**NASA DEVELOP National Program**



NASA Ames Research Center

*Spring 2017*

Chile Water Resources

Integrating NASA Earth Observations into the Google Earth Engine Platform to Enhance Drought Monitoring in Chile

 **Technical Report**

Final Draft – March 30, 2017

Garrett McGurk (Team Lead)

Billy Babis

Mariana Webb

Dr. Juan Torres-Perez, Bay Area Environmental Research Institute (Science Advisor)

Dr. Eduardo Bendek, NASA Ames Research Center (Science Advisor)

# 1. Abstract

Chile is characterized by extreme variations in climate. Situated along the Southwestern part of continental South America, Chile spans nearly 4,300 km from North to South, yet is only 350 km wide at its widest point.  This creates diverse climate regions across latitudes, ranging from the Atacama Desert in the northern part of the country, one of the driest places on earth, to the particularly wet Lake District in the southern region that averages 2,500mm of rainfall annually (Valdés-Pineda et al., 2014).  In recent years, Chile has experienced abnormal climate conditions as record droughts and rapidly receding glaciers place an added strain on a historically limited water supply (Bennett et al, 2016). With stakeholders from the water resources and agricultural communities relying on this limited reserve, the management of water in Chile has become increasingly relevant as the Ministry of Agriculture prepares for future climate variability and increasing demands.

This project provided the Chilean Ministry of Agriculture with a case study evaluating the effectiveness of incorporating new Earth observation datasets into the Google Earth Engine API (GEE) platform to enhance drought monitoring capabilities in Chile. To highlight the potential benefits of utilizing GEE to aid in drought monitoring and decision-making, the team developed a tool within the GEE platform that aggregates data from the Soil Moisture Active Passive (SMAP) radiometer, the Moderate Resolution Imaging Spectroradiometer (MODIS), and the Advanced Microwave Scanning Radiometer for EOS (AMSR-E) and Advanced Microwave Scanning Radiometer 2 (AMSR2) radiometers. Parameters derived from the remotely-sensed data include soil moisture (SM), snow cover (SC), and snow water equivalent (SWE). The advantage of hosting this data in the GEE platform is that it is a collaborative cloud-based analysis tool, which allows users to share, store, and perform analysis on large volumes of data. GEE has the potential to transform the way in which geospatial information is studied and disseminated, improving upon the accessibility and breadth of the data currently available.

**Keywords**

Water resources, agriculture, climate monitoring, SMAP, MODIS, AMSR-E, AMSR2

# 2. Introduction

* 1. ***Background Information***

Seasonal precipitation trends consist of the majority of rainfall events taking place May through September, with increased variability during El Niño Southern Oscillation (ENSO) years (Meza, 2013). Within the past few decades, Chile has experienced abnormal climate fluctuations in the form of record droughts, varied precipitation patterns, and increased rates of glacial retreat. Additionally, national measuring stations have observed a negative trend in annual precipitation in the densely populated south-central Chile, which is consistent with future climate scenarios modeled by General Circulation Models (GCMs), which predict a 10-20% decrease in precipitation levels by the end of the 21st Century (Valdés-Pineda et al., 2014).

With growing demands from water resources and agricultural stakeholders, accurate climate data has played an increasingly important role in the Chilean Ministry of Agriculture’s decision-making process. Chile sets aside almost 80% of consumptive water rights for agriculture and irrigation purposes. Moreover, by 2022, the government plans to increase the amount of irrigated land by 57% (Valdés-Pineda et al., 2014). Historically, Chile has relied on snowmelt and glacier-fed rivers as the primary water source. Increased demand, coupled with persistent drought conditions and abnormally low snowpack levels as a consequence of the rising 0°C isotherm, has resulted in the utilization of groundwater reserves (Vicuña et al., 2012; Valdés-Pineda et al., 2014). Hydrologic resources in Chile are unevenly distributed; although the wet southern regions generally have a surplus in the water budget, there is no infrastructure to facilitate the transportation of these resources to the arid north-central region where the domestic and agricultural demands are highest (Valdés-Pineda et al., 2014).

* 1. ***Project Partners & Objectives***

The Ministerio de Agricultura (Chile Ministry of Agriculture), is responsible for “promoting, guiding, and coordinating the agricultural activities of Chile” (Gobierno de Chile, 2017). In terms of water resource management, the Ministry is tasked with making policy decisions addressing the use of water resources, modifying the existing Water Code, and developing measures to combat drought. Current drought-monitoring information and tools are hosted on the Climate Data Library (CDL), and include the Combined Drought Index (CDI), the Standardized Precipitation Index (SPI), and a collection of precipitation and soil moisture measurements recorded by a network of stations across the country.

# As a result of increased strain on water resources, the Ministry of Agriculture was interested in exploring a more efficient and organized platform for data visualization and analysis. Currently, SC, SWE, & SM are measured at ground stations distributed throughout the country. These data lack spatial continuity and have not been kept up-to-date in the CDL. The potential strength of using GEE, as opposed to the current water resource monitoring tools, is the platform’s ability to host, process, and visualize large volumes of data in an efficient and accessible manner. Furthermore, GEE is a collaborative, cloud-based platform, which allows end-users to customize and share scripts. The overall objective for this project was to conduct a case study for evaluating the strengths and weaknesses of utilizing GEE as a drought-monitoring platform in Chile. To do this, the team needed to upload relevant SC, SWE and SM data to the GEE platform, develop algorithms to visualize and perform analysis on the data, develop a series of tutorials for end-users, and lastly communicate the benefits and challenges of using the GEE platform as a drought-monitoring tool to the project partners.

# This project directly addressed both Agriculture and Water Resources National Application areas. With much of the agricultural production taking place in the semi-arid central part of the country, agricultural stakeholders could benefit from a tool that provides accurate and accessible water resource information (Valdés-Pineda et al., 2014).

# 3. Methodology

***3.1 Data Acquisition***

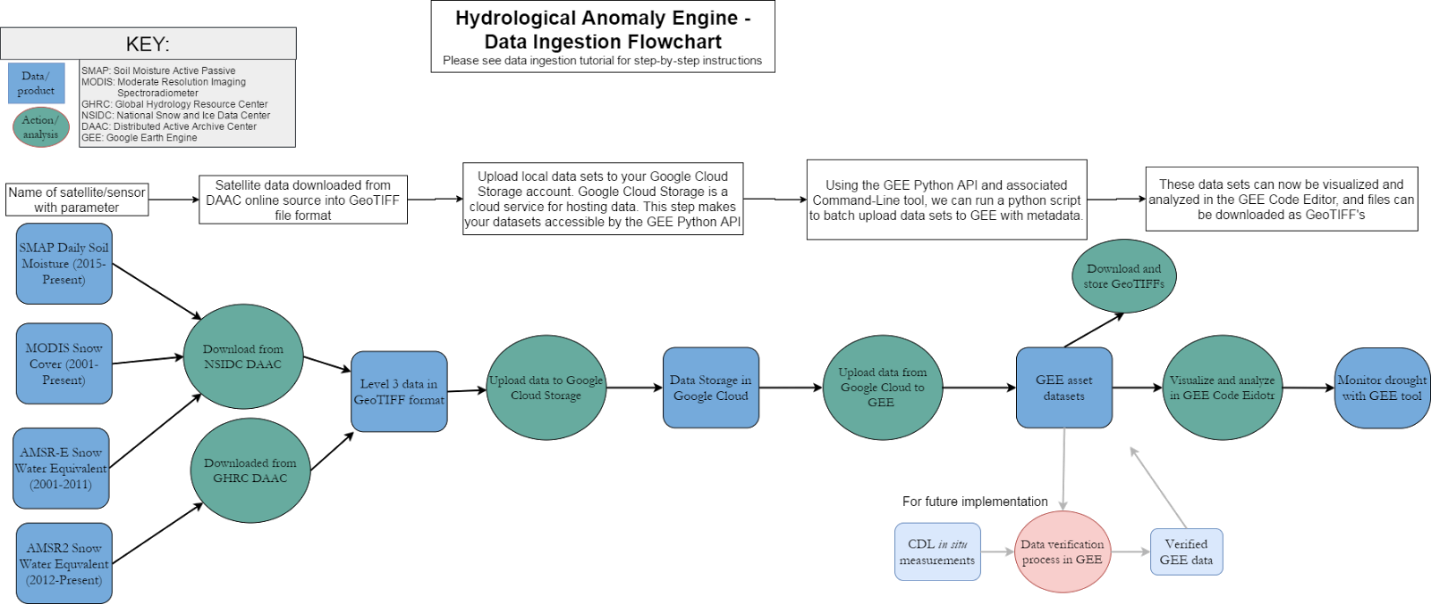
The project focused on SC, SWE, and SM as important hydrologic variables that could be used to perform quantitative analysis on drought conditions in Chile. The study period was January 2001 to February 2017. While SC data from MODIS, and SWE data from AMSR-E and AMSR 2 date back to January 2001, SM data from the SMAP radiometer are only available beginning January 2015. The geospatial platform used in the study was Google Earth Engine (GEE), an emerging cloud-based analysis tool, which allows users to import, store, and perform analysis on large volumes of data. GEE has the potential to transform the way in which geospatial information is studied and disseminated, improving upon the accessibility and breadth of the data currently available. While there have been a number of feasibility studies conducted to assess the usefulness of GEE, they are limited to dealing with datasets hosted on the GEE public data catalog (Dong et al., 2016; Johansen et al., 2015).

The public GEE data catalog hosts Landsat, MODIS, SENTINEL-1 and other popular datasets, which are made available to all users. The version 6 of MODIS/Terra Snow Cover Daily L3 Global 500m Grid became publically available on the GEE data catalog February 2017, which was then integrated into the project. The SM and SWE datasets were not yet available in the GEE data catalog, which required the team to download the datasets from the Distributed Active Archive Center (DAAC) and uploaded them to GEE. The AE\_DySno AMSR-E/Aqua Daily L3 Global Snow Water Equivalent EASE-Grids, Version 2 data (2001 to 2011) was reprocessed and downloaded in GeoTIFF format from the Earthdata Search Tool. SWE data from the AMSR2/JAXA GCOM W1 Daily L3 Snow dataset (2012 to 2017) was retrieved from the Global Hydrology Resource Center (GHRC) DAAC (https://ghrc.nsstc.nasa.gov/home/). This dataset lacked a defined extent, which required reprocessing these GeoTIFFs to assign extent values using the Geospatial Data Abstraction Library (GDAL). Lastly, the SM data, SPL3SMP\_E SMAP Enhanced L3 Radiometer Global Daily 9 km EASE-Grid Soil Moisture, Version 1 (2015 to 2017), was reprocessed and downloaded in GeoTIFF format from the Earthdata Search tool.

***3.2 Data Processing***

The data was uniformly downloaded and processed as L3, which provided consistency across the entire data set for spatial analysis. The data was available in .h5 and .hdf formats directly from the Earthdata Search tool, although it required the use of the HDF-EOS to GeoTIFF converter (HEG) tool to reprocess the data as GeoTIFFs. To filter out granules with high cloud cover, an Fmask was run on the MODIS SC data in GEE after being uploaded.

The first step was to create a shared asset repository to host the datasets in Google Earth Engine. This was accomplished by creating a Google Group (chile-water-resources@googlegroups.com) and requesting the GEE team at Google to create a shared asset repository named Chile-water. This allowed users with proper credentials to access this folder to upload and access satellite data. The GEE interface allows for import of local GeoTIFF files individually. However, due to the volume of the datasets, manually importing each individual file is impractical. This problem was addressed by using the publicly available GEE script library on GitHub, Google Earth Engine Batch Asset Manager (GEEBAM) (https://github.com/tracek/gee\_asset\_manager), which does allow for batch upload of local GeoTIFF files, although this tool did not automatically load the GEE asset with date metadata at the time of this project.



**Figure 1.**Data ingestion pipeline.

The GEE Python API and Command-Line Interface (CLI) tool allow for programmatic upload of files within a loop to automatically upload files in batch. Neither plug-in supports local GeoTIFF files, but the GEE CLI tool can upload files from Google Cloud Storage (GCS) to GEE Assets. Thus, local GeoTIFF datasets needed to be uploaded to a new GCS account to be accessible by the GEE CLI tool.

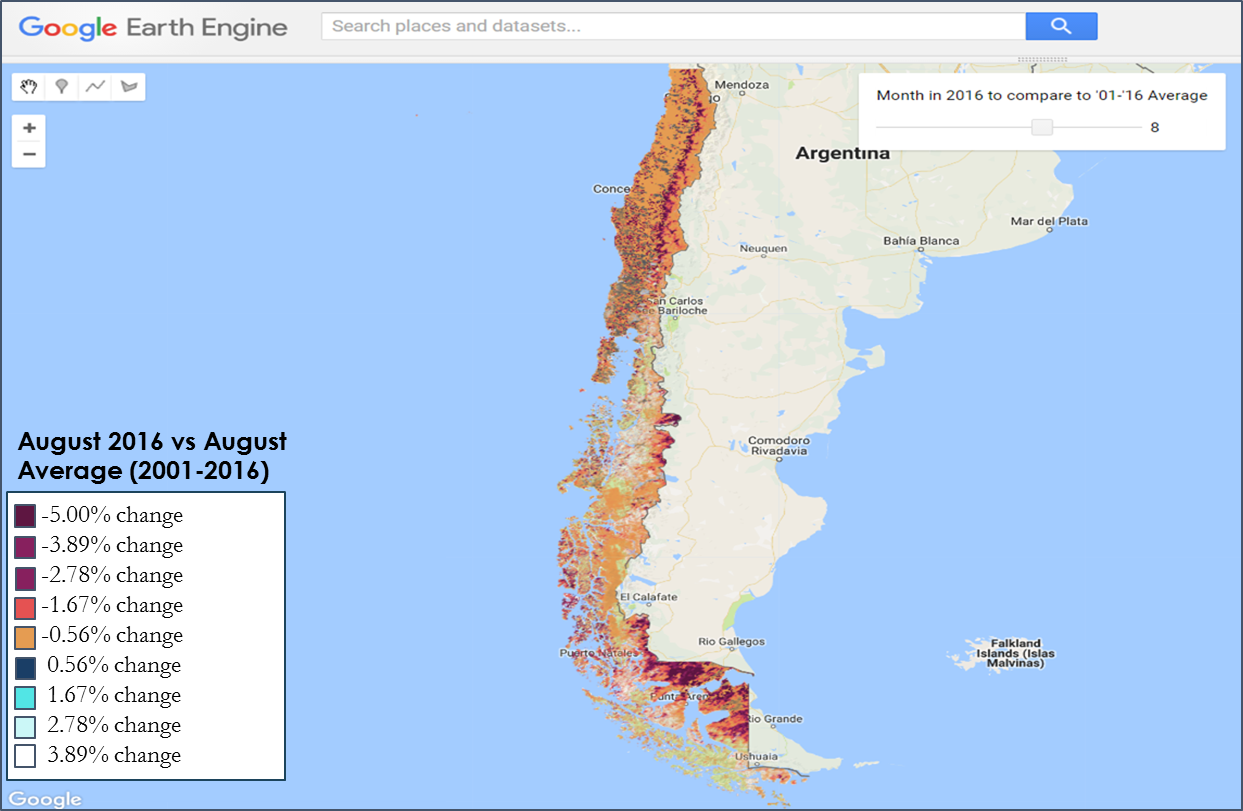
The GEE Python API can export Earth Engine image (EE Image) objects from GEE into GEE assets, but cannot upload GCS GeoTIFF files to GEE assets. Thus, the GEE CLI was used to upload GCS GeoTIFF files to the shared Chile-water asset repository. Specifically, the appropriate EE Image Collection was created in the Chile-water repository using the GEE interface and then individual GeoTIFFs were added to that collection in a loop with the GEE CLI tool. However, this process does not load metadata into the EE Image objects, which presents issues when performing analysis on the data. To address this issue, a Python script was written to fetch the date of individual GeoTIFF files and upload each image with static properties using the GEE CLI tool. While the GDAL Python library is the proper way to automatically fetch the date of a given GeoTIFF file, the team was forced to use this method of manipulating the file name to find the date due to ten-week project timeline.

Once in the GEE Chile-water asset folder, satellite images were analyzed and visualized using GEE. Some of the datasets required minor modification before proper analysis. The SMAP dataset includes AM and PM measurements every day which covered the entire study area after 2-3 days of measurement. Thus, GEE was used to mosaic these images into a SMAP 3-day collection for proper analysis. For the full SWE dataset, the AMSR-E dataset (2002-2011) was combined with the AMSR2 dataset (2012-2017) using GEE.

***3.3 Data Analysis***

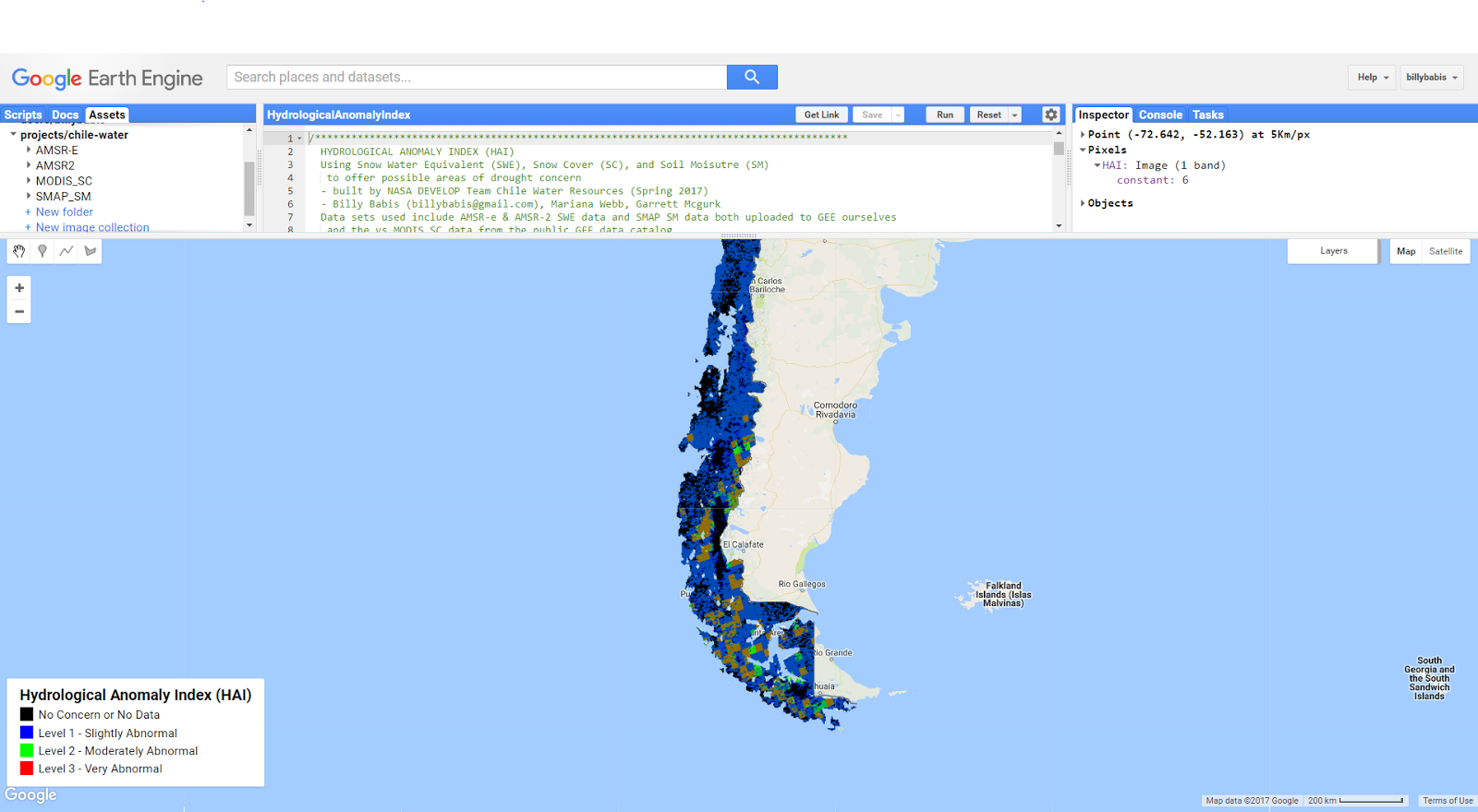
*3.3.1 Hydrological Anomaly Engine (HAE)*

The cumulative product built for this feasibility study was the Hydrological Anomaly Engine (HAE). This product includes the shared asset repository (Chile-water) discussed above, a shared script repository (ARC-ChileWater), and an in-depth data ingestion tutorial. The HAE was made accessible to the Chilean Ministry of Agriculture for hosting and accessing data, version-controlled sharing of analysis files, and exemplary guidance for uploading and visualizing Earth observations.

Although performing analysis and verification of the datasets was not an objective for this term, the HAE includes scripts to manipulate and visualize SC, SWE, and SM datasets to demonstrate the capabilities of GEE. The SMAP 3-day collection did not offer reliable anomaly assessments due to the short amount of time that the data had been available. However, GEE visualizes the SM data with a time-series chart to portray the short-term anomalies for each pixel. The MODIS SC visualization example compares the average SC of the ‘current’ month with the average of the same month for the years that data were available (2001- 2017). After generating a monthly average SC collection, the average value for the present month was compared to previous averages for that month, by taking the average of those monthly averages. Similarly, the coalesced SWE dataset generated annual averages of SWE with the ability to visualize different years with a slider. These scripts demonstrate some interactive User Interface tools GEE can build for visualizations including a slider bar, legends, color bars, and charts.

**Figure 2:**MODIS Snow Cover Visualization.

*3.3.2 Hydrological Anomaly Index (HAI)*

The HAI is an analysis tool that integrates SC, SWE, and SM datasets into a single index. The HAI compares ‘current’ SWE, SC, and SM values with averages of the past discussed above. A discrete distribution was established where a divergence from average of 40% of a given parameter yields a 3 HAI value (highly abnormal), a divergence from average of 25% of a given parameter yields a 2 HAI value (moderately abnormal), and over a 10% divergence from average of a given parameter yields a 1 HAI value (slightly abnormal). The HAI values for each parameter are combined into one HAI value where the SMAP HAI value is weighted less heavily due to the increased statistical inaccuracies of the shorter dataset. While the team performed no verification of the visualization, this tool demonstrates the ability to determine potential areas of hydrological concern.

**Figure 3:**Hydrological Anomaly Index (HAI).

# 4. Results & Discussion

***4.1 Analysis of Results***

*GEE Assessment*

As a feasibility study for the utilizing the GEE platform to monitor drought conditions in Chile, emphasis was placed on providing a qualitative assessment of the platform and communicating challenges that need to be overcome before this tool can be adopted by the Ministry of Agriculture. Ultimately, the HAE determined GEE to be very powerful for hosting, processing, and visualizing hydrological phenomena for the Chilean Ministry of Agriculture. However, there are various caveats that must be taken into consideration before using this tool.

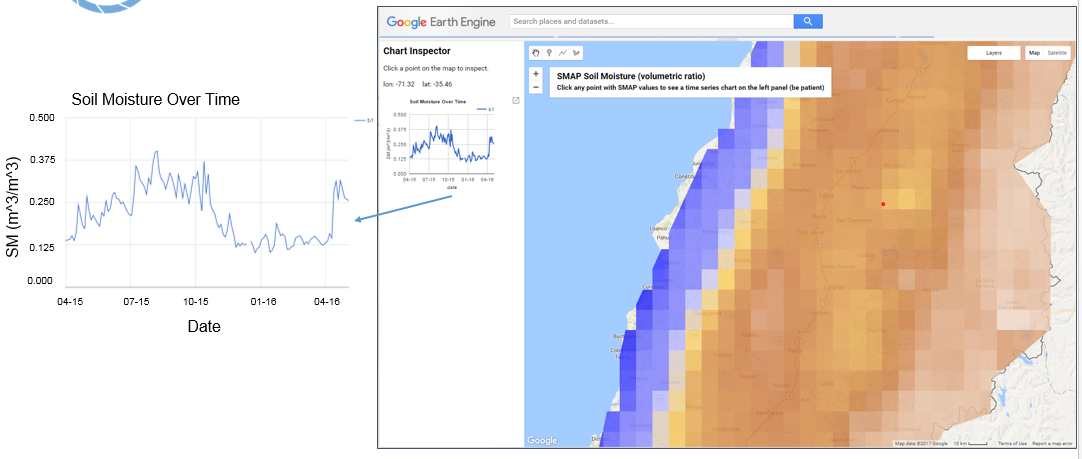
Finding, formatting, analyzing, and visualizing satellite data can often be clumsy and cumbersome. However, GEE solves many of these tedious issues, which maximizes the time a user can spend assessing the data. The GEE public data catalog is an informative and seamless way to access, import, and visualize vast satellite datasets and will only continue to grow. The GEE Code Editor interface (Figure 3) includes many helpful tools for effective analysis including an inspector tab to find pixel-by-pixel values of each band and very detailed documentation to simplify programming with GEE JavaScript objects. These GEE objects and methods are sensible and offer powerful server-side processing capacities. The computing power of GEE is perhaps its most powerful feature; the speed and ease of manipulating and visualizing large satellite datasets is immense. The HAI tool is able to process, mathematically alter, and visualize 15 years of daily SWE data, 16 years of daily SC data, and 2 years of bi-daily SM data in seconds.

*Data Ingestion Problems*

While GEE is an increasingly powerful assessment tool, the cloud-based service contains some problems that the GEE team will need to improve. The main issue with building the HAE was uploading the data for analysis. Once satellite data is on GEE either in a private repository or the public GEE data catalog, computation and visualization of these datasets are relatively seamless and intuitive for most users. In contrast, the process of uploading new data sets to GEE was much more difficult. The data pipeline described above is what the team determined to be the best option for uploading GeoTIFF datasets. However, it involves many steps that may exceed some user’s technical capabilities.  There were also several formatting challenges when seeking the proper file format to upload, with the proper projection, extent, and scale. Several manipulations with the HEG tool and the GDAL Python library made the files appropriate for ingestion to GEE.

The main problem with this data ingestion process at the time of this project was the difficulty of uploading datasets with appropriate metadata. While the GEE CLI tool supports assignment of a date attribute and any list of properties, it did not support customizable band names, image or EE Image Collection descriptions, provider info, and units for bands. Once an asset was loaded into GEE, some of these metadata properties could be edited and loaded into a new GEE asset. While this was an option for enhancing the data sets, adding another layer to the data pipeline seemed overly meticulous, especially considering the quickly improving GEE features. The GEEBAM library is a great tool for uploading local GeoTIFF files in batch and as the GEE CLI tool continues to develop, this library will be a powerful tool for uploading datasets with metadata.

However, the GEE team at Google was an extremely valuable resource in solving many of these issues. Questions asked on the GEE Developers Google Group received timely responses from GEE team members and experienced GEE users. As the GEE platform grows the GEE team at Google continues to update the product with meaningful improvements.



**Figure 4:**SMAP Visualization.

*Sample Visualizations*

The HAE includes several visualizations and analytic tools demonstrating the capabilities of the GEE platform. While these tools have not been validated, they serve as a demonstration of what future verifiable models can do. The images shown below illuminate some analytical and visual features of GEE. The SWE equivalent visualization (Figure 5) demonstrates a useful visualization of data in which a comparison of SWE values for 2016 is compared to 2002. The interactive slide bar at the top of the screen and color bar above the images provide an example of features that can be incorporated into the GEE platform using programmable scripts. The SMAP SM visualization (Figure 4) demonstrates the interactive chart feature of the GEE tool.  Users are able to click a point on the map triggering the generation of a SM time-series chart for that individual pixel. These tools are powerful, easy to implement, and effective in visualizing hydrologic data.

**Figure 5:**Snow Water Equivalent Visualization

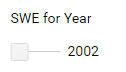
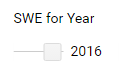
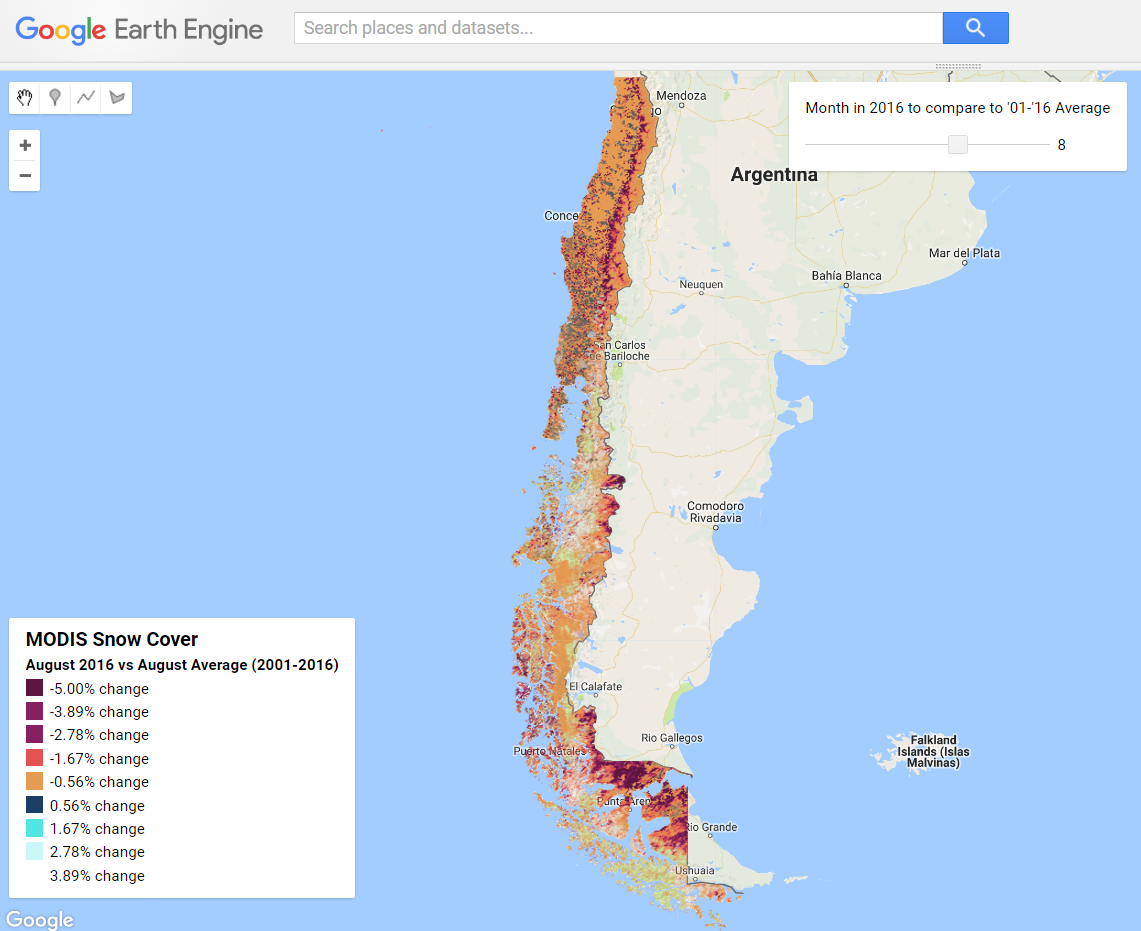


0mm

40mm

2016

2002



***4.2 Future Work***

As a feasibility study, the main objective of this project was to demonstrate the capabilities of GEE to the Chilean Ministry of Agriculture. To accomplish this, the primary focus was centered around creating a web based data visualization and analysis platform built within Google Earth Engine. Future terms of this project could focus on adding functionality to the tool and developing a more polished product that end-users can implement into their decision-making process.

# While the HAE was useful in demonstrating GEE’s capabilities, it was far from a completed project. Future drought-monitoring tools built within GEE will need to undergo an in-depth data verification process that assesses the accuracy of the visualizations. Validation for this tool should compare in situ measurements with remotely sensed data on the GEE platform for accurate analysis. Future work may draw from data assimilation strategies developed by the Global Land Data Assimilation Systems (GLDAS) tool developed at the NASA Goddard Center. Ultimately, we hope this case study highlights the many benefits of utilizing GEE as a drought-monitoring platform and has shown the Ministry of Agriculture what a more finalized tool could be capable of as they look to modify these algorithms for future use.

# Future work on the HAE should focus on minor improvements that improve user experience and functionality. The HAE currently does not support live ingestion and analysis of Near Real-Time (NRT) satellite data. Implementing this feature should be a priority in during future work with this product. Additionally, the HAI could greatly benefit from these improvements and ultimately be hosted on Google App Engine as a user-friendly, publically available tool for assessing hydrologic data in Chile. Additionally, as the GEE data ingestion process continues to improve, the datasets hosted on the HAE will grow more descriptive and user-friendly.

# 5. Conclusions

The Hydrological Anomaly Engine (HAE) was created to serve as a key element of the feasibility study demonstrating the capability of the Google Earth Engine (GEE) Platform to perform hydrologic analysis. Using remotely sensed soil moisture, snow water equivalent, and snow cover data, GEE was used to host, process, and visualize statistical anomalies in different regions throughout Chile. While the GEE public data catalog hosts many popular, easily accessible satellite datasets, uploading supplementary datasets was a difficult process. The HAE hosts these datasets and associated visualization scripts in a shared repository that is easily shareable and importable. The processing power of GEE’s server-side computations make the free cloud-based service stand out among competing geospatial analysis tools.

Chile’s diverse biomes create a complex system of water resource management. As stakeholders hydrologic needs intensify, the Ministry of Agriculture may look to implement a more technologically advanced solution for drought-monitoring. The GEE platform has great potential for serving as a useful addition to the Ministry of Agriculture’s drought-monitoring toolset. When interactive tools are properly developed and validated, the HAE can be used to serve as a valuable resources in the Ministry of Agriculture’s decision-making process.

# 6. Acknowledgments

* Dr. Juan Torres-Perez, Bay Area Environmental Research Institute
* Dr. Eduardo Bendek, NASA Ames Research Center
* Dr. Kenton Ross, NASA Langley Research Center
* Brittany Zajic, NASA Ames Research Center
* Jenna Williams, NASA Ames Research Center

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Aeronautics and Space Administration.

This material is based upon work supported by NASA through contract NNL16AA05C and cooperative agreement NNX14AB60A.

# 7. Glossary

**AMSR –** Advanced Microwave Scanning Radiometer

**CDL –** Climate Data Library

**CLI** - Command-Line Interface

**DAAC –** Distributed Active Archive Center

**DGA –** General Directorate of Water Resources

**GCS** - Google Cloud Storage

**GDAL** - Geospatial Data Abstraction Library

**GEE –** Google Earth Engine

**GHRC** - Global Hydrological Resource Center

**GLDAS** - Global Land Data Assimilation Systems

**HAE** - Hydrological Anomaly Engine

**HAI** - Hydrological Anomaly Index

**JAXA –** Japanese Aerospace Exploration Agency

**MODIS** **–** Moderate resolution Imaging Spectroradiometer

**NRT** - Near Real-Time

**NSIDC –** National Snow and Ice Data Center Data

**SMAP –** Soil Moisture Active Passive

**SC –** Snow Cover

**SM –** Soil Moisture

**SWE –** Snow water equivalent

# 8. References

Bennett, M., New, M., Marino, J., & Sillero-Zubiri, C. (2016). Climate complexity in the Central Andes: A study case on empirically-based local variations in the Dry Puna. *Journal of Arid Environments, 128*, 40-49.

Dong, J., Xiao, X., Menarguez, M. A., Zhang, G., Qin, Y., Thau, D., . . . Moore, B. (2016). Mapping paddy rice planting area in northeastern Asia with Landsat 8 images, phenology-based algorithm and Google Earth Engine. *Remote Sensing of Environment, 185*, 142-154. doi:10.1016/j.rse.2016.02.016

Gobierno de Chile. (n.d.). Retrieved February 16, 2017, from http://www.minagri.gob.cl/

Hall, D. K. and G. A. Riggs. 2016. MODIS/Aqua Snow Cover Daily L3 Global 500m Grid, Version 6. [Indicate subset used]. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center, accessed February 13, 2016. doi:10.5067/MODIS/MYD10A1.006

Johansen, K., Phinn, S., & Taylor, M. (2015). Mapping woody vegetation clearing in Queensland, Australia from landsat imagery using the Google Earth Engine. *Remote Sensing Applications: Society and Environment, 1*, 36-49.

Meza, F. J. (2013). Recent trends and ENSO influence on droughts in Northern Chile: An application of the Standardized Precipitation Evapotranspiration Index. *Weather and Climate Extremes, 1*, 51-58. doi:10.1016/j.wace.2013.07.002

O'Neill, P. E., S. Chan, E. G. Njoku, T. Jackson, and R. Bindlish. 2016. SMAP Enhanced L3 Radiometer Global Daily 9 km EASE-Grid Soil Moisture, Version 1. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center, accessed 13 February 2017. doi:10.5067/ZRO7EXJ8O3XI

Tedesco, M., R. Kelly, J. L. Foster, and A. T. Chang. 2004. AMSR-E/Aqua Daily L3 Global Snow Water Equivalent EASE-Grids, Version 2. [Indicate subset used]. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center, accessed 17 February 2017. doi:10.5067/AMSR-E/AE\_DYSNO.002

Valdés-Pineda, R., Pizarro, R., García-Chevesich, P., Valdés, J. B., Olivares, C., Vera, M., ... & Abarza, A. (2014). Water governance in Chile: Availability, management and climate change. *Journal of Hydrology, 519*, 2538-2567.

Vicuña, S., Mcphee, J., & Garreaud, R. D. (2012). Agriculture Vulnerability to Climate Change in a Snowmelt-Driven Basin in Semiarid Chile. *Journal of Water Resources Planning and Management, 138*(5), 431-441. doi:10.1061/(asce)wr.1943-5452.0000202