Automatic Aura Remote Sensing Satellite Data Acquisition System

Olivia Maftukhaturrizqoh[1], Budhi Gustiandi[1]

[1)]Remote Sensing Technology and Data Center, Indonesia National Institute of Aeronautics and Space (LAPAN)
Jalan LAPAN No. 70, Pekayon, Pasar Rebo, 13710 Jakarta Timur, Indonesia

e-mail: olivia.maftukhaturrizqoh@lapan.go.id

Abstract. Aura is one of National Aeronautics and Space Administration (NASA) Earth Observing System (EOS) satellite series. It measures aerosols, ozone, and other gases from space to gain knowledge about the chemistry of our atmosphere. Although it was launched on 15 July 2004, there is still no system available to acquire its data and products specifically for Indonesian region. Scientists usually download the data through Goddard Earth Science (GES) Data and Information Services Center (DISC) or NASA Langley Atmospheric Sciences Data Center manually if they need them. An automatic data acquisition system was established to obtain Aura satellite data in near real-time. The system includes storage management sub system that mimics the data management system in GES DISC. Selenium package in Python programming language was used to program the automatic acquisition through browser crawling and bash shell scripting was used to build the storage management sub system. Analysis of system time in acquiring the data will be done to ensure that the system fulfill the near real-time criteria. Further, analysis of the data volume will also be taken to calculate how many storage that will be required to store the data. In the end, the system will be integrated into low-resolution remote sensing satellite data processing system that is already available in Remote Sensing Technology and Data Center, National Institute of Aeronautics and Space (LAPAN).

1. Introduction

Aura (Latin for "breeze", formerly EOS CHEM) is one and the last of NASA's core satellites for Earth Observing System (EOS) program. It is dedicated to understand how chemistry composition changes in the Earth's atmosphere, especially in stratosphere and troposphere, by measuring ozone, aerosols, and key gases using innovative space technologies. The satellite was successfully launched from the Vandenberg Western Test Range about 15 years ago, exactly on 15 July 2004, using a Delta II 7920-10L rocket. The platform orbits the Earth at 705 km (438 miles) altitude in sun-synchronous polar orbit with a 98.2o inclination. It crosses the equator in ascending (north-going) mode at 1:45 p.m. (±15 min) and cycles the planet in 98.8 period (sixteen-day repeat cycle with 233 revolutions per cycle) [1].

There are four instruments that fly with the Aura platform, viz. the High Resolution Dynamics Limb Sounder (HIRDLS) [2], the Microwave Limb Sounder (MLS) [3], the Ozone Monitoring Instrument (OMI) [4] and the Tropospheric Emission Spectrometer (TES) [5].The platform and its instruments are shown in Figure 1. The instruments and their configuration were selected because they can make complementary measurements based on previous proven technologies so that they can bring new capabilities in measuring the Earth's atmosphere[6]. Summary of each instrument is listed in Table 1.

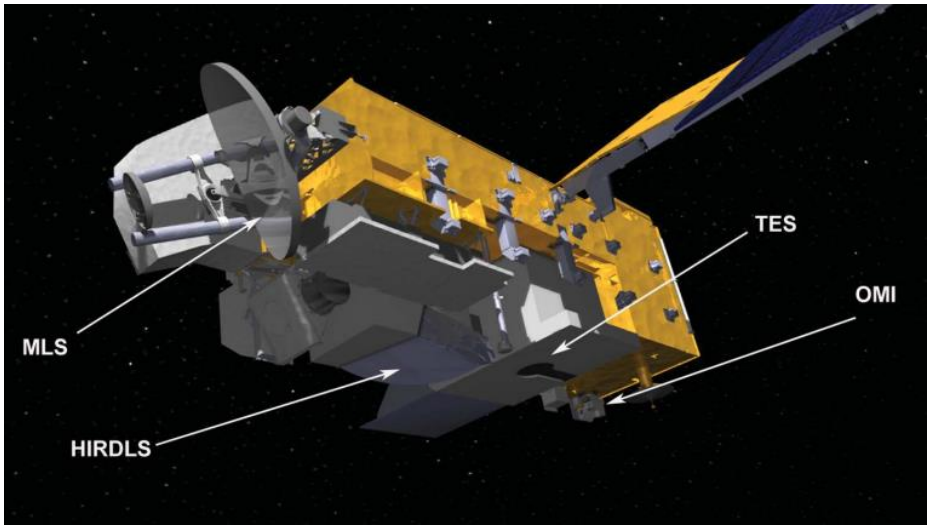

Figure 1. The Aura platform and its four instruments' location in a computer model that was made by Jesse Allen from NASA Earth Observatory[7].

Table 1. Summary of four instruments flying in the Aura satellite.

Instrument Name	Type	Spectral Characteristics	Constituents Measured
High Resolution Dynamic Limb Sounder (HIRDLS)	Limb infrared filter radiometer	6.2 to 17.76 μm	Profiles of temperature, O_3, H_2O, CH_4, N_2O, NO_2, HNO_3, N_2O_5, CF_3Cl, CF_2Cl_2, $ClONO_2$, and aerosols (2)
Microwave Limb Sounder (MLS)	Microwave limb sounder	118 GHz to 2.5 THz	Profiles of temperature, H_2O, O_3, ClO, BrO, HCl, OH, HO_2, HNO_3, HCN, N_2O, CO, and cloud ice (3)
Ozone Monitoring Instrument (OMI)	Hyperspectral nadir imager	270 to 500 nm	Profiles of O_3 and UV-B; Column O_3, SO_2, aerosols, NO_2, BrO, OclO, HCHO, and cloud top pressure [8]
Tropospheric Emission Spectrometer (TES)	Limb and nadir infrared Fourier transform spectrometer	3.2 to 15.4 μm	Profiles of temperature, O_3, NO_2, CO, HNO_3, CH_4, and H_2O (5)

The satellite instruments take data and store them on board until broadcast them to several major downlink facilities that are located in Poker Flat (Alaska), McMurdo Base (Antarctica), and Svalbard (Norway) via a highspeed data link [7]. Then, the data are sent to processing facilities at NASA's Goddard Space Flight Center (GSFC) through optical fiber network. After that, the data are processed into geophysical measurements by the instrument teams. Processed data are sent and store in NASA Data Active Archive Center (DAAC).

Aura data users can access the NASA DAAC through several channels. They can access HIRDLS, MLS, and OMI instruments data via the Goddard Earth Science Data and Information Services Center (DISC) (https://daac.gsfc.nasa.gov/) or the Giovanni system (https://giovanni.gsfc.nasa.gov/giovanni/)[9]. Daily OMI images can also be retrieved from the OMI Sulfur Dioxide Group (https://so2.gsfc.nasa.gov/). The users can access TES data through NASA Langley Atmospheric Sciences Data Center (https://eosweb.larc.nasa.gov/).

However, users has to check the availability of new data provided by the distribution systems manually. The users also has to do some several steps provided by the systems to download required data manually. Some data are updated several times daily and it will take a significant time for a user to check and download the new required data.

In this paper, we propose an automatic system to search for new required data and download them automatically if they are available. It is hoped that the proposed system can help users in their operational activities by efficiently reducing time required to do the search and download processes.

2. Method

Aura satellite was designed to have mission life for five years with operational goal of six years [10]. However, it has operated for almost 15 years. Therefore, the first step in designing an automation system to search and download its data or products is by ensuring which instruments that are still operational nowadays.

2.1. Aura Satellite's Operational Instrument

Soon after the HIRLDS instrument was activated after the satellite launch, something was blocking the optical path of the instrument so that it could only view a small portion of the Earth's atmosphere. Attempts to remove the blocking material were failed. Although the measurements were conducted with 80% blockage, the data or products are only available until 17 March 2008 in the NASA GSFC DISC system.

Similar problem was also found in the TES instrument. Ice was covering its detectors soon after the instrument was activated. Signs of bearing wear were also found after several months of operation. The instrument only works in nadir mode since then. NASA decommissioned the instrument operation on 31 January 2019 [11]. Therefore, the automation system will not focus to search and download the HIRDLS and TES instruments data or products.

On the other hand, there was no problem that has been found since the MLS and OMI instruments were activated. Both instruments are operating flawlessly and their data are currently available in the NASA GSFC DISC system. As a result, the automatic system will focus to data or products that are derived from these instruments.

The next step is to identify types of MLS and OMI data or products that are available in the NASA GSFC DISC system. Some data or products are produced from multiple instruments other than those that are included in the Aura satellite. The automatic system will only focus to data or products that are derived from the MLS and/or the OMI instruments.

2.2. Work of the System

The system was built for Windows environment using Python 3.7 programming language. In general, the automatic system was designed in three main modules of processing. They are search module, download module, and classification module. The modules are shown in Figure 2.

The search module consists of program routines with main purpose to search and list all available data or products from the MLS and/or the OMI instruments that are available in the NASA GSFC DISC system. The module crawls through the web site automatically, check the data availability updates, and list all the required links to be used as inputs to the next module (download module). Furthermore, the search module also limits the search criteria built on the web site to only occupy the Indonesian region. The module was built using Python version 3.7 and selenium web driver package inside it. The block diagram for the search module is illustrated in Figure 3.

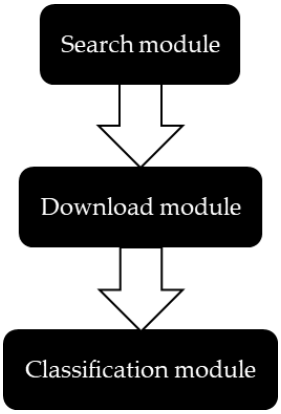

Figure 2. Three main modules that build the automatic system.

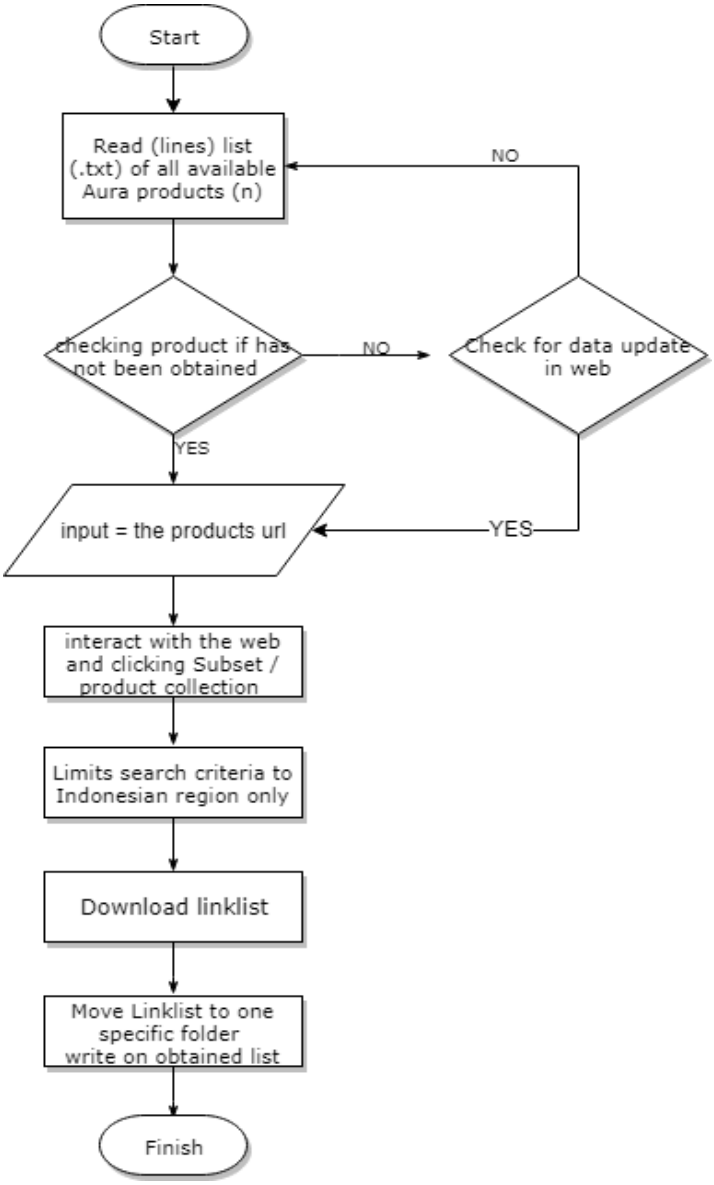

Figure 3. Block diagram of the search module in the automatic system

The download module filters information that is provided by the search module. Lists of data or products provided by the search module will be compared with lists of data or products that have been downloaded by the system. The block diagram for the download module is depicted in Figure 4.

The classification module classifies the downloaded data or products into different directories based on instrument, processing level, and type of product. This is intended to make users easier in searching the downloaded data or products in their storage system. In general, the directory structure that will be produced as the result of the classification module is shown in Figure 5.

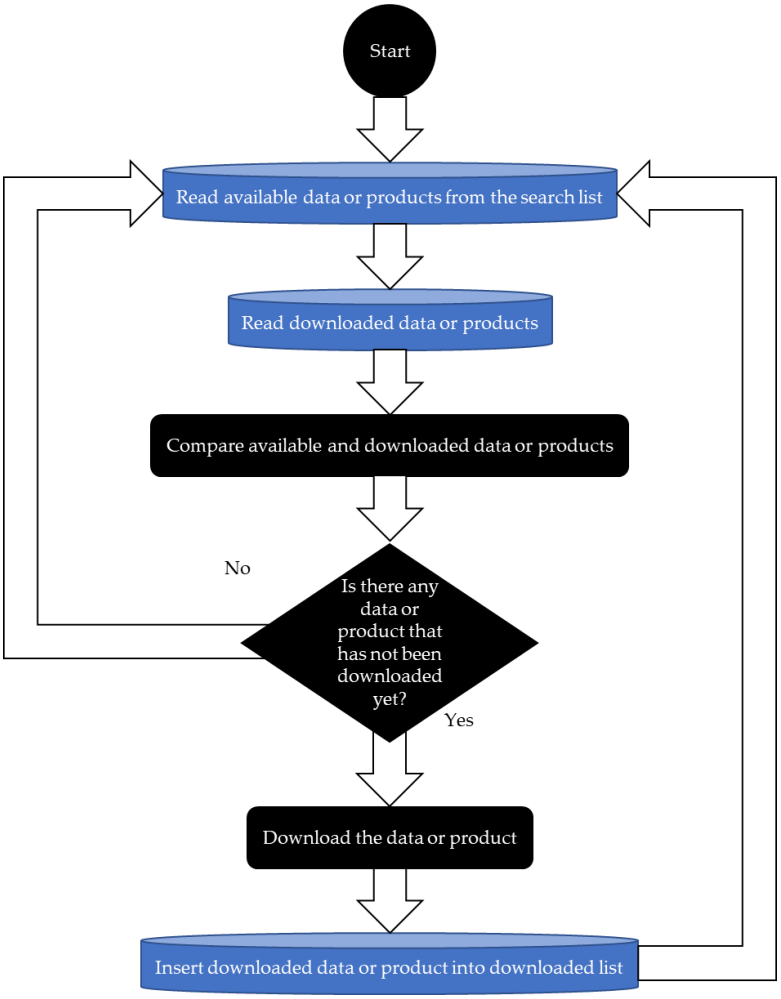

Figure 4. Block diagram of the download module in the automatic system.

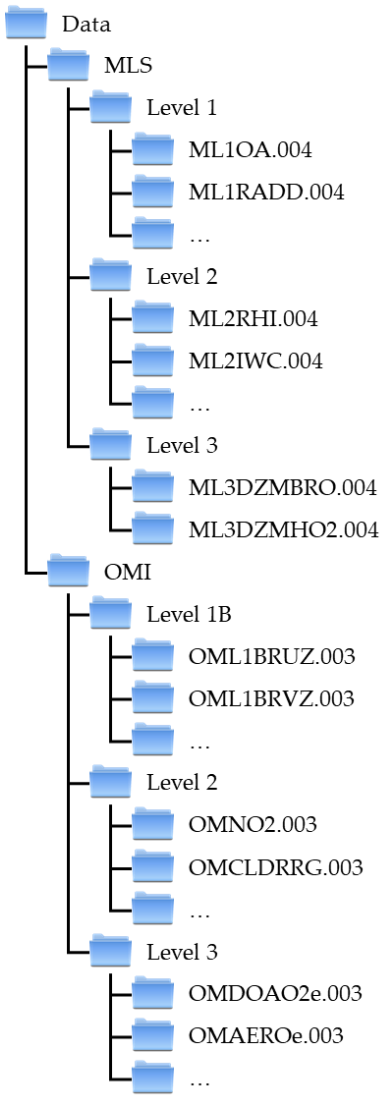

Figure 5. Directory structure as output of the classification module.

3. Results

There are 73 types of MLS data or products that are identified in the NASA GSFC DISC web site and there are 57 types of OMI data or products that are disseminated through the web site. The focus of the automatic system designed is to data or products that are still updated until now. Therefore, we only focus on 34 types of MLS data or products and 43 types of OMI products. A list of the data or the products is shown in Table 2.

Table 2. List of data or products that become the focus of the automatic acquisition system.

	MLS
Level 3	Bromine Monoxide (BrO) Daily 10degrees Lat Zonal Mean; Hydroperoxy (HO2) Daily 10degrees Lat Zonal Mean
Level 2	Relative Humidity With Respect To Ice; Cloud Ice Product; Near-Real-Time Water Vapor (H2O) Mixing Ratio; Diagnostics, Geophysical Parameter Grid; Near-Real-Time Sulfur Dioxide (SO2) Mixing Ratio; Water Vapor (H2O) Mixing Ratio; Near-Real-Time Ozone (O3) Mixing Ratio; Near-Real-Time Nitric Acid (HNO3) Mixing Ratio; Methanol (CH3OH) Mixing Ratio; Geopotential Height; Nitric Acid (HNO3) Mixing Ratio; Near-Real-Time Temperature; Ozone (O3) Mixing Ratio; Hydrogen Chloride (HCl) Mixing Ratio; Hypochlorus Acid (HOCl) Mixing Ratio; Chlorine Monoxide (ClO) Mixing Ratio; Near-Real-Time Carbon Monoxide (CO) Mixing Ratio; Methyl Chloride (CH3CL) Mixing Ratio; Bromine Monoxide (BrO) Mixing Ratio; Temperature; Nitrous Oxide (N2O) Mixing Ratio; Near-Real-Time Nitrous Oxide (N2O) Mixing Ratio; Diagnostics, Miscellaneous Grid; Carbon Monoxide (CO) Mixing Ratio; Hydrogen Cyanide (HCN) Mixing Ratio; Hydroperoxy (HO2) Mixing Ratio; Methyl Cyanide (CH3CN) Mixing Ratio; Sulfur Dioxide (SO2) Mixing Ratio
Level 1	Orbit/Attitude and Tangent Point Geolocation Data; Radiances from Digital Autocorrelators; Radiances from Filter Banks for THz; Radiances from Filter Banks for GHz
OMI	
Level 3	Ozone (O3) DOAS Total Column 1 day 0.25 degree x 0.25 degree; Multi-wavelength Aerosol Optical Depth and Single Scattering Albedo 1 day Best Pixel in 0.25 degree x 0.25 degree; Sulfur Dioxide (SO2) Total Column 1 day Best Pixel in 0.25 degree x 0.25 degree; TOMS-Like Ozone and Radiative Cloud Fraction 1 day 0.25 degree x 0.25 degree; Surface UVB Irradiance and Erythemal Dose Daily Global Gridded 1.0 degree x 1.0 degree; TOMS-Like Ozone, Aerosol Index, Cloud Radiance Fraction 1 day 1 degree x 1 degree; NO2 Cloud-Screened Total and Tropospheric Column Global Gridded 0.25 degree x 0.25 degree; Near UV Aerosol Optical Depth and Single Scattering Albedo 1 day 1.0 degree x 1.0 degree
Level 2	Nitrogen Dioxide (NO2) Total and Tropospheric Column 1-orbit Swath 13x24 km; Effective Cloud Pressure and Fraction (Raman Scattering) Daily Global Gridded 0.25 degree x 0.25 degree; Ozone (O3) Profile 1-Orbit Swath 13x48km; Formaldehyde (HCHO) Total Column Daily Global Gridded 0.25 degree x 0.25 degree; Ozone (O3) Total Column 1-Orbit Swath 13x24 km; Aerosol product Multi-wavelength Algorithm Zoomed 1-Orbit Swath 13x12km; Cloud Pressure and Fraction (Raman Scattering) 200-km swath subset along CloudSat track; Cloud Pressure and Fraction (Raman Scattering) 1-Orbit Swath 13x24 km; Sulphur Dioxide (SO2) Total Column Daily Global Gridded 0.125 degree x 0.125 degree; Cloud Pressure and Fraction (O2-O2 Absorption) Daily Global Gridded 0.25 degree x 0.25 degree; Surface UV Irradiance 1-orbit Swath 13x24 km; Near UV Aerosol Optical Depth and Single Scattering Albedo 1-orbit Swath 13x24 km; NO2 Total and Tropospheric Column Daily Global Gridded 0.25 degree x 0.25 degree; Sulphur Dioxide (SO2) Total Column 1-orbit Swath 13x24 km; Ozone (O3) DOAS Total Column 1-Orbit Swath 13x24 km; Formaldehyde (HCHO) Total Column 1-orbit Swath 13x24 km; Cloud Pressure and Fraction (O2-O2 Absorption) 1-Orbit Swath; Zoom-in Ground Pixel Corners 1-Orbit Swath 13x12km; Global Ground Pixel Corners 1-Orbit Swath 13x24km; Near UV Aerosol Optical Depth and Single Scattering Albedo Daily Global Gridded 0.25 degree x 0.25 degree; Surface UVB Irradiance and Erythemal Dose Daily Global Gridded 0.25 degree x 0.25 degree; Ozone (O3) DOAS Total Column Daily Global Gridded 0.25 degree x 0.25 degree; DOAS Total Column Ozone Zoomed 1-Orbit Swath 13x12km; Cloud Pressure and Fraction (O2-O2 Absorption) Zoomed 1-Orbit Swath 13x12km; Multi-wavelength Aerosol Optical Depth and Single Scattering Albedo 1-orbit Swath 13x24 km; Chlorine Dioxide (OClO) Total Column 1-orbit Swath 13x24 km; Multi-wavelength Aerosol Optical Depth and Single

	Scattering Albedo Daily Global Gridded 0.25 degree x 0.25 degree; Bromine Monoxide (BrO) Total Column 1-orbit Swath 13x24 km; Ozone (O3) Total Column Daily Global Gridded 0.25 degree x 0.25 degree; Near UV Aerosol Index, Optical Depth and Single Scattering Albedo 1-Orbit 13x24km
Level 1B	UV Zoom-in Geolocated Earthshine Radiances 1-orbit L2 Swath 13x12 km; VIS Zoom-in Geolocated Earthshine Radiances 1-orbit L2 Swath 13x12 km; UV Global Geolocated Earthshine Radiances 1-orbit L2 Swath 13x24 km; Solar Irradiances; VIS Global Geolocated Earth Shine Radiances 1-orbit L2 Swath 13x24 km

The automatic acquisition system that was designed successfully crawls all the data or the products listed in the Table 2 above. The data or the products that are successfully downloaded into storage system are classified based on their instrument, processing level, and type. The directories follows the convention that is described in Figure 5 above. Directory names use the code of each data or product as described in the NASA GSFC DISC web site.

4. Discussions

Although the automatic acquisition system runs successfully to search and download all required data or products from the NASA GSFC DISC web site, there were some problems faced during the development of the system. One of the primary problems was update in the web site. Since the search module uses crawl method that searches for certain tag in the web site structure, every change in the structure will affect the behavior of the module. For example, when the web site added an additional pop up menu about tour to the web site feature in its home page, the search module had to be modified to skip the pop up. Another example was when the results from the search box in the web site not only include results for an exact code of required data or products but also any relevant data or products, the module had to be modified again to exclude irrelevant data or products.

An analysis was conducted to know how the automatic system affects time required in obtaining an Aura satellite product compared with a manual process. Table 3 shows the result of this comparison and Figure 6 depicts the chart of the comparison.

Table 3. Comparison of manual and automatic time required to obtain Aura satellite product. Style and spacing

Checking steps run	amount of data	time required	
		manual	automatic
1 data without update	1	00:01:03	00:00:32
3 data without update	3	00:02:27	00:01:00
1 data update and download linklist	1	00:02:17	00:01:07
3 data update and download linklist	3	00:05:12	00:04:17

As shown in the chart below, the automatic system always requires less time than manual processes. Although the analysis process was only conducted using only several data and the time difference was not significant, the automatic process will help users when they have to deal with thousands of products. A manual process will take hours to check all the available data while the automatic system can do the same process in just under 60 minutes.

Additionally, analysis of the data volume also has been conducted to calculate how many storage that required to store the data. According to data volume calculation the total storage volume of Aura for level 1, 2, and 3 products which is acquired by this system is approximately 20 TB. While, data volume

of Aura OMI from 2004 to 2019 required about 16 TB storage. Accordingly, LAPAN has to provide approximately 36 TB to store Aura data from 2004 to 2019.

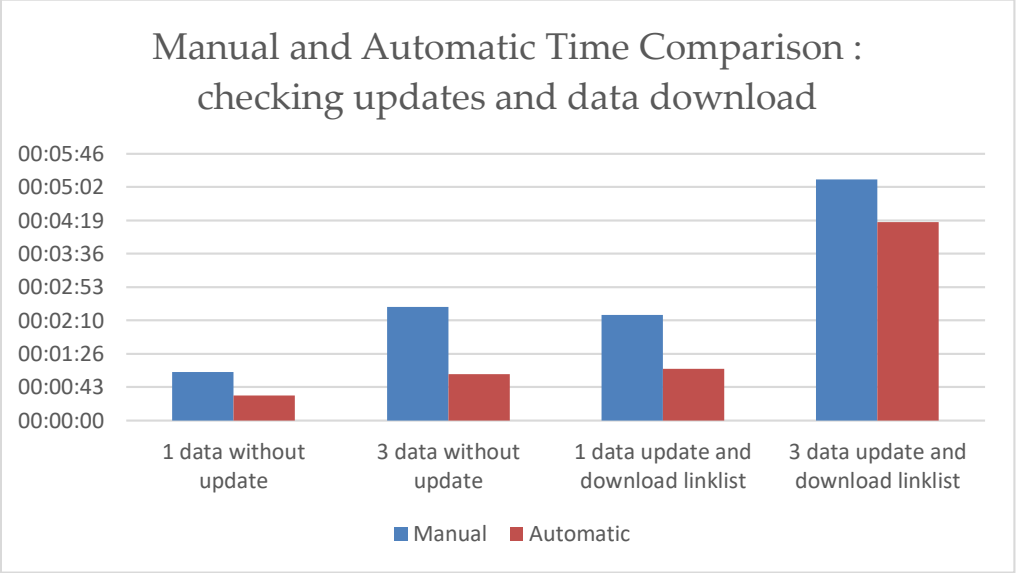

Figure 6. Comparison chart of the automatic system efficiency compared to manual process.

5. Conclusions

An automatic acquisition system was successfully built to obtain Aura remote sensing satellite data or products. The system has capabilities to search for required Aura MLS and OMI data or products from NASA GSFC DISC dissemination system, to download them, and to classify them based on their instrument, processing level, and types of data or products.

Currently, the automatic acquisition system was only designed to be run in a computer with Windows operating system. In the future, the system will be developed further so that it can run in a computer, especially a server, with Linux or Unix operating system. It is hoped that the system can be integrated into the current low resolution remote sensing data processing system that is available in Remote Sensing Technology and Data Center, LAPAN. The system design concept can also be implemented to be used in developing other remote sensing data acquisition systems that require search and download steps from a website-based dissemination system.

6. References

[1] NASA, *Our Changing Atmosphere: Discoveries from EOS Aura.* 2010.
[2] Gille, J., et al., *High Resolution Dynamics Limb Sounder: Experiment overview, recovery, and validation of initial temperature data.* Journal of Geophysical Research: Atmospheres, 2008. **113**(D16).
[3] Waters, J.W., et al., *The earth observing system microwave limb sounder (EOS MLS) on the Aura satellite.* IEEE Transactions on Geoscience and Remote Sensing, 2006. **44**(5): p. 1075-1092.
[4] Levelt, P.F., et al., *The ozone monitoring instrument.* IEEE Transactions on geoscience and remote sensing, 2006. **44**(5): p. 1093-1101.
[5] Beer, R., *TES on the Aura mission: Scientific objectives, measurements, and analysis overview.* IEEE Transactions on Geoscience and remote sensing, 2006. **44**(5): p. 1102-1105.
[6] Beer, R., T.A. Glavich, and D.M. Rider, *Tropospheric emission spectrometer for the Earth Observing System's Aura satellite.* Applied optics, 2001. **40**(15): p. 2356-2367.

[7] Schoeberl, M., et al., *Earth observing system missions benefit atmospheric research.* Eos, Transactions American Geophysical Union, 2004. **85**(18): p. 177-181.
[8] Ahmad, S.P., et al. *Atmospheric products from the ozone monitoring instrument (OMI)*. in *Earth Observing Systems VIII*. 2003. International Society for Optics and Photonics.
[9] NASA. *GES DISC*. 2004.
[10] Schoeberl, M.R., et al., *Overview of the EOS Aura mission.* IEEE Transactions on Geoscience and Remote Sensing, 2006. **44**(5): p. 1066-1074.
[11] Kusterer, J.M. *TES Data and Information* 2019.