



## Report

## User needs and assessing the impact of low latency NASA Earth observation data availability on societal benefit

Molly E. Brown <sup>a,\*</sup>, Mark L. Carroll <sup>b</sup>, Vanessa M. Escobar <sup>b</sup><sup>a</sup> Biospheric Sciences Laboratory, Code 618, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA<sup>b</sup> Sigma Space Corporation, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA

## ARTICLE INFO

## Article history:

Received 26 March 2014

Accepted 5 May 2014

Available online 28 June 2014

## Keywords:

NASA

Latency

Earth observation

Applications

Operational

Satellite data

## ABSTRACT

Since the advent of NASA's Earth Observing System, knowledge of the practical benefits of Earth science data has grown considerably. The community using NASA Earth science observations in applications has grown significantly, with increasing sophistication to serve national interests. Data latency, or how quickly communities receive science observations after acquisition, can have a direct impact on the applications and usability of the information. This study was conducted to determine how users are incorporating NASA data into applications and operational processes to benefit society beyond scientific research, as well as to determine the need for data latency of less than 12 h. The results of the analysis clearly show the significant benefit to society of serving the needs of the agricultural, emergency response, environmental monitoring and weather communities who use rapidly delivered, accurate Earth science data. The study also showed the potential of expanding the communities who use low latency NASA science data products to provide new ways of transforming data into information. These benefits can be achieved with a clear and consistent NASA policy on product latency.

Published by Elsevier Ltd.

## 1. Introduction

Since the advent of NASA's Earth Observing System, knowledge of the practical benefits of Earth science data has grown considerably. The community using NASA Earth science observations in applications has grown significantly to serve national interests with increasing sophistication. The National Research Council's 2007 Earth Science Decadal Survey report stated that the planning for applied and operational considerations in the missions should accompany the acquisition of new knowledge about Earth [1]. The scientific and applications-oriented goals of the Decadal Survey and NASA Climate Initiative can complement one another, as science enables improved decision-making and the practical benefits of using Earth observations build interest in further scientific discovery [2]. As support for applications has increased, the new user base has begun requesting data sooner after acquisition to increase the types and operational nature of applications and to improve the timing of decision-making. Data latency refers to the time it takes from data acquisition on a satellite until it reaches a user in an

actionable format. This study focuses on understanding the impact of the length of delay between acquisition and delivery of data, with a focus on operational users.

Data latency, to date, has not been a prominent concern in the science research objectives of NASA Earth science missions. The primary goal of NASA Earth science missions is to create science quality datasets that can be used in analyses to characterize processes of the Earth system and to answer research questions. Algorithms are used to process raw science instrument observations into interpreted or thematic outputs with the basic research user in mind. These algorithms begin as theoretical processes and pass through several stages of maturity where the results are verified against ground "truth" until they achieve a level of consistency called validated [3]. The Committee on Earth Observation Satellites (CEOS) Validation Hierarchy identifies many stages to this process, from beta (initial results) to full validation (well characterized with known errors) [4]. To achieve the best quality outputs, it is often necessary to utilize information from more than one source in the algorithm. There can be additional time lag for each external satellite dataset that is used in the algorithm. This can contribute significantly to the data latency when creating the highest quality, most stable products.

As applied and operational considerations are increasingly taken into account, however, the question of how to best meet the

\* Corresponding author. Tel.: +1 301 614 6616; +1 703 855 6190 (mobile)

E-mail addresses: [molly.brown@nasa.gov](mailto:molly.brown@nasa.gov), [mebrown69@gmail.com](mailto:mebrown69@gmail.com) (M.E. Brown).

latency needs of the broad user communities (both scientific and operational) of NASA data products is now demanding more attention. The release of the National Research Council's first Earth Science Decadal Survey report in 2007, set the foundation for applied and operational considerations to accompany the acquisition of new knowledge about the Earth. Among the report's key recommendations are 1) strengthening the dual role of Earth system science, science and applications, and 2) considering socioeconomic factors in the planning and implementation of Earth observation missions [1]. The latter recommendation includes 'establishing priorities of the applications community in space-based missions' and 'creating closer institutional relationships between the science and applications (user) communities' [1]. Establishing the latency needs of user communities, as well as reviewing options for attaining low latencies on NASA Earth Science missions are goals that respond to the National Research Council's recommendations. This report summarizes information on current and future data latencies, how these meet user needs, and summarizes the results of a professional review focused on characterizing how NASA data is used in the broader community. The full document, that includes an analysis of all questions in the survey and a comprehensive assessment of the comments received can be found online in Brown et al. (2013), a NASA technical report [5].

## 2. Summary of review findings

The authors of this report conducted a study in 2013 that focused on better understanding the needs of the satellite data user community for rapidly delivered data products. First a review of the technical aspects involved in provision of data to users was completed that investigated current and upcoming missions. Second, a questionnaire was developed and circulated among the known user communities to get direct feedback from users of the data on what their needs were. A steering committee with representatives from many of the NASA Earth Science directorates was formed to provide guidance, review documents, and answer questions from the study team. The online survey sent to approximately 7000 people who are part of a broad community of practice who know about or use NASA satellite data. The review consisted of fourteen questions and was open from April 1 through May 31, 2013. The total number of respondents was 526, approximately 13 percent of the people who received the review.

The responses from the questionnaire, conversations with the NASA the steering committee and interactions with individual users all indicate that the need for low latency is determined on a product basis rather than on a mission basis. For example, it is necessary to know where clouds are as quickly as possible but it is not necessary to have a rapid evaluation of the land cover for the same area. Both data products in this example can be generated by the Terra mission MODIS instrument but have very different latency requirements.

We identified three basic "classes" of latency needs from the user communities that were polled:

- Operations: need data in less than 3 h.
  - Numerical weather forecasting.
  - Disasters management, such as fighting fires, anticipating floods and landslides, and responding to the aftermath of disasters.
  - Field campaigns.

- Near real time: need data within a day.
  - Strategic allocation of resources for forest/wildfire management.
  - Environmental monitoring.
    - Agriculture – food security, food production and weather impacts.
    - Human health – air quality, conditions for propagation of disease vectors.
  - Field campaigns.
- Research and science: need best quality data, time/latency of data delivery is less relevant.

The study team and the steering committee recognize that there are many other uses that fit within these classes that are not listed above, or who were not reached or were able to respond to our review. We also acknowledge that the approach of surveying the professional community already engaged with NASA excludes the larger community of practice that could use satellite data were it to be of lower latency, easier to access or provided commercially preprocessed to exactly the parameter of interest. Given these caveats, the results obtained are representative of the community interested in using low latency NASA data. Full details on the survey results, along with the questions asked and analysis of responses can be found in the full NASA report [5].

## 3. Achieving low latency satellite data

Direct broadcast is the most rapid way to get satellite data from a sensor to the ground processing system. To achieve low latency satellite product delivery, direct broadcast is not enough, however. Significant investment of the user facility designed to accommodate rapid data processing of higher-level data products is also necessary. This investment can be part of a mission and its goals and objectives from the start, as was the case with Suomi National Polar-orbiting Partnership (NPP) sensor system, or can be established after the launch of the satellite. For example, NASA developed the Land Atmosphere Near real-time Capability for EOS (LANCE) after the launch of its Earth Observing System [6]. LANCE provides products from the Moderate Resolution Imaging Spectroradiometer (MODIS), Ozone Monitoring Instrument (OMI), Atmospheric Infrared Sounder (AIRS), and the Microwave Limb Sounder (MLS) instruments in near real time, or less than 3 h after acquisition [7]. The system supports applications interested in monitoring and analyzing a wide variety of natural and man-made phenomena, but that need the data in near real time in order to meet operational timelines imposed by their systems.

Reducing the latency for real-time operations is largely a hardware issue. To reduce the latency from 3 h to 1 h would require use of Tracking and Data Relay Satellite (TDRS), funding additional ground receiving stations, upgrading ground based computer networks, or other system level units to speed up the process of getting the physical data from the satellite to the ground processing system. Onboard processing of raw sensor output to a usable, geo-located product could take advantage of the time between acquisition and downlink opportunity and would give the user a product that is immediately useful. If onboard processing occurs there would need to be a priority system in place to determine the order in which data is sent to the ground.

Reducing latency for near real time products of 1–3 h is largely an algorithm level issue. The primary delay in processing Level 2 and higher products is related to algorithmic needs for ancillary information to perform their calculations. Science requirements

have accuracy levels defined by the science teams. To reduce latency for products requiring ancillary information it may be possible to lower requirements for accuracy of “preliminary” data products by allowing the use of static (or less timely) ancillary data products. Other options for marginal gains include on-the-fly reformatting and repackaging for users and direct push of data to the application or user facility.

In many cases, user communities have not reached consensus on what is needed in terms of latency. There will continue to be trade-offs between the quality of satellite data products and the timeliness of their delivery. Some users may decide that a more accurate, less timely product is preferable. Experience with the EOS data products has shown that once a good quality, low latency data product has been developed and is consistently available, new ways of using the data and new communities who rely on the data will emerge that will have significant societal value. Clearly defined and reliable low latency data products on key environmental and climate variables would broaden the NASA data user community into new areas not yet known [8].

#### 4. Policy implications

Improving the number and quality of low latency of datasets available will increase the value of the services delivered to society by fire fighters, weather forecasters and other users who require low latency data products to be effective. With a clearly articulated policy on latency for each new mission it funds, NASA could expand the number and type of satellite data users.

The latency study was undertaken to understand the user community needs for low latency data from missions. We have found that there are a number of communities, such as weather forecasters, agricultural communities, and disaster response teams, that have applications that need or would benefit by rapid provision of data products from NASA missions. These needs are for specific data products and not necessarily for the entire suite of data products from a given mission. At present NASA supports these needs either through direct broadcast capabilities or through independent rapid turnaround processing like LANCE that is outside of the standard mission data processing scheme. Based on this information we have four primary conclusions.

First, clear guidance is needed from NASA headquarters regarding the support that will be provided for operational and near real time applications for all its data products. This guidance should address expectations on science missions to provide support for applications so that the user facilities can plan for available or coming data products. If latency requirements are defined during mission development at the product level, then during the development of the mission, the science team member responsible for that product can design the algorithm with that in mind and look for ways to speed the process. The algorithm developer is best suited for identifying ways to generate a data product faster and understanding the consequences to the product integrity if changes are made in the processing scheme.

Second, from the applications perspective, latency needs are based on specific data products, which may come from multiple

missions with different designs and product characteristics. Data products need to have specific and well-defined latencies to maximize their usefulness to the broader applications community. A reduction in latency may be possible by prioritizing the order of production of products based on latency requirements. This includes processing data on-orbit if possible, so that data can be ready to use when it reaches the ground. If low latency can only be met by producing near real time products separately from standard products, then the difference between low latency products and standard products needs to be clearly described and documented.

Third, clear guidance is needed from NASA headquarters regarding the use of direct broadcast on future satellite data systems. If the capability is to be included some investment in modularization could make it easier to incorporate into spacecraft engineering. There is a group of applications users that will continue to rely on direct broadcast systems regardless of any changes or improvements to either operational or near real time systems. This can be due to a lack of network infrastructure in their region or because of a need for immediate response data for model operation or emergency response.

Finally, serving the applications community by providing low latency data products is of significant value to NASA and to its missions. By clearly defining the latency of each NASA data product, and linking new missions to heritage observations, new applications will be developed that cannot envisioned today. The environment sector of the US economy, which includes disaster response, services, regulatory activities, and the insurance sector, is a significant area where NASA can contribute directly to society through its observations of the Earth. NASA's investment in pre-mission applications development for Decadal Survey missions is a significant step in promoting the use of NASA data in applications and operational processes. Although current uses of low latency data are primarily in fire, navigation, agriculture and weather applications, there is an enormous opportunity to expand these uses far beyond what is done today. By working to reduce latencies, NASA can greatly increase its impact on the US economy.

#### References

- [1] NRC. Earth science and applications from space: national imperatives for the next decade and beyond. Washington DC: The National Academy of Sciences; 2007.
- [2] NASA. NASA 2010 science plan. NASA; 2010. p. 86.
- [3] Morisette JT, Privette JL, Justice CO. A framework for the validation of MODIS land products. *Remote Sens Environ* 2002;83:77–96.
- [4] Morisette JT, Pinzon JE, Brown ME, Tucker CJ, Justice CO. Initial validation of NDVI time series from AVHRR, VEGETATION and MODIS. In: *Proceeding of the 2nd SPOT VEGETATION users conference*. Antwerp, Belgium; 2004. pp. 149–54.
- [5] Brown ME, Carroll ML, Escobar VM. Study on data latency needs and requirements. Greenbelt, MD: NASA Goddard Space Flight Center; 2013. p. 31.
- [6] Kaufman YJ, Herring DD, Ranson KJ, Collatz GJ. Earth observing system AM1 mission to earth. *IEEE Trans Geosci Remote Sens* 1998;36.
- [7] Michael K, Murphy K, Lowe D, Masuoka E, Vollmer B, Tilmes C, et al. Implementation of the land, atmosphere near real-time capability for EOS (LANCE). In: *IEEE geoscience and remote sensing symposium*. Honolulu, HI: IEEE International; 2010. pp. 1442–5.
- [8] GCOS. Guideline for the generation of datasets and products meeting GCOS requirements. Geneva: Global Climate Observing System, United Nations World Meteorological Organization; 2010. p. 10.