Study on Data Latency Needs and Requirements

Author Team
Molly E Brown, Mark L. Carroll, Vanessa M. Escobar
Biospheric Sciences Laboratory
NASA Goddard Space Flight Center
Greenbelt, MD 20171

HQ Leadership
Francis Lindsay, Applied Sciences Division

Latency Steering Committee Membership
Jeanne Behnke, Stephen Berrick, Jared Entin, Michelle Gierach, John Haynes, Kenneth Jucks, Allison Leidner, Daniel Mandl, Edward Masuoka, Chis Mattman, Kevin Murphy, John Murray, Michael Seablom, Chi Wu
Executive Summary

This Latency Study was initiated by the Earth Science Division (ESD) under the auspices of the Flight Program with the Applied Science Program to act as the study implementation lead. The objective of the study is to quantify the anticipated latency needs of application and science users for the upcoming Decadal Survey as well as current Earth Science Missions. Special attention is given to the development of synergies between latency requirements across communities and where advances in ground systems, data processing and other means can elicit greater use of these data. Overall, the ESD seeks to assess options for meeting latency desires of the community without unduly driving-up the cost of missions. In most cases, the provision of data in short time steps has been done as an afterthought after the satellite was on orbit and may not represent the most efficient or cost effective solutions given technological developments and potential improvements in ground-based data delivery mechanisms.

Provided in this report is a complete description of the findings with definitions and explanations of what goes into measuring latency as well as how users and applications utilize NASA data products. The approach included a review of published materials, direct interviews with mission representatives, and an online professional review that was distributed to over 7000 individuals. We have identified 3 classes of users: operational (need data in 3 hours or less), near-real time (need data within a day of acquisition), and scientific users (need highest quality data, time independent). Operational users are typically expert data users in the weather modeling and data assimilation communities and near-real time users are emergency response and fire managers, among others. We have also determined that most users with applications are interested in specific types of products that may come from multiple missions. These users will take the observations when they are available, however the observations may have additional applications value if they are available either by a certain time of day or within a period of time after acquisition. NASA has supported the need for access to low latency data on an ad-hoc basis and more substantively in stand-alone systems such as the MODIS Rapid Response system and more recently with LANCE.

The increased level of support and advertising of that support has grown the community of individuals and organizations that use low-latency science data to supply decision-making processes with updated information. These applications are increasingly high profile and have high societal value, which is of importance to NASA, the US Government and the broader society. The primary conclusions of this report revolve around clarifying NASA’s intentions to support operations and near real time needs for quick access to data after acquisition. Specifically, clear statements from NASA Headquarters are needed to indicate the requirement for Decadal Survey missions to provide low latency products, and the level of support NASA is willing to provide to make these products available. Technology solutions such as direct broadcast should be balanced with software or science algorithm solutions that will include defining new low latency product accuracy and precision requirements. The results of the analysis clearly show the significant benefit to society for serving the needs of the agricultural, emergency response, environmental monitoring and weather communities who use low latency data today, and to grow the use of low latency NASA science data products into new communities in the future. These benefits can be achieved with a clear and consistent NASA policy on product latency.
## Contents

Executive Summary .................................................................................................................................. 1

1. Introduction ......................................................................................................................................... 4

2. Background and Definitions .................................................................................................................. 5
   2.1. Data Latency ...................................................................................................................................... 5
   2.2. Research and Science Validated Products ...................................................................................... 7
   2.3. Applications and Operational Processes ......................................................................................... 7
   2.4. NASA Earth Science Missions Data Access .................................................................................... 7

3. Mission Latencies ................................................................................................................................... 8
   3.1. Case Study: Visible/NIR Observations ............................................................................................ 10
   3.2. Case Study: Altimeter Observations ............................................................................................... 10

4. Direct User Feedback ............................................................................................................................ 10
   4.1. Respondent Characteristics ........................................................................................................... 11
   4.2. Latency Needs ................................................................................................................................. 11
   4.3. How Low Latency Improves Results ............................................................................................. 16
   4.4. Why Low Latency Improves Results ............................................................................................. 17

5. Discussion ............................................................................................................................................ 19

6. Conclusions .......................................................................................................................................... 21

References ................................................................................................................................................. 23

Appendix A. .............................................................................................................................................. 24

Appendix B. .............................................................................................................................................. 27

Appendix C. .............................................................................................................................................. 30
1. Introduction
Since the advent of the Earth Observing System, knowledge of the practical benefits from 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. The National Research Council’s 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 (NRC, 2007). The scientific and applications-oriented goals of the Decadal Survey and Climate Initiative missions can complement one another, as science enables improved decision-making and the practical benefits of using Earth observations build interest in further scientific discovery. As support for applications has increased the new user base has begun requesting data sooner after acquisition 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 data latency on users with a focus on applications users.

NASA missions are designed with research science objectives. These objectives are used to derive “Level 1 Science Requirements” which are set early in the process and are the basis for the design decisions that are made throughout the development of the mission. Typically the research science objectives consider how to create the most accurate data that can answer the science questions posed for the mission. In this context, data latency is not always a primary concern and can be subjugated to the importance of having the highest possible accuracy and precision. Data latency is, however, a major factor in the utility of data products for applied and operational uses as well as some scientific investigations. Provision of data describing a disaster, such as forest fire or flood, a week after the event has concluded is not useful to the decision makers for that event. Data from NASA’s current missions have been incorporated into applications that support operational agencies such as the National Oceanographic and Atmospheric Administration (NOAA) numerical weather forecasting, USDA – Forest Service wildfire monitoring, and USDA Foreign Agriculture Service agriculture forecasting just to name a few. The tolerance for data latency is often explicit to a given application. Many of NASA’s future missions have data products that may be extremely valuable if they can reach the applied communities quickly after collection.

This Latency Study was initiated by the Earth Science Division (ESD) under the auspices of the Flight Program with the Applied Science Program to act as the study implementation lead. The study seeks insight into the anticipated latency needs of application and science users for the upcoming Decadal Survey and other Earth Science Missions. Special attention is given to the development of synergies between latency requirements across communities and where advances in ground systems, data processing and other means can elicit greater use of these data. Overall, the ESD seeks to assess options for meeting latency desires on the missions without unduly driving-up the cost of missions. In most cases, the provision of data in short time steps has been done as an afterthought after the satellite was on orbit and may not represent the most efficient or cost effective solutions given technological developments and potential improvements in ground-based data delivery mechanisms. Thus the following are the two objectives of this study:
• *Technical Capabilities and Data Latency.* Examine possible methods and mechanisms for delivering data that meets data latency targets for data users.

• *Data Latency & User Needs.* Assess the probable data latency targets for the suite of ESD planned missions through 2020.

This study was conducted in two phases. 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 are their needs. A steering committee with representatives from many of the Earth Science directorates was formed to provide guidance, review documents, and answer questions from the study team. The team was led by Molly Brown with assistance from Mark Carroll and Vanessa Escobar with Frank Lindsay providing programmatic guidance from NASA Headquarters. We will describe how data latency is controlled by instrument design and data processing along with where the largest delays in product delivery usually occur. We also provide quantitative and qualitative evidence of the user requirements, from the questionnaire, and will link these needs back to the data latencies planned by NASA.

### 2. Background and Definitions

In this section we will discuss the characteristics of the current state of data acquisition and delivery and provide definitions of terms used in the study.

#### 2.1. Data Latency

Data latency is the time it takes to get data acquired by a satellite into the hands of a user in an actionable format. There are four components that when taken together add up to the time it takes to deliver Earth science remote sensing data products:

- Data acquisition and on-board storage of data on the satellite until down-link begins;
- Transmission of data from satellite to ground and from the downlink location to the primary data processing center;
- Data product processing, or converting Level 0 raw to usable form Level 2 or higher (see Table 1); and
- Providing the data to the customer through a database, direct download, FTP or other method.

The length of time that each of these steps takes can be minimized through investment of resources, but there are significant constraints on the ability to reduce the time spent in some of the stages. Data acquisition times are determined by the design of the instrument and are often measured in nano-seconds. On board storage is determined by the transmission method from satellite to ground. There are 3 general methods for transmission to the ground:
• Direct downlink to NASA ground stations
  o Requires onboard data storage until the satellite is within line of sight to the
ground station; typical latencies around 90 minutes

• Tracking Data Relay Satellite (TDRS) uses orbiting communications satellites to relay the
instrument observation from the Earth observing satellite to a NASA ground station
(see http://tdrs.gsfc.nasa.gov/ for a complete description)
  o Requires onboard data storage until satellite is within line of sight to TDRS
satellites (space network) which can transmit to the ground stations; typical
latencies around 50 minutes.

• Direct Broadcast sent from the satellite directly to a user facility that has their own
satellite data receiving station
  o Requires no onboard storage, facility receives only data that the satellite “sees”
when it is within line of sight of the receiving station; typical latencies of a few
minutes

Data received at the ground receiving station, from each of the 3 methods above, are raw
satellite observations usually defined as Level 0 (Table 1) and will require further processing on
the ground before it is usable for research or applications. This processing is not done at the
NASA ground receiving stations so the data must be transmitted via ground computer networks
to processing facilities. The ground data processing facilities are defined in the mission and
tend to vary by science discipline, for example much of the data for land remote sensing from
Visible and Near Infrared (Vis/NIR) sensors occurs at Goddard Space Flight Center whereas
much of the radar data are processed at the Alaska Satellite Facility.

### Table 1 Scientific Data Level definitions from the EOSDIS Data Panel and the Committee on Data Management, Archiving and Computing (CODMAC)

<table>
<thead>
<tr>
<th>Level Name</th>
<th>Processing Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 0 (L0)</td>
<td>Level 0 data products are reconstructed, unprocessed instrument/payload data at full resolution; any and all communications artifacts, e.g. synchronization frames, communications headers, duplicate data removed.</td>
</tr>
<tr>
<td>Level 1A (L1A)</td>
<td>Level 1A data products are reconstructed, unprocessed instrument data at full resolution, time-referenced, and annotated with ancillary information, including radiometric and geometric calibration coefficients and georeferencing parameters, e.g., platform ephemeric, computed and appended but not applied to the Level 0 data.</td>
</tr>
<tr>
<td>Level 1B (L1B)</td>
<td>Level 1A data that have been processed to sensor units (not all instruments will have a Level 1B equivalent).</td>
</tr>
<tr>
<td>Level 2 (L2)</td>
<td>Level 2 data products are derived geophysical variables at the same resolution and location as the Level 1 source data.</td>
</tr>
<tr>
<td>Level 3 (L3)</td>
<td>Level 3 data products are variables mapped on uniform space-time grid scales, usually with some completeness and consistency.</td>
</tr>
<tr>
<td>Level 4 (L4)</td>
<td>Level 4 data products are model output or results from analyses of lower level data, e.g. variables derived from multiple measurements.</td>
</tr>
</tbody>
</table>

In the ground processing facility there are individual algorithms to generate specific output
products at the levels of processing shown in table 1. For most missions, the latency for raw, unprocessed Level 0 data is very low, less than three hours and in many cases shorter periods. Processing the data through the different levels is an additive process whereby Level 2, for example, cannot be made until Level 0 and Level 1 have been produced, hence latency increases with each level of processing.

From an applications perspective there is an additional component of latency which incorporates the time to the next observation. This measure is important to the application because in most cases they will need to update the information in order to make the “next” decision. For example, if the satellite goes over a location once a day but this overpass is an hour after a decision needs to be made, then the data will always be 23 hours old, even with very low latencies.

2.2. Research and Science Validated Products

The primary goal of NASA Earth science missions is to create science quality data sets that can be used to answer research questions and in analyses to characterize processes or the present state of the Earth system. Algorithms are used to process satellite data into interpreted or thematic outputs. The 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. There are many stages from beta (initial results) to full validation (well characterized with known errors) for satellite derived data products according to the CEOS Validation Hierarchy (Morisette et al., 2004). 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 (to the satellite data) dataset that is used in the algorithm. This can contribute significantly to the data latency.

2.3. Applications and Operational Processes

Applications are defined as innovative uses of mission data products in decision-making activities for societal benefit. Operational processes are a subset of applications that are typically characterized by needs for continuous data inputs and often needs for quick turnaround on decision processes. Agricultural monitoring to project yields and potential food shortages is an example of an application of remotely sensed data. Other applications of satellite data require very low latencies. The classic example is the numerical weather forecast community. Recent research has shown that operational assimilation of direct atmospheric measurements such as the data from the Advanced Microwave Sounding Unit (AMSU) has increased the Global Forecast System anomaly predictive capacity from 50% to 80% of the variance from the 1970s to the 2000s (Marshall et al., 2007). These improvements are only due to the ability of models to assimilate observations that are available at low latencies of less than three hours. This time requirement is driven by the regular re-running of weather models every hour to predict conditions three or more days ahead. The initial conditions of the model, set by satellite observations, are critical for the accuracy of the forecast.

2.4. NASA Earth Science Missions Data Access
NASA’s Earth Science Directorate currently operates 23 missions and has 15 planned for launch in the next 8 - 10 years. These missions are driven by research needs that are articulated in the level 1 science requirements for each mission. These requirements are used to plan the mission, design the instrument, and define the products derived from the measurements from the instrument. Since science is rarely conducted in near-real time, low latency is not usually considered a high priority when setting mission science requirements. After the Decadal Survey, applications needs have been introduced as a consideration in mission planning, but is not of primary importance in setting priorities. Changing latencies after the Level 1 Science requirements are set can be costly and detrimental to the mission.

Currently there are 3 functional ways to get NASA data: standard production, near real time production, or direct broadcast. Standard products are generated by the mission science team and the algorithms developed in conjunction with the instrument and satellite design by the mission before launch. These products are well characterized, have gone through a peer review process for the methodologies included in Algorithm Theoretical Basis Document (ATBD), and are typically accessed through a NASA supported data distribution facility such as the Distributed Active Archive Centers (DAAC). The standard products serve the broad research science community as well as applications that involve long term monitoring.

In response to requests from other federal agencies for faster access to satellite data NASA has provided two mechanisms for rapid access to data. The direct broadcast systems, onboard the satellite, enable an individual user facility to download data directly from a satellite that is within line of sight. These data will represent the region where the facility is located and will be Level 0 data that require further processing at the user facility to create actionable data. More information on direct broadcast can be found at http://directreadout.sci.gsfc.nasa.gov/ . NASA has also provided rapid access to data to other federal agencies and users on an ad-hoc basis and through stand-alone systems. The Land Atmosphere Near real-time Capability for EOS (LANCE) uses algorithms that are based on the standard product algorithms but with reduced requirements to provide data faster than the standard production system. When using data other than the standard mission products it is necessary to consider the impact of changes in the processing approach on the validation of these products will be required, since changes to the processing approach may introduce unforeseen bias, spatial variability or geometric differences between the near real time product and the original validated science product (Morisette et al., 2002).

3. Mission Latencies

For the purposes of this study, several missions currently in operations or in development were polled to determine what their product latencies are. There is no one document or location where latency information is held, and it required speaking directly with each mission to find this information. The missions shown in table 2 below represent different types of sensor systems that are both on orbit and in several different stages of development for the NASA
Earth Science Directorate. This is clearly not a complete list of missions but rather is intended to show examples of the breadth of latencies NASA products have.

Table 2. NASA ESD Mission latencies

<table>
<thead>
<tr>
<th>Mission Name</th>
<th>GPM</th>
<th>ICESat-2</th>
<th>Landsat 8</th>
<th>MODIS</th>
<th>SMAP</th>
<th>VIIRS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of Mission</td>
<td>Microwave Radiometer and Dual Frequency Precip RADAR</td>
<td>Altimetry</td>
<td>Spectrometer</td>
<td>Spectrometer</td>
<td>RADAR and Microwave</td>
<td>Spectrometer</td>
</tr>
<tr>
<td>Measurements</td>
<td>Rainfall amount and intensity</td>
<td>Range and elevation measurements</td>
<td>Land cover and processes; Land cover change</td>
<td>Land cover and processes; Ocean processes; Atmospheric Composition</td>
<td>Soil moisture, surface state, water detection</td>
<td>Land cover and processes; Ocean processes; Atmospheric Composition</td>
</tr>
<tr>
<td>Status</td>
<td>Jun, 2014</td>
<td>Fall 2016</td>
<td>operations</td>
<td>operations</td>
<td>Oct, 2014</td>
<td>operations</td>
</tr>
<tr>
<td>Products available in 0 - 3 hrs</td>
<td>L0 - L1 (Brightness Temp; Radar Enhanced Precipitation)</td>
<td>none</td>
<td>L0</td>
<td>L0 - L2 (raw telemetry, calibrated radiance, Preliminary products: SR, VI)</td>
<td>none</td>
<td>L0 - L2 (raw telemetry, calibrated radiance, Preliminary products: SR, VI)</td>
</tr>
<tr>
<td>Products available within 12 hrs</td>
<td>L2 (Reflectivities; Precipitation with Vertical Structure)</td>
<td>none</td>
<td>L1A-L1B</td>
<td>L1 - L2 (Science products: SR, LST, fire)</td>
<td>L1 (calibrated Radar)</td>
<td>L1 - L2 (Science products: SR, LST, fire)</td>
</tr>
<tr>
<td>Products available within 1 day</td>
<td>none</td>
<td>L0 (raw telemetry)</td>
<td>L1T</td>
<td>L3 (Science products gridded dailies: snow, LST)</td>
<td>L2 (soil moisture swath)</td>
<td>L3 (Science products gridded dailies: snow, LST)</td>
</tr>
<tr>
<td>Products available within 3 - 4 days</td>
<td>none</td>
<td>L1 - L2A (calibrated range measurements)</td>
<td>none</td>
<td>none</td>
<td>L3 (Freeze/Thaw, Soil Moisture daily)</td>
<td>none</td>
</tr>
<tr>
<td>Products available within 1 week</td>
<td>none</td>
<td>none</td>
<td>none</td>
<td>none</td>
<td>L3 (Science products gridded composites: SR, NDVI, LST, LAI, Albedo)</td>
<td>L3 (Science products gridded composites: SR, NDVI, LST, LAI, Albedo)</td>
</tr>
</tbody>
</table>

Abbreviations: SR – spectral reflectance, LST – land surface temperature, NDVI – normalized difference vegetation index, LAI – leaf area index, VI – vegetation index
Table 2 shows that similar types of missions have similar latencies, and the approach the mission takes in producing its data products is influenced by the research and investment in previous missions.

3.1. Case Study: Visible/NIR Observations

There are numerous sensor systems that collect observations in the visible and Near Infrared regions. The MODIS instrument on Terra has improved its data latency from 30 – 60 days shortly after launch to a matter of a few hours today. VIIRS is able to produce its Earth Data Record products within 20 minutes after data are received at the ground receiving station. The low latencies with VIIRS are mandated in the mission to support operational needs of the ocean navigations and weather forecasting communities. The rapid processing comes at the cost of lower accuracies from predicted ephemeris and ancillary data that may be 12 – 24 hours old. Science data products are then generated within 24 hours.

3.2. Case Study: Altimeter Observations

Altimeter observations are made by the Jason-1 sensor and in the future by ICESat-2. Certain products from Jason-1 can be made in just a few hours (Level-2 Sea Surface Height Anomaly can be made within 7 hours, for example). While ICESat-2 is estimating that their interpreted Level-3 products will be available within 30 – 60 days. The latencies for these missions are very different due to the different sensor architecture, requirements from the science team and user community, and due to the extremely high accuracy requirements for altimetry measurements. The missions in development provide projected latencies based on level 1 science requirements, updated with system timing information from test data with developmental algorithms in production systems and test environments. Actual achievable data latencies will not be fully characterized until the instruments are on orbit and the production systems are fully operational.

4. Direct User Feedback

For the purposes of this study, we created a professional review online to engage users’ interest in latency of data products. The review was entitled ‘ESD Decadal Survey System Trade Study’ (Questions can be found in Appendix A). The review focused on determining the data latency needs of the broader community. Each user has a set of requirements that determines the maximum length of time they can wait between acquisition of data and receipt of that data.

For some users, this is a very long, flexible period since their primary mission is to do basic research without a time constraint. Other users have operational responsibilities with hard timelines without any flexibility. Data that does not arrive before the deadline cannot be used within their system.

We distributed the survey to as many NASA satellite data users as possible by providing the on-
line link to the review to all known scientific and satellite user communities, and by requesting
others to provide information on as many users as possible. We estimate that the survey was
sent to 7000 people. The questions in the review are provided in Appendix A at the end of this
report. The review was open from April 1 through May 31. The total number of respondents
was 526, approximately 13 percent of the people who received the review.

4.1. Respondent Characteristics

The review began with questions on the respondent’s organization and requests for
information on what kind of satellite data is used. The first few questions were focused on the
type of work they did and institution respondents worked for. 34% of respondents worked for a
Federal institution, 31% for a University, and the remaining 35% worked for state and local
governments (13%), non-governmental organizations (10%) and private industry (12%). Of the
526 individuals who responded, 92% stated they used satellite remote sensing data in their
work. 56% of respondents used or analyzed directly science data products, whereas 33% used
satellite data to support decision-making or policy-related products after the data has been
processed in their system. 36% of the respondents stated they were a user of decision support
products and did not work directly with the satellite data themselves.

4.2. Latency Needs

Question 3 focused on asking which satellite data the community used. Table 3 shows the
satellite data products used by the respondents of the survey. The most commonly used
satellite data by the community who responded to the survey are the visible/near infrared
products derived from MODIS, VIIRS, AVHRR, POES/GOES, Landsat and ASTER. These datasets
are the oldest, most well-known and have the greatest number of applications.

<table>
<thead>
<tr>
<th>Instrument/Observation</th>
<th>Percent</th>
<th>Total responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>MODIS</td>
<td>65.33%</td>
<td>343</td>
</tr>
<tr>
<td>Landsat ASTER</td>
<td>15.81%</td>
<td>83</td>
</tr>
<tr>
<td>AVHRR GOES POES</td>
<td>12.95%</td>
<td>68</td>
</tr>
<tr>
<td>MODIS Rapid Fire</td>
<td>9.90%</td>
<td>52</td>
</tr>
<tr>
<td>VIIRS</td>
<td>8.95%</td>
<td>47</td>
</tr>
<tr>
<td>Altimetry</td>
<td>8.57%</td>
<td>45</td>
</tr>
<tr>
<td>Other</td>
<td>8.57%</td>
<td>45</td>
</tr>
<tr>
<td>TRMM 3B42</td>
<td>4.57%</td>
<td>24</td>
</tr>
<tr>
<td>High Res Commercial</td>
<td>3.62%</td>
<td>19</td>
</tr>
<tr>
<td>SeaWIFS</td>
<td>3.62%</td>
<td>19</td>
</tr>
<tr>
<td>AIRS</td>
<td>3.43%</td>
<td>18</td>
</tr>
<tr>
<td>TOMS OMI</td>
<td>3.05%</td>
<td>16</td>
</tr>
<tr>
<td>Aquarius</td>
<td>2.86%</td>
<td>15</td>
</tr>
<tr>
<td>CloudSAT</td>
<td>2.86%</td>
<td>15</td>
</tr>
<tr>
<td>AMSR/E</td>
<td>2.48%</td>
<td>13</td>
</tr>
</tbody>
</table>
Question 5 was focused on determining the timeliness of data, as defined by the period between data acquisition and product delivery to the user (Figure 1). The results show that

<table>
<thead>
<tr>
<th>Category</th>
<th>Percentage</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>JASON</td>
<td>2.48%</td>
<td>13</td>
</tr>
<tr>
<td>RADAR SAR</td>
<td>2.48%</td>
<td>13</td>
</tr>
<tr>
<td>SeaWINDS WindSAT</td>
<td>2.48%</td>
<td>13</td>
</tr>
<tr>
<td>Calipso</td>
<td>2.10%</td>
<td>11</td>
</tr>
<tr>
<td>GMES Products</td>
<td>2.10%</td>
<td>11</td>
</tr>
<tr>
<td>AURA MLS OMI</td>
<td>1.71%</td>
<td>9</td>
</tr>
<tr>
<td>SSMI</td>
<td>1.71%</td>
<td>9</td>
</tr>
<tr>
<td>QUIKSCAT scatterometer</td>
<td>1.52%</td>
<td>8</td>
</tr>
<tr>
<td>DMSP OLS</td>
<td>1.52%</td>
<td>8</td>
</tr>
<tr>
<td>ASCAT</td>
<td>1.14%</td>
<td>6</td>
</tr>
<tr>
<td>ICESat GLACE IceBridge</td>
<td>1.14%</td>
<td>6</td>
</tr>
<tr>
<td>SMOS</td>
<td>1.14%</td>
<td>6</td>
</tr>
<tr>
<td>GRACE Water</td>
<td>0.76%</td>
<td>4</td>
</tr>
<tr>
<td>MISR</td>
<td>0.57%</td>
<td>3</td>
</tr>
<tr>
<td>Uncategorized</td>
<td>10.86%</td>
<td>57</td>
</tr>
</tbody>
</table>

Figure 1. Data latency requirements. Total responses for this question was 766, since each user could provide up to 2 answers. Percentages given are the percent of responses in each category of the total (766). 525 respondents provided answers to the question.

40% of the respondents stated that their application requires data in less than 12 hours. This question permitted multiple responses since most respondents use multiple datasets. A significant number of respondents (50%) stated that product latency of over 24 hours was
acceptable. 10% stated that time is not a relevant factor for some of their analyses. This variability was expected, but it does indicate that if a mission planned very long latencies they would fail to serve a portion of the user community.

The survey then asked what the optimal latency was for the user’s application, regardless of whether or not it was actually available today. Figure 2 shows that far more respondents selected the ‘Less than 2 hours’ response than previously, showing that there are potentially far more users of low latency data products if these products were provided. A subsequent question asked “Will more timely delivery of satellite data improve the outcomes of your current analyses?” Of the 462 individuals who responded, 76% positively, reinforcing these conclusions.

![Optimal Latency](image)

Figure 2. Responses length of optimal data latency. 462 individuals responded to this question.

Comments on how low latency data can improve the respondent’s system or process, included the following:

- Improved active fire analysis, allowing for improved tactical response to containing brush and forest fires (28% of 127 comments on Question 8);
- Ocean and coastal applications, especially during fieldwork (12%);
- Weather Forecasting, which has strict latency requirements (10%);
- Operational data assimilation cannot be done without low latency data (3%);
- Park Monitoring of deforestation, poacher camps (3%); and
- Agriculture assessment provided within a day of observation can be more useful (2%).

Many comments said that NASA data provided at longer than 4 hour latency makes our data products less useful (16% of 127 comments), and often these products are excluded even if they could increase accuracy or detail of application because the products are not available
when they are needed.

Question 9 asked if the user would want higher latency products if the product had increased uncertainty such as reduced absolute geolocation, increased noise in a measurement, or other effects. The response to this question was that 54% of the respondents said no, they would use the standard science products with longer latency, and 46% said yes, they would be willing to use these lower-accuracy products. Question 10 asked the same question, but focused on data quality (data calibration, more atmospheric contamination, other issues), the response was 54% said yes, they would use these products available within a few hours of acquisition, and 46% said no, they would use the standard science product available with existing latency.

Comments under Question 9 included the following, which are reasons why higher noise/lower quality are acceptable:

- “Observations are always useful if the error is less than the next-best estimate, usually a model prediction”;
- “If the uncertainty is properly characterized, we can account for it in our process/system”;
- “For VIS/NIR and temperature products, low latency products have high accuracies that are acceptable for most users. This is not the case for other satellite products, in particular Altimetry products that require very precise observations”; and
- “Provided more accurate data was available later, as with LANCE-MODIS and standard MODIS products”.

Many respondents stated that the need for accuracy and timeliness vary for each product and use, and thus the nearly equal split in the answers received in Questions 9 and 10 is understandable. For most LANCE products, the difference between their products and the standard science products are small and well characterized, thus the trade-off for these users is considered acceptable. Since over half of all respondents to this survey use LANCE products, this trade-off is seen as one that is acceptable for these users. If lower latency data products are provided with careful validation and characterization of their accuracies, many new uses for these products will be found.

Question 11 states ‘If you were given a data product with a low confidence value (the product is not fully processed and some corrections still need to be made) quickly (within hours), would you be willing to use this product to meet your immediate needs, and then apply the same polished product, with greater confidence value, at a later time?’ For this permutation, 62% said yes and 38% said no. This question focuses on the value of NASA Missions providing incompletely calibrated or corrected Level 3 data within four hours of acquisition, to be replaced by completely calibrated and corrected Level 3 data 12 hours later, for example, or having an independent system like LANCE do such processing. 62% of respondents thought that they could use these products. 80% of respondents stated that they could replace the lower quality data with the final product at a later time, and redo their analysis with benefit to their decision-making, scientific or policy process or system (Question 12).
Figure 3 shows the impact of low latency data on the respondent’s institution or to society. Multiple responses were permitted for this question; the total number of responses was 703, where most respondents provided more than one answer. Comments under Question 13 included many references to the potentially improved weather forecasts that could be obtained if many more NASA data products could be assimilated into weather forecasts. Currently only a few products can be used because of the longer latencies for science data products. Extensive research has shown the large societal value of accurate forecasts of extreme events through weather models (Katz and Murphy, 2006, Emanuel et al., 2012).

![Figure 3. Value added due to low latency of NASA satellite data](image)

Other comments provided on Question 13 mention the life-saving value of accurate ice maps for ocean navigation, accurate and detailed fire line maps for wildfire response, and accurate air quality forecasts using atmospheric and temperature data for cities. These are just a few examples of the importance to our society of operational systems that need low latency data.

Question 14, the final question of the survey, asked people to describe in their own words how and why rapid delivery of satellite data products would improve their work. Responses varied in detail and, if not explicitly stated by the respondent, it was left to our discretion to decide what in the responses corresponded to answering *why* and *how*. The analysis of the responses consisted in manually matching each response to categories. A total of 33 categories were used to map responses to either *how* or *why* answers. See Appendix B for response characteristics.

We analyzed responses from 323 (62%) respondents to Question 14; 12 (3.7%) respondents provided no clear answer to either *how* or *why*. Responses could be matched to a *how* answer in 304 (94.1%) cases and to a *why* answer in 190 (58.8%). Table 4 shows, for example, that lower latency will improve the work of 34 (10.8%) of respondents through improved forecasting and model simulation from rapid delivery of satellite data; 18 (34%) stated it would benefit
their target audience.

Table 4. Number (%) of respondents whose answer could be matched to how \((n=315)\) and why \((n=189)\) categories

<table>
<thead>
<tr>
<th>Category</th>
<th>Number (%) how respondents</th>
<th>Number (%) why respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Improved Forecasting and Model Simulation</td>
<td>34 (11.2)</td>
<td>5.3 (10)</td>
</tr>
<tr>
<td>Provide early warning</td>
<td>10 (3.3)</td>
<td>4.2 (8)</td>
</tr>
<tr>
<td>Improved decision making process</td>
<td>23 (7.6)</td>
<td>11.1 (21)</td>
</tr>
<tr>
<td>Benefit target audience</td>
<td>4 (1.3)</td>
<td>17.9 (34)</td>
</tr>
<tr>
<td>Improved and/or faster results</td>
<td>8 (2.6)</td>
<td>0.5 (1)</td>
</tr>
<tr>
<td>Accuracy and/or quality is important</td>
<td>5 (1.6)</td>
<td>7.9 (15)</td>
</tr>
<tr>
<td>Current data system works</td>
<td>9 (3.0)</td>
<td>0.5 (1)</td>
</tr>
<tr>
<td>Enhanced rapid response capacity</td>
<td>14 (4.6)</td>
<td>9.5 (18)</td>
</tr>
<tr>
<td>No clear answer</td>
<td>12 (3.7)</td>
<td>0.0 (0)</td>
</tr>
<tr>
<td>Reduced impact/mitigation</td>
<td>3 (1.0)</td>
<td>2.1 (4)</td>
</tr>
<tr>
<td>Support field work</td>
<td>20 (6.6)</td>
<td>6.8 (13)</td>
</tr>
<tr>
<td>Improved spatial and/or temporal coverage</td>
<td>4 (1.3)</td>
<td>0.5 (1)</td>
</tr>
<tr>
<td>Able to access data within a practical timeframe</td>
<td>6 (2.0)</td>
<td>0.5 (1)</td>
</tr>
<tr>
<td>Enhanced assessment and/or analysis</td>
<td>25 (8.2)</td>
<td>3.2 (6)</td>
</tr>
<tr>
<td>Within specific time requirement</td>
<td>25 (8.2)</td>
<td>0.0 (0)</td>
</tr>
<tr>
<td>Timely detection</td>
<td>8 (2.6)</td>
<td>0.5 (1)</td>
</tr>
<tr>
<td>Allow for provision of value added information</td>
<td>2 (0.7)</td>
<td>1.1 (2)</td>
</tr>
<tr>
<td>Cost-effective operations and/or competitiveness</td>
<td>0 (0.0)</td>
<td>1.1 (2)</td>
</tr>
<tr>
<td>Only source of adequate/needed information</td>
<td>5 (1.6)</td>
<td>1.1 (2)</td>
</tr>
<tr>
<td>Improved planning</td>
<td>3 (1.0)</td>
<td>2.1 (4)</td>
</tr>
<tr>
<td>Allow for provision of accurate and/or timely data</td>
<td>23 (7.6)</td>
<td>5.3 (10)</td>
</tr>
<tr>
<td>Enhanced operations management</td>
<td>7 (2.3)</td>
<td>5.8 (11)</td>
</tr>
<tr>
<td>Used in training</td>
<td>1 (0.3)</td>
<td>0.0 (0)</td>
</tr>
<tr>
<td>Allow for preventative actions</td>
<td>3 (1.0)</td>
<td>0.0 (0)</td>
</tr>
<tr>
<td>Enhanced situational awareness</td>
<td>24 (7.9)</td>
<td>5.8 (11)</td>
</tr>
<tr>
<td>Allow for consistency/validation</td>
<td>2 (0.7)</td>
<td>0.5 (1)</td>
</tr>
<tr>
<td>Used supplementary to surveillance/detection system</td>
<td>3 (1.0)</td>
<td>0.5 (1)</td>
</tr>
<tr>
<td>Other factor is important</td>
<td>5 (1.6)</td>
<td>1.1 (2)</td>
</tr>
<tr>
<td>Improved monitoring</td>
<td>16 (5.3)</td>
<td>2.6 (5)</td>
</tr>
<tr>
<td>Readily accessible data</td>
<td>4 (1.3)</td>
<td>0.5 (1)</td>
</tr>
<tr>
<td>Enhanced nowcasting</td>
<td>1 (0.3)</td>
<td>0.5 (1)</td>
</tr>
</tbody>
</table>

Question 14: In your own words, please tell us how and why rapid delivery of satellite data products will improve your work.

* Percentages are based on total sum of responses within why and how criterion

#### 4.3. How Low Latency Improves Results

Besides improved forecasting and model simulation, other respondents suggested that low latency data will improve their work through: enhanced assessment and/or analysis (8.2%), enhanced situational awareness (7.9%), improved decision making process (7.6%), by being able to provide accurate and/or timely data (7.6%), as well as by enhanced support of their field work (6.6%). More detail on specific user responses can be found in Appendix C of how and why...
reduced latency will contribute to societal benefit.

As can also be observed in Table 4, 25 (8.2%) respondents provided a specific time requirement to explain how their work would improve. Overall, there were 40 (12.4%) responses in Question 14 that defined a specific desired latency. Figure 5 shows that, just as in Question 6 (optimal latency), the majority of these respondents (22.5%) desire data in less than 2 hours. Other responses, not included in Figure 5, correspond to users who desire data ‘within hours’ or at a ‘scale of hours’, in near real time or real time, as well as “direct broadcast” or “direct real time” data.

![Desired Data Latency](image)

There were 55 (17%) respondents who explicitly mentioned either using or needing real time or near real time data. The majority of these respondents (21.8%) referred to enhanced situational awareness through real time or near real time delivery of data. For 20% of respondents improvement of their work through real time or near real time data will be in the form of support to field work and for 18.2% improved forecasting and model simulation.

**4.4. Why Low Latency Improves Results**

Of the 190 respondents answering why rapid delivery of satellite data will improve their work, the majority pointed to benefiting their target audience, 34 (18%) responses. Some responses that describe typical responses are:

- “We would be able to produce integrated analysis results more quickly and with less reliance on models to extrapolate to real time, hence benefiting our users community.”
- “For burned areas, people in the regions would usually ask for total area burned in hectares for their reporting purposes”
- “The U.S. Navy has global interests with a large number of operational users. If products are older than 3 hours, they will mostly refuse to use them.”
• “It would provide a quicker response to customers and make us, my group, appear more responsive”
• “Better and timely advice to fishermen.”

Overall, the target audience as shown in Figure 6, consisted mainly of the general public (e.g. “society”, “public”, “people in the countryside”) and operational users (e.g. “Antarctic aviator”, “operational marine users”, “offshore fishing community”).

![Target Audience](image)

**Figure 5. Target Audience in why responses**

Other major reasons for rapid delivery of data included improved decision making process (11%), enhanced rapid response capacity (9%), fieldwork support (7%), and enhanced situational awareness (6%).

For 15 (8%) respondents, accuracy and/or quality was as important or more important than time. The following are some of the reasons given:
• “Accuracy is more important to me that time of delivery. The current system works for me.”
• “We are assembling a climate record. Data record length and fidelity are more important than rapid delivery.”
• “Weekly altimetry data with good spatial resolution are critical for decision making. Take the BP oil spill as an example.”
• “Access to near real time products is already good enough. I would like standard products with quality control and documentation in a shorter time frame-- a few weeks to a few months depending on complexity of data.”

Of the 55 (17%) respondents, mentioned in the previous section, who explicitly stated either using or needing real time or near real time data, 10.9% said it would benefit their target audience. Another 10.9% explained that rapid delivery would enhance their situational awareness.
Out of the 323 people that provided an answer to Question 14, 17% explicitly mentioned either using or needing real time or near real time data. The majority of these respondents referred to needing the data for strategically coordinating, directing, planning and making decisions during fieldwork (24%), e.g. research cruise support, research aircraft flights, and adaptive sampling from the ocean. Four specific references were made to latency requirements for cruise support, as summarized here:

1. Real time with less than 4 hours without the need for high accuracy until months later
2. Near real time: 24 hour delay for sub-mesoscale in situ sampling; shortest possible delay (i.e. hours), if there are few usable images during a cruise
3. Real time on a semi-daily basis
4. Day to day mission based on real time satellite imagery

Another 13% of respondents stated needing real time or near real time data for surveillance or monitoring, e.g. monitoring of illegal forest activity, pollution monitoring, tracking fires. Indications as to why rapid delivery would improve the respondent’s work in these cases included increased transparency of issues, broader community/institutional participation in monitoring and reporting, and better temporal resolution. In the case of tracking fires, “almost” real time was defined as less than 2 hours.

5. Discussion

Direct broadcast is the most rapid way to get satellite data from a sensor to the ground processing system. To achieve low latencies, direct broadcast is not enough, however. Significant investment at the user facility to accommodate rapid data processing of higher level data products is also necessary. Systems such as LANCE fill an important gap between the data available in standard products and the needs of those users who require data quickly. There are significant possibilities for expanding the pool of satellite data users by providing data with shorter latencies. Matching the needs identified by the review to the actual latencies available requires understanding the explicit data needs of different user communities. The responses from the questionnaire, 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. In Table 6 below some latencies have been described by product rather than by mission. The same product can be made from multiple satellite systems and the users will have the same latency requirements regardless of the sensor used to make the measurements.
Table 6. Latencies by product from NASA data production systems including LANCE (excludes possible latencies from direct broadcast systems).

<table>
<thead>
<tr>
<th>Product</th>
<th>Source Sensor(s)</th>
<th>Provided Latency</th>
<th>Desired Latency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surface Reflectance (VIS/NIR)</td>
<td>MODIS, VIIRS, AVHRR, Landsat</td>
<td>1 - 5 hours</td>
<td>1 - 5 hours</td>
</tr>
<tr>
<td>Vegetation Indices</td>
<td>MODIS, VIIRS, AVHRR</td>
<td>3 - 12 hours</td>
<td>12 - 24 hours</td>
</tr>
<tr>
<td>Atmosphere (cloud, aerosols, ozone, etc.)</td>
<td>MODIS, VIIRS, CALIOP</td>
<td>2 - 12 hours</td>
<td>1 - 5 hours</td>
</tr>
<tr>
<td>Fire Detection Products</td>
<td>MODIS, VIIRS, Landsat</td>
<td>2 - 12 hours</td>
<td>1 - 5 hours</td>
</tr>
<tr>
<td>Inundation</td>
<td>JASON, MODIS</td>
<td>1 - 24 hours</td>
<td>3 – 36 hours</td>
</tr>
<tr>
<td>Temperature (LST)</td>
<td>MODIS, VIIRS, AVHRR, Landsat</td>
<td>2 - 4 hours</td>
<td>1 - 5 hours</td>
</tr>
<tr>
<td>Sea Ice</td>
<td>MODIS, ICESat (ICESat-2)</td>
<td>2 - 4 hours</td>
<td>1 - 5 hours</td>
</tr>
<tr>
<td>Winds</td>
<td>SeaWINDS, WindSAT, MISR</td>
<td>1 - 3 hours</td>
<td>1 - 5 hours</td>
</tr>
<tr>
<td>Soil Moisture</td>
<td>ASCAT, SMOS, (SMAP)</td>
<td>3 days to 3 weeks</td>
<td>1 - 5 hours</td>
</tr>
</tbody>
</table>

In this study, we have identified three basic “classes” of latency needs from the user communities that were polled:

- Operations: need data in less than three hours
  - Numerical weather forecasting
  - Disasters management
  - Field campaigns
- Near real time: need data within a day
  - Strategic allocation of resources for forest/wildfire management
  - Environmental monitoring
    - Agriculture – food security
    - Human health – air quality, conditions for propagation of disease vectors
  - Field campaigns
- Research and science: need best quality data, time/latency is less relevant

We recognize that there are many other uses that fit within these classes that are not listed, or who were not reached to respond to our survey.

Research science is well supported through the existing systems and management structures. Support for both operations and near real time applications has been evolving over the last decade. To adequately support operations over large geographic regions will require a sensor
system to reduce time to downlink to either use TDRS or an expanded network of ground stations or perhaps both. We must couple these rapid downlinks with rapid product processing in systems like LANCE in order to adequately meet the needs of the near real time users.

Reducing the latency for real time operations is largely a hardware issue. To reduce the latency from 3 hours to 1 hour would require use of TDRS, an additional ground receiving station, 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 L0 data to a usable product (Level 1 or higher) 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 is largely an algorithm level issue. The primary delay in processing L2 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 the user communities have not reached consensus on what is needed in terms of latency. As applications continue to develop we need to identify specific user needs and processes that will improve from decreased latency. Experience with MODIS 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 would broaden the NASA data user community into new areas not yet known.

6. Conclusions

This study was undertaken to understand what the user community needs for data from missions with low latency are. We have found that there are a number of communities which 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 that is outside of the standard mission 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. This guidance should address
expectations on the 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 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 the needs are based on specific data products, which may come from multiple missions with different designs and product requirements. 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 production order of the 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.

Third, clear guidance is needed from NASA headquarters regarding the use of direct broadcast on future 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 disasters 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 latency at the product level and linking new missions to heritage observations, new applications will be developed. 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


Appendix A.

Questions on the Professional Review

1. What type of institution do you work for?
   ____Private Industry
   ____NGO (non-government organization)
   ____State/Local Government
   ____Federal Government
   ____University
   (Optional) Please tell us the name of the organization you work for. If you would like us to contact you, please include your email and contact information.

2. Do you use remote sensing data?
   ____Yes
   ____No

3. What type of satellite instrument products do you use? (eg. Terra MODIS MOD09). If you do not know the name of the data you use, please provide us with a brief description of the data source(s) you rely on.

4. How would you characterize your data use?
   ____Science Data Development: user that creates or analyzes science data products.
   ____Post Development: users that create decision/policy products, decision support products after raw data has been processed.
   ____Applications User: a consumer/user of decision support products. No data processing is conducted but rather the products are brought into an existing system or process.
   ____Other (Please specify)

5. For purposes of your work, how timely are your data requirement needs? Here, we define timeliness/latency as the period between data acquisition by the instrument and the delivery of the data product to the user. Note: Users with both real-time and long-term data requirement needs may select up to 2 answers.
   ____In less than 2 hours
   ____In less than 4 hours
   ____Within 12 hours
   ____24-36 hours
   ____36-96 hours
   ____1 week
   ____Time is irrelevant

6. If you could order a satellite data product that you use to best suit your work/responsibilities, what is the optimal delivery time you want for your data?
   ____In less than 4 hours
7. How often do you need data (frequency of data acquisition)?
   ____ Several times a day
   ____ Daily (early morning hours in order to use data for daily reports)
   ____ Daily by close of business
   ____ Several times a week
   ____ Weekly
   ____ Monthly
   ________________________________ Other (Please specify)

8. Will more rapid delivery of satellite data (decreased latency) improve the outcomes of your analysis (decision making process, product delivery to other operational users, faster model outputs critical to operational needs)?
   ____ Yes
   ____ No
   Please give a brief explanation (optional)

9. If producing data products faster resulted in an accuracy trade-off, (increased uncertainty such as reduced absolute geolocation, increased noise in a measurement, etc) while maintaining the quality science algorithms and quality checks to produce mission validated products, would you still want the data faster? (Can we give an example of the reduction that can be a representative sample?)
   ____ Yes
   ____ No
   Please give a brief explanation (optional)

10. If a data product was available quickly (within hours of acquisition) but that same data product was of lesser quality (e.g., noisier, only quick-look calibration, other potential issues), would you be willing to use this product to meet your immediate needs?
    ____ Yes
    ____ No

11. If you were given a data product with a low confidence value (meaning the product is not fully processed and some corrections still need to be made) quickly (within hours), would you be willing to use this product to meet your immediate needs, and then apply the same polished product, with greater confidence value, at a later time?
    ____ Yes
    ____ No
12. This question references your answer to Question 11. Would you replace the low confidence value product at a later time with a more refined higher-confidence version of the same product and redo your analysis?
   ____ Yes
   ____ No

13. What is the value added to your work for having more rapid delivery of satellite data products?
   ____ Direct societal impact with faster delivery of modeled products
   ____ Institutional benefit
   ____ Financial benefit to institution/society
   ____ Emergency decision-making benefits
   ____ Time is not an issue.

14. In your own words, please tell us how and why a lower latency will improve your work (1000 characters or less).
Appendix B.
Question 14 Response Characteristics

Responses to Question 14 followed a “free list in context” type of text, which describes a text with either limited response length consisting of a concise “list” format, or a short narrative that allows respondents to “vent” or explain themselves (Jackson and Trochim, 2002). Examples of both types of responses are given next.

Concise “List” Format
While some responses consisted of only one phrase or sentence, such as the two following responses:

<table>
<thead>
<tr>
<th>ID</th>
<th>Date/Time</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>126</td>
<td>May 4, 2013 3:10 AM</td>
<td>watch the weather</td>
</tr>
<tr>
<td>147</td>
<td>May 2, 2013 6:34 PM</td>
<td>Time is not a factor.</td>
</tr>
</tbody>
</table>

Others provided a list to answer the question:

```
14   May 10, 2013 7:26 PM
"For the following reasons: - early morning purpose
- accurate and timely information providing
- real time training data
How:
- through mails alerts
- information on the availability of data through magazine, scientific papers etc."
```

Response No. 14 is separated into why and how sections; however, most responses do not explicitly provide answers to these questions. In general, if not explicitly stated by the respondent, it was left to our discretion to decide what in the responses corresponded to answering why and how. Removing context from concepts in sparse answers is also typically a difficult endeavor (Jackson and Trochim, 2002).

Short Narrative
Short narratives typically vary in length from a few phrases to a couple of paragraphs (Jackson and Trochim, 2002:308), as shown next.

```
133  May 3, 2013 10:28 AM  "We produce weather forecast products several time daily and have a very tight schedule for our forecasting system. We acquire observations from many sources which feed into our analysis of the state of the atmosphere. Reducing the latency of observations increases the number of observations used in our atmospheric analysis. This should, in general, improve the atmospheric analysis which should, in turn, improve the quality of our forecast products.

I answered ""No"" to Q11. We understand there is often a slight degradation on quality if latency is reduced. We need to analyse the effect of this reduced quality. If the quality is poor, or flagged as ""low confidence"" we may not use the data. If the reduction in quality is ""reasonable"" and the data still give a useful contribution to our atmospheric analysis, then we will use the data."
```

27
No. 133 indicates a specific timeframe, as well as introduces a negative correlation between latency and the number of observations used in atmospheric analysis leading to a positive correlation between the quality of atmospheric analysis and the quality of forecast products. The response also references the Yes or No Question 11.

If you were given a data product with a low confidence value (meaning the product is not fully processed and some corrections still need to be made) quickly (within hours), would you be willing to use this product to meet your immediate needs, and then apply the same polished product, with greater confidence value, at a later time?

The respondent links Question 11 with Question 14 and uses it as an opportunity to explain why a “No” answer was selected. Overall, answers to Question 14 align with what has already been established as characteristics of responses for open-ended questions: varying response domains and orientation within those domains, frequent or infrequent mention of topics, and a wide variety of concepts varying in frequency and detail (Jackson and Trochim, 2002; Langer Tesfaye, 2011). The following definitions are used for topics and concepts:

- **Topics**: “natural clustering of terms that people will have coherent and consistent opinions about” (Ingersoll, Morton, and Farris, 2013).
- **Concepts**: subsets of topics defined in context.

Other characteristics of the responses were if-then statements, positive or negative qualifications, and conditional judgments. Some examples are shown next:

- **If-Then**
  faster is not always better but if it is the only option then something is better than nothing
  For our forecast cycle, if the data are not available within 24 hours, then they are not used in the forecast.
  We do perform hindcasts frequency to exploit late arriving and higher quality datasets

- **Positive/Negative Qualifications**
  I am not using satellite data in daily operations, so rapid delivery is not necessary for me. Instead, I am using satellite data for climate research, and so I can wait for several weeks or months to add more accurate and high resolution data to my time series.
  Today my work time is not important since I work with in situ data and time series to validate products in our area. But, I’m sure that the private company, as well as public institutions will make greater use of remote sensing scenes for management and organization tasks if near real time data were available.

- **Conditional Judgments**
  The name says it all: "Daily Fire Alert". It’s not much of an alert if it arrives several hours after the satellite overpass.
  The ability to obtain satellite data product with less latency will result....is even a few hours old. If we could obtain data with latencies of less than an hour (for example), it would open up a great number of possibilities for additional application of NASA satellite data to operational end users.

Our response analysis was conducted by manually mapping each response (n=323) to a why or how category. Each category was derived successively from each of the responses; a total of 33 categories were derived in this manner, as show in the next table. Those responses not
matching to either why or how criteria, were mapped to a No clear answer category.

<table>
<thead>
<tr>
<th>ID</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Improved Forecasting and Model Simulation</td>
</tr>
<tr>
<td>2</td>
<td>Provide early warning</td>
</tr>
<tr>
<td>3</td>
<td>improved decision making process</td>
</tr>
<tr>
<td>4</td>
<td>Benefit target audience</td>
</tr>
<tr>
<td>5</td>
<td>Improved and/or faster results</td>
</tr>
<tr>
<td>6</td>
<td>Accuracy and/or quality is important</td>
</tr>
<tr>
<td>7</td>
<td>Current data system works</td>
</tr>
<tr>
<td>8</td>
<td>Enhanced rapid response capacity</td>
</tr>
<tr>
<td>9</td>
<td>No clear answer</td>
</tr>
<tr>
<td>10</td>
<td>Reduced impact/mitigation</td>
</tr>
<tr>
<td>11</td>
<td>Support field work</td>
</tr>
<tr>
<td>12</td>
<td>Improved spatial and/or temporal coverage</td>
</tr>
<tr>
<td>13</td>
<td>Able to access data within a practical timeframe</td>
</tr>
<tr>
<td>14</td>
<td>Enhanced assessment and/or analysis</td>
</tr>
<tr>
<td>15</td>
<td>Within specific time requirement</td>
</tr>
<tr>
<td>16</td>
<td>Timely detection</td>
</tr>
<tr>
<td>17</td>
<td>Allow for provision of value added information</td>
</tr>
<tr>
<td>18</td>
<td>Cost-effective operations and/or competitiveness</td>
</tr>
<tr>
<td>19</td>
<td>Only source of adequate/needed information</td>
</tr>
<tr>
<td>20</td>
<td>Improved planning</td>
</tr>
<tr>
<td>21</td>
<td>Allow for provision of accurate and/or timely data</td>
</tr>
<tr>
<td>22</td>
<td>Enhanced operations management</td>
</tr>
<tr>
<td>23</td>
<td>Used in training</td>
</tr>
<tr>
<td>24</td>
<td>Allow for preventative actions</td>
</tr>
<tr>
<td>25</td>
<td>Enhanced situational awareness</td>
</tr>
<tr>
<td>26</td>
<td>Allow for consistency/validation</td>
</tr>
<tr>
<td>27</td>
<td>Used supplementary to surveillance/detection system</td>
</tr>
<tr>
<td>28</td>
<td>Other factor is important</td>
</tr>
<tr>
<td>29</td>
<td>Improved monitoring</td>
</tr>
<tr>
<td>30</td>
<td>Readily accessible data</td>
</tr>
<tr>
<td>31</td>
<td>Enhanced nowcasting</td>
</tr>
<tr>
<td>32</td>
<td>Not applicable to current work</td>
</tr>
</tbody>
</table>
Appendix C.

Question 14 Response specifics.

Enhanced assessment and/or analysis

- “better real time differentiation between dust and smoke”
- “Analyze the variability of environment every day”
- “Improving the timeliness and quantity of SST observations that can be used in a blended Level 4 dataset”
- “Immediate analysis of algorithm or algorithm changes - Timely examples from current weather situations”

Enhanced situational awareness

- “We have to have direct real time imagery showing regional weather behavior to properly support aviation movement in the Antarctic...”
- “I work for emergency response and disaster reduction institutions, the faster we can get the data the faster we can know where the affected areas are and help them”
- “During high fire danger periods, the Aqua and Terra FIRMS data is invaluable for assessing the extent of a particular fire, or for determining the number of large fires which are burning within a particular region.”
- “We rely on MODIS hotspots, snow and ice, and LANCE imagery to gain a strategic-level sense of fire activity and landscape condition across the province.”

Improved decision making process

- “Conservation issues need to be addressed on a priority basis. Rapid delivery of satellite data improves the decision making process more efficient and timely deliverable.”
- “Faster decision making to notify role players and land owners of wild fires or at least potential wild fires...”
- “Making informed decisions as far as Park Management Initiatives are concerned.
- “Weekly altimetry data with good spatial resolution are critical for decision making. Take the BP oil spill as an example.”

Allow for provision of accurate and/or timely data

- “It can help me to deliver services to my clients in a timely manner thereby giving a competitive edge to my company.”
- “Reliable information would help better (accurate) news coverage and analysis of forest fires.”
- “I provide information about ongoing wildfires to the public. Fast and accurate information is critical.”
- “Better and timely advice to fishermen.”

Support field work

- “If I was involved in a field project, then quick access to the data could be helpful.”
• “Rapid delivery will quicken the ground level action”
• “Faster data availability enables us to route our ships navigating in ice more effectively and wisely which has direct impacts on ship and crew safety, fuel consumption and environment footprint.”
• “War fighters rely on rapid, reliable oceanographic optical properties to accomplish their mission safely and more efficiently.”

**Improved decision making process**
• “In training forecasters to use polar orbiting data, having the data as fast as possible will improve forecaster decisions.”
• “It will improve the response capacity to decision making”
• “Scope for operational usage in decision support is improved”
• “Faster delivery of decision-making-products, with a great influence over the field work.”

**Enhanced rapid response capacity**
• “because it’s the only way to get adequate information about bush fires in our area and only through that we are able to get in action to fight the fires in time and to be able to get emergency procedures in line”
• “It will allow us to respond to issues on time”
• “Disaster response/management imagery services would be improved upon.”
• “we are working on Environmental Emergency department and responses in Kuwait. Your data will be useful for immediate response of our Environmental Alerts”

**Field work support**
• “Support field experiments from the tropics to the Arctic in near real time, based on which the right decisions can be made to maximize science returns.”
• “We monitor ice shelf conditions in Antarctica, and occasionally support field parties with weather and sea ice conditions. Rapid delivery of satellite data products would improve their usefulness for both applications.”
• “oceanographic campaign survey in NRT”

**Enhance situational awareness**
• “To gain better, rapid idea of impact of changes in land cover on sites of conservation importance following detection of events.”
• “monitor fast changing phenomena in atmosphere, ocean, and land”
• “Rapid delivery of satellite products for near real-time monitoring complements the periodic remote sensing-based analysis of ecosystem condition and change...”