where STEM learning and Earth science data services meet** (Bugbee)

In addition GHRC staff chaired or co-chaired several sessions:

IN51C **Spatial Data Infrastructure for Earth and Space Sciences: Analyzing, Visualizing, and Sharing Multidimensional Earth Science Data I Posters** (Ramachandran, Maskey)

IN53E **Spatial Data Infrastructure for Earth and Space Sciences: Analyzing, Visualizing, and Sharing Multidimensional Earth Science Data II** (Ramachandran, Maskey)

IN51D: **Innovative Tools and Services to Enable Data Use across Broad User Communities I** (Weigel)

IN53C: **Innovative Tools and Services to Enable Data Use across Broad User Communities II Posters** (Weigel)

IN33C: **Collaborations, Partnerships, and Alliances that Are Building, Sustaining, and Stewarding Data and Research Infrastructures in Support of the New Era of the Big Data Transdisciplinary World II** (Bugbee)

AMS Annual Meeting 2017

Education poster session, **Stimulating Remote Sensing Education through Knowledge Augmentation Services** (Weigel)

Observation Symposium (theme of 2017 meeting) poster session, **Visualize, Discover, and Analyze: A Data Center's Innovative Services for Addressing Observing System Challenges** (Smith)

Recommendation #1

(previously Recommendation #3)

GHRC should continue to hold AMS and AGU town halls, develop and distribute information brochures that describe their capabilities to potential data providers (e.g. field campaign PIs) and data users, utilize the NASA Hyperwall, and pursue other opportunities (BAMS) to enhance GHRC visibility. Old Recommendation #3 (New Rec #1) was to attend several venues and increase GHRC visibility. GHRC staff attended a number of professional conferences and workshops, but it was unclear exactly what informational brochures or exact material was presented in town halls. (The annual report did communicate the new vision and mission of the DAAC.) Several UWG members suggested hands-on activities might be of value, particularly in the context of an AGU hyperwall/NASA booth. GHRC seems to have a plan for AGU this year and should continue to have a presence at such venues on an annual basis. The UWG recommends establishing relationships with PIs and Program Managers to get access to invited program-level meetings and relevant data workshops. Having an internal plan to do this would be an ongoing recommendation, but does not require any specific actions vis-a-vis the UWG.

Recommendation #2, #2a

(merger of previous Recommendations #4, #5 and #6)

We feel that the GHRC has largely addressed the initial UWG recommendations from 2014-2015, which were broadly geared towards developing a unified theme for the DAAC, fulfilling that theme through current and planned data holdings, and developing tools that will better facilitate their users interaction with those datasets. The UWG feels that the next major issue for the GHRC is to continue this momentum by further developing their 5-10 year plan. The GHRC is in a unique position to be a data provider for “Storm Induced Hazards”, which already includes the following major categories: lightning, hurricanes and floods.

New Recommendation #2a: As part of their 5-10 year plan, the UWG recommends that GHRC should become the data subject matter expert on those categories – either through datasets that they provide in house or links to outside data – and become known within the community as the go-to place for datasets related to these subjects. The UWG has decided to merge Recommendations #4, #5, and #6 into this New Recommendation, as responses to these recommendations should contribute to the 5-10 year plan.

Recommendation #2b

(merger of previous Recommendations #4, #5 and #6)

New Recommendation #2b (previously Recommendation #4): Carry out dataset holdings analysis and create a reporting structure that categorizes what is available at GHRC and possibly elsewhere. This compilation should enable prioritization of efforts that will fill the most significant data voids, where these efforts align with the GHRC mission. The first half of this recommendation from 2015 was closed. The second half of the recommendation was left open. Specifically, Recommendation #4 (2015) suggested that “it would be helpful to see data holdings and mission broken down into a simple chart describing the data holding inventory broken out by and related to GHRC mission components.” The chart (or 5 equivalent) would provide a mission-driven means to gap-fill the dataset inventory and assist in developing tactical and strategic planning related to dataset acquisition. Working between the 2015 and 2016 UWG meetings, there has been more focused “bottom up” thinking on behalf of GHRC as it pertains to dataset inventory considerations (“what, who, where, why” approach) in the context of integrating data inventory with the GHRC mission.

Recommendation #2b *continued*

(merger of previous Recommendations #4, #5 and #6)

It is also clear that thought has been given to considering how GHRC could benefit from potential new dataset holdings in a strategic sense. For example, discussion of several potential new datasets discussed on day 1 “fit” the mission of GHRC and were presented in the context of mission when the question “why” was answered. This was especially true for tropical cyclone/convection focus where the bulk of the new datasets mentioned applied (e.g, SMAP Ocean Winds, TMI V7, TCIS, inter-calibrated TBs, CPEX field campaign). As another example, active pursuit of the SWOT dataset is ongoing and a good strategic move for supporting hydrologic hazards emphasis and synergies with ongoing precipitation dataset collections. Hence there has been significant forward movement in addressing Recommendation #4, though it is not clear that the recommendation should be closed completely. A chart (or equivalent) that clearly and succinctly demonstrated the linkages or traceability of current/proposed future data holdings to the GHRC mission components should be completed. Perhaps a hierarchical chart that showed traceability to specific GHRC mission components and then continued to NASA Earth Science focus area(s) would be useful. In addition to strategic planning such a chart would be useful for presentation to higher-level managers as well. Recommendation 4 could be closed quickly when this “traceability” chart is completed.

Recommendation #2c

(merger of previous Recommendations #4, #5 and #6)

New Recommendation #2c: Update public dataset information pages to include data holding analysis results that might be helpful to the user community (previously Recommendation #5). Determine a set of useful user metrics, with feedback obtained from the UWG, which can be routinely updated and made available to the NASA sponsor, UWG and broader community. Analysis of these metrics should inform the 5 -10 year plan (previously Recommendation #6). Declared to be in progress for the time being, but not specifically addressed. It would have been nice to see a page on the GHRC site that simply displays user metrics, hot datasets, etc., but it is not obvious that this exists. Also, there was no mention of the Dashboard tool that was presented last year, which looked promising; however, there is a move towards using Google Analytics to analyze what parts of the site are being used, where they're from, etc. which was presented by Deborah Smith.

Recommendation #3

(previously Recommendation #8)

Create a data lifecycle process for GHRC that can be applied to current and future holdings. Ask NSIDC and PODAAC for their policies and assess utility within GHRC. Publish the data lifecycle on the website, along with a contact, to provide clarity on the process for investigators interested in providing data. The questionnaire, lifecycle plan and work flows have been completed and are now on the website. So this part of the recommendation has been met. However, the location of the publish data information is not intuitive. The UWG recommends that lifecycle plans and workflows should be moved up a level (e.g under Resources instead of Resources>Documents)

Recommendation #4

(previously Recommendation #11)

Determine LIS technical specifications for data products, latency, formats, etc. Publicize this future data source at appropriate venues. This recommendation addresses the future importance of LIS on ISS data to the GHRC, also emphasizing outreach with regard to the GHRCs lightning holdings. While LIS on ISS is an upcoming mission, the UWG would like to see the GHRC be more proactive about future data holdings, and this is a prime example. To close the recommendation, the UWG would like to see this future dataset publicized, and technical specifications of the data products and latency available to potential users before the start of the mission. The GHRC presentation claims that this effort is in progress, though not much ISS/LIS information can be found on GHRC websites. Therefore, the recommendation remains open. The UWG envisions GHRC becoming a hub of all lightning data (i.e., New Recommendations #2 and #8). In addition to the existing lightning datasets at GHRC, they should also provide the services of virtual holdings, including brief introductions and links to all available lightning data from other resources, such as GLM data at CLASS, or other future lightning datasets from other Countries.

Recommendation #5

(previously Recommendation #12)

Develop a single tool that can provide broad use to multiple field campaigns and data types. Significant progress has been made, but the scope of the effort needs to be better defined. The Python tools are very helpful to those that use Python, but other code for other languages should also be considered. The Field Campaign Explorer is very impressive, although work remains to be done to make it more useful for more users. For example, it would be useful to apply such a tool to other field programs serviced by the GHRC (e.g., the numerous GPM field campaign datasets it archives). The change from Oracle to Open database saves money and allows for improved flexibility, as well as normalized and simplified database schema. Hydro 2.0 is considerably more superior to previous version, and certainly appears greatly simplify accessing desired datasets. While significant progress has been made, the UWG recommends that there is considerable work remaining on these tools, so this recommendation will remain open.

Recommendation #6

(previously Recommendation #17)

Create Data bundles for scientists who want to study processes. Demonstrate such bundling capabilities for review by the UWG. The committee was impressed by GHRC efforts towards bundling data for users who may be interested in processes. The virtual collection is a great effort and the case presented in Bugbee's talk on the GCPEX snow microphysics case study is an example of the GHRC addressing this recommendation. It appears to be a prototype for this type of data bundling. The micro-articles are another great way of offering the users an example of how to bring multiple data holdings together, however, it's not clear to the committee how this effort will be continued (i.e., who will determine the virtual collections in the future, and who will do the work? GHRC or the PIs?). We recommend that the GHRC continue to pursue this effort, and in particular, feel free to use the UWG as a sounding board for micro-articles that the GHRC feels would be interesting for publication on the website. The UWG envisions a process by which the GRHC proposes a topic for a micro-article to the UWG Chair, who will then seek the advice from the subject-matter experts on the UWG. Also, the idea of using an automatic program to do data bundling needs more research. Giovanni could be a great tool, but lacks advanced logic in that it seems to have trouble identifying the most relevant datasets for a specific event. For these reasons, the recommendation remains open.

Recommendation #7

(previously Recommendation #19)

Discuss the possibility of getting land data from SWOT mission at GHRC to complement hazardous weather related to floods caused by excess precipitation. This would complement other flood and extreme event (including precipitation) data sets. Some progress has been made on this recommendation based on the presentation. GHRC has identified specific datasets from the SWOT mission (pass-based lake and river levels and 21-day averaged levels for each). In addition to the planned datasets, they have identified the field campaign mission, AirSWOT, and any value-added products from the early-adopter community as potential datasets. GHRC should make the pursuit of these dataset a priority during upcoming years, as they are relevant to this “storm induced hazard” aspect of GHRCs mission. It is recommended that GHRC develop a plan to obtain the targeted SWOT datasets. Specific steps for the SWOT data holding plan can include expanding surface water hazard products and pursue existing flood products, such as the MODIS flood product, to develop a user base and reputation within the field. In addition to developing the product collection, GHRC should be in contact with PIs and program managers prior to the SWOT mission to begin discussions on the SWOT products.

Recommendation #8

(previously Recommendation #20)

GHRC should include GOES GLM data in its portfolio of accessible data whether stored in house or as a virtual data set. Functionality should be seamless with other holdings. GHRC should continue planning to be the first stop for accessing global lightning data. GLM should be managed as a virtual data collection since it is funded for data stewardship by NOAA. GHRC should plan to coordinate with NOAA CLASS/NCEI-NC and also maintain awareness of on-going coordination through the WMO of the upcoming operational lightning data from operational space agencies (e.g., refer to the SATURN Satellite User Readiness portal). The GHRC has the only historic data archive for the NASA LIS data since launch of OrbComm-1 with OTD (1995-2000) and TRMM with LIS (1997-2015), as well as ancillary reference data from the commercial lightning data providers (e.g. Vaisala, Earth Networks, WWLLN, and select LMA regional data sets). Users can be found worldwide. These should be continued to be preserved with the launch of the GOES-R series that will provide an additional 20+ years of lightning data from the NOAA Geostationary Lightning Mapper for the western hemisphere, as well as the ISS-LIS data with planned launch in 2016 (also providing coverage to 54° latitude as does GLM).

Recommendation #8 continued

(previously Recommendation #20)

Also, China will host a Lightning Imager on their next series of FY-4 GEO satellites with first launch in December 2016, as will EUMETSAT with the launch of a Lightning Imager on their next series of MTG GEO satellites. The GLM, MTG-LI and CMA-GLI L1B and L2+ data (even, group, flash data components are in the same format as the LIS data structure). Thus, the new GEO satellites provide an extension of the initial LIS climate data set. The GLM L1B and L2+ are operational NOAA data products with long-term stewardship funded by NOAA and archived in the NOAA Comprehensive Large Array Storage System (CLASS) managed by the National Centers for Environmental Information, NCEI-NC). GHRC should coordinate with CLASS so the data are viewable at GHRC and point to CLASS, perhaps providing readers for GHRC user communities. The GHRC also has field campaign holdings supporting field campaigns. The reference data collected during the GOES-R Field Campaign should also be bundled with the satellite data for one-stop shopping. The GOES-R Program is funding GHRC to develop a web portal for the 2017 GOES-R field campaign. NOAA CLASS is making plans for permanent stewardship after the campaign is completed. Again, coordination between GHRC and CLASS is desirable and should be a metric/milestone for FY17.



Thank you!

Questions?

2017 GHRC User Working Group Meeting
Sept 26-27, 2017

