Arctic Image Downloader

GEOM4009 - Team Report

Alana Beauchamp, Leo Keenan, Zack Ward

GEOM4009 - April 8th, 2025

Introduction

This project aimed to automate the process of collecting near real-time satellite imagery so our client can easily and instantly view it to determine weather conditions. Our client works for the Polar Continental Shelf Program (PCSP) under Natural Resources Canada (NRCan). NRCan is the federal department responsible for developing policies and programs related to Canada's natural resources. The PCSP is concerned with coordinating logistical support and providing related assistance for the purposes of advancing scientific knowledge and management of the Arctic region. Our client is a logistics coordinator who helps organize projects for people who want to work in the field in northern regions. A major part of our client's role is to enable research and field operations across the Arctic by providing aircraft support. The goal of this project was to develop a tool to monitor weather conditions to enable safe aircraft operations. This can be challenging due to frequent and severe changes in weather in the Arctic. They have found that the best way to monitor weather is through near real-time satellite imagery.

This project successfully developed a satellite image downloader tool so that PCSP can automatically retrieve near real-time satellite imagery (within the past 1-3 hours). The script downloads the most recent imagery from MODIS Aqua and Terra satellites within a set time period, and projects it to an appropriate coordinate reference system. It then composites images from both 250m and 500m spatial resolutions into a single multi-band image and saves it as a GeoTIFF that can be viewed in GIS software. The visualization parameters of the image will need to be adjusted manually in QGIS to achieve true colour and false colour visualizations, and we have provided instructions to guide the client through this process. This script is set to be run in the background at set intervals using a task scheduler so that the imagery can be instantly accessed when needed. This project involved several data types. The MODIS imagery files are downloaded from NASA EarthData as Hierarchical Data Format (HDF) files. When creating multiband composites, these files are converted to GeoTIFFs.

This report provides a comprehensive overview of our project. It addresses how we met our client’s needs, the workflow and methods used, challenges and limitations encountered and suggestions for future improvements.

Workflow

The script executes the following steps:

- Authentication:
    - Uses NASA Earth Data credentials to authenticate with Earth Data Server.
- Data search and retrieval:
    - Determine the time period for acquiring recent imagery.
    - Searches for imagery.From both Terra and Aqua satellites based on the provided area of interest. (AOI) and resolution requirements. (250 m for QKM bands, 500 m for HKM bands.
- Download and processing:
    - Downloads the identified HDF files.
    - Processes each file using GDAL to extract subdatasets.
    - Applies geospatial transformations to project the imagery into an Arctic polar stereographic coordinate system.
- Composite image creation:
    - Matches QKM and HKM files by satellite pass

    - Resamples HKM data to match the QKM dimensions.
    - Takes matching images from both resolutions and creates a single multi-band composite.
    - Applies scaling to enhance visualization and saves the result as a Geotiff.
- Clean up:
    - Removes temporary files and intermediate products to maintain an organized output directory.

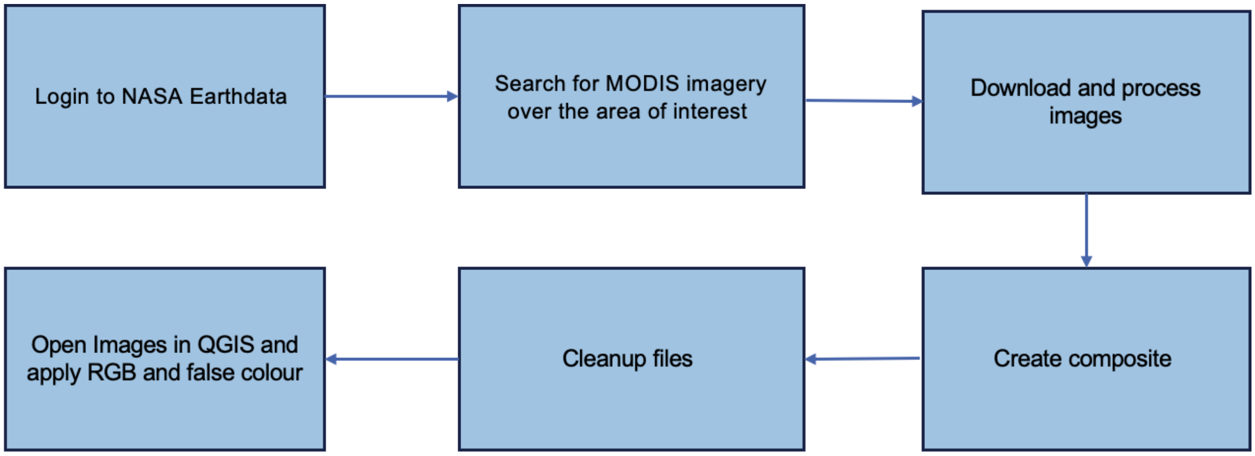
Login to NASA Earthdata
Search for MODIS imagery over the area of interest
Download and process images
Create composite
Cleanup files
Open Images in QGIS and apply RGB and false colour

Figure 1: Flowchart outlining the workflow of the image downloader tool

Documentation

Dependencies

The Arctic Image Downloader tool requires a standard desktop or server environment capable of running python-based geospatial processing. A windows operating system is ideal and is required for task scheduler. The required softwares include Python, Anaconda, and the ImageDownloader environment file that is provided on GitHub. This environment includes the following required packages:

- GDAL with HDF4 support
- GeoPandas
- Earthaccess
- NumPy
- Requests
- Shapely
- Rasterio

Installation/setup

A “README” file is available in the GitHub repository that explains the installation process for the required softwares and dependencies. A supplementary document is also provided on GitHub to guide the user through the installation process. This document also provides instructions on how to set up the script with the users NASA EarthData credentials, and how to configure the users desired aoi, and time range. Additionally, this document provides instructions on how to change visualization parameters in QGIS to achieve true colour and false colour visualizations, and how to set up the windows task scheduler to automate the script. A user guide with a step-by-step tutorial is also available in the repository to further support the user during setup.

Troubleshooting/FAQ:

1. **Why am I seeing "Warning 1: Unable to compute source region for output window 6836,0,2279,2433, skipping."**

   This is not a cause for concern. This warning message is just your system's way of handling edge cases where pixel mapping can't be properly calculated. The rest of the image should still display correctly, but some pixels near the edges might be cut off

2. **Why am I seeing "ERROR 1: Too many points (529 out of 529) failed to transform, unable to compute output bounds.**"

   This is not a cause for concern. This message just means that the transformation failed, this can occur when the input data does not have a proper coordinate reference system. This error can be ignored since the image will still display correctly. Some pixels particularly near the edges might be misaligned.

3. **Why am I seeing "No images available to process"**

   There are no images acquired in the set time interval. Try increasing the `hours_ago` parameter in the `get_modis_imagery` function call.

4. **Authentication issues**

   Double-check your NASA Earthdata credentials to ensure they are correct.

5. **GDAL Errors**

   Make sure you're using the conda environment, which has the proper GDAL configuration. GDAL requires HDF4 support to handle MODIS HDF files correctly. You can check this is installed by entering the following command in the Anaconda prompt:

   ```
   conda list gdal
   ```

6. **File Size Errors**

   The script uses BIGTIFF=YES, but extremely large areas may still cause issues.

Sphinx Documentation

A html document is available on the GitHub repository, which contains auto generated documentation from the source code. This documentation outlines the docustrings provided for each function within the script.

Discussion

Throughout the project, our team encountered several challenges that influenced the development and refinement of our image downloading tool. One major challenge we faced was that the bands required for our image visualization were on two different sensors. We overcame this by creating a multiband composite that combined the bands from both sensors into one image and saved it as a GeoTIFF file. Another significant challenge we faced was that the image files are quite large, and therefore, processing them took a long time. We attempted to speed up processing, but despite our efforts, we were unable to significantly reduce these times. We also faced challenges with applying the desired visualizations to the images in QGIS, particularly in determining appropriate min and max values. This made it difficult to produce visually consistent images. We were unable to set appropriate visualization parameters in the script, meaning this process has to be done manually in QGIS. This adds additional steps for the client and reduces the efficiency and overall ease of use.

One of the main limitations of our tool is that it only supports a single area of interest. This approach was taken due to file size and processing time constraints. Although our client initially requested coverage of all of Northern Canada, we determined that this was not feasible within the scope of our project. Additionally, this tool is limited in that it will only acquire high-quality imagery during the summer months, as MODIS is an optical satellite, meaning that it requires sunlight to take images. The lack of sunlight in the winter makes near real-time acquisition of imagery challenging in northern regions.

Throughout the project, we maintained contact with our client and had several interactions. We met with our client at the beginning of the project before we started to clarify expectations and outline our plans. We then met a few weeks later, about halfway through our project to provide an update on our progress and discuss our plans until the end of the term. At this time she had some feedback that we were able to incorporate into our project.

In the future, there are several improvements we recommend that will enhance the functionality of the tool. First, optimizing the processing time would help reduce the time required to handle image files. This could be done by scaling the files down to a smaller size, as the script currently produces a 32-bit float output, which is quite large. Another option to reduce file sizes could be to apply LZW compression or convert the files to virtual rasters. Additionally, implementing better error handling could provide more robust error detection and messaging, making it easier for the user to troubleshoot any issues. Another major improvement would be to incorporate automated visualizations for the images. This could be done by creating an additional function that extracts the minimum and maximum values of the bands and applies the true colour and false colour parameters. This would streamline the process by allowing the images to be opened with the correct visualizations applied. Expanding this tool to support a larger AOI could also be done to better support the clients needs and provide more comprehensive coverage. Finally, adding an auto cleanup feature that deletes files after a set amount of time would help manage storage.

Conclusion

This project successfully automates the retrieval and processing of near real time (NRT) MODIS Satellite imagery (Terra and Aqua) for weather monitoring in the Arctic. The tool's ability to composite images from different sensors into a single GeoTIFF output enhances the availability of critical weather data for field operations facilitated by the PCSP. While challenges remain in processing speed and scalability, the script provides a strong foundation for further developments and optimization to meet client requirements.

Acknowledgements

We would like to extend our thanks to the following individuals, organizations, and programs for their support in our project. Tanya Lemieux, our client, who brought this project to us and played a crucial role in communicating the needs and requirements of the project deliverables. Derek Mueller, who provided guidance in the course, from python workshops to support for our project, and offered valuable feedback on progress reports to help us improve our script. We are grateful to NASA for providing free and open access to MODIS satellite imagery, which was essential for our project, and to the earthaccess library for providing simplified access to EarthData. A special thanks to the GDAL developers for creating an excellent geospatial processing library, without which this project would have been much more challenging. Finally, we would like to thank Claude.ai for troubleshooting and debugging the code as we advanced through development.

Sources

- AI Assistance: Code development support from Claude.ai, an AI assistant by Anthropic.
- Licensing Information: MIT License details from Choose a License
- Documentation Tools: Sphinx auto-documentation framework for Python code documentation.
- Reference Implementation: Based on previous project code from GEOM4009/24WImage-Downloader
- Earth Access Library: Python library for simplified access to NASA Earth Science data from NASA EarthData Blog
- NASA Earth Data Catalog: Data catalog from NASA EarthData