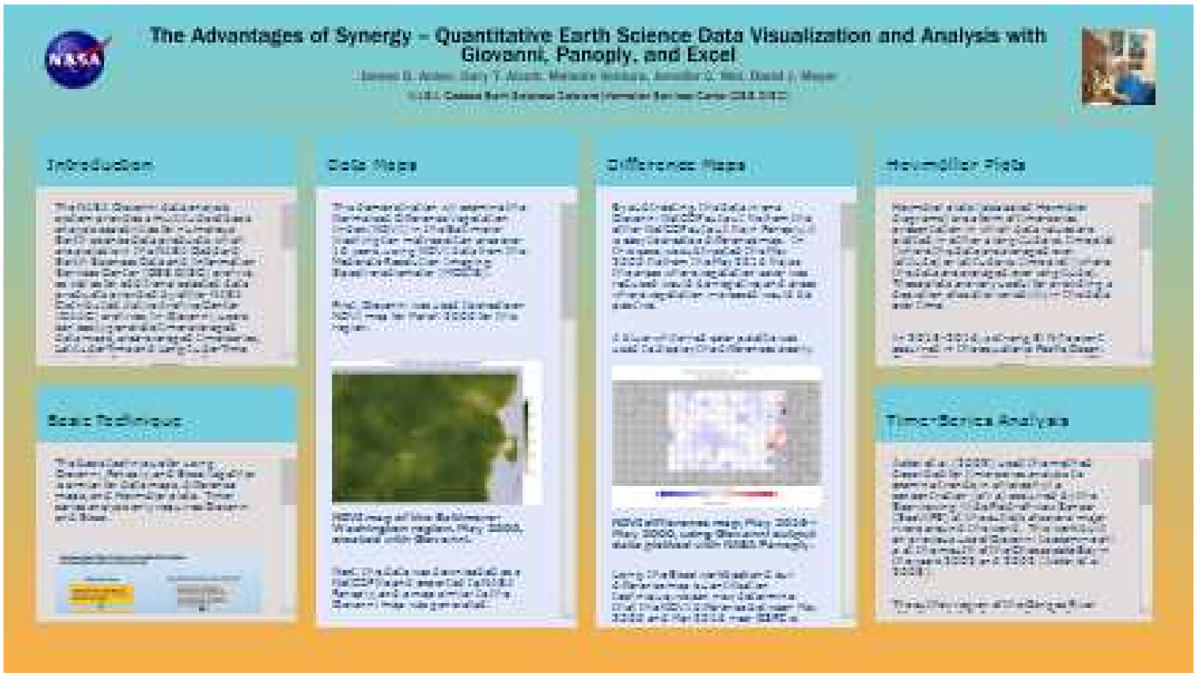
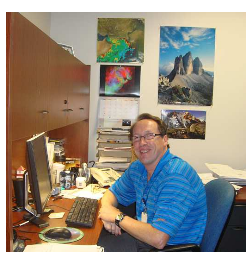
The Advantages of Synergy – Quantitative Earth Science Data Visualization and Analysis with Giovanni, Panoply, and Excel



James G. Acker, Gary T. Alcott, Melanie Ventura, Jennifer C. Wei, David J. Meyer

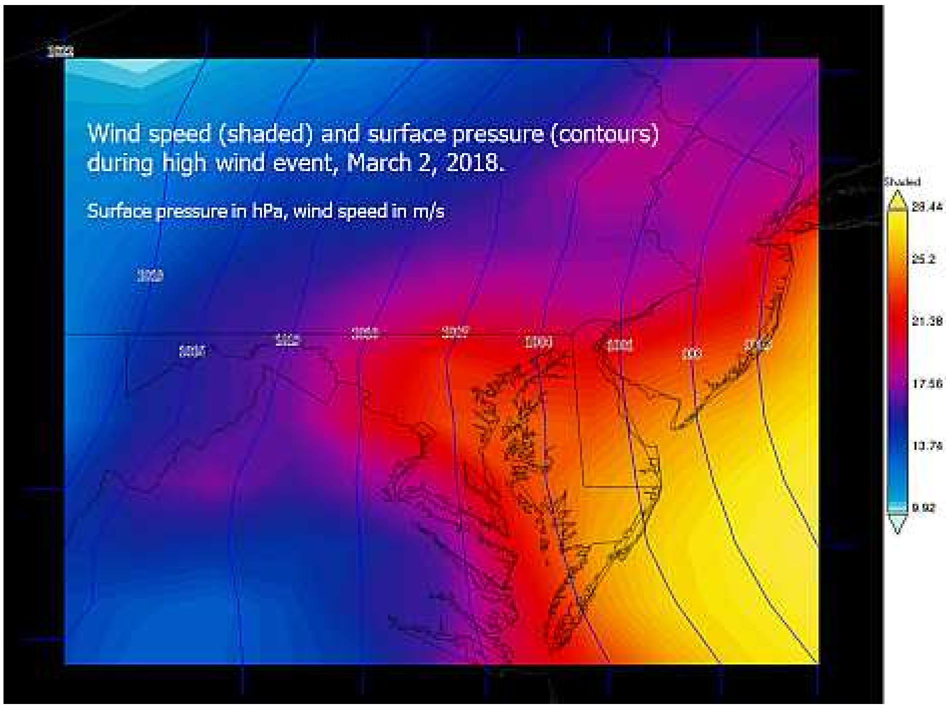
NASA Goddard Earth Sciences Data and Information Services Center (GES DISC)



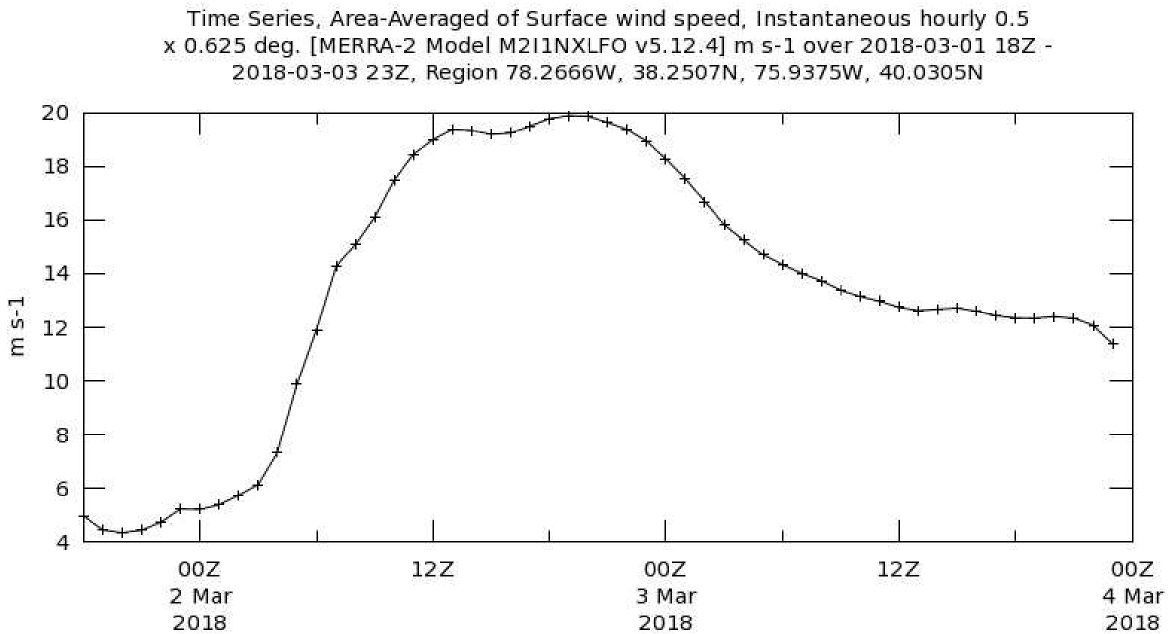
PRESENTED AT:

INTRODUCTION

The NASA Giovanni data analysis system provides a multitude of basic analysis capabilities for numerous Earth science data products which are available in the NASA Goddard Earth Sciences Data and Information Services Center (GES DISC) archive, as well as for additional selected data products provided by other NASA Distributed Active Archive Center (DAAC) archives. In Giovanni, users can easily generate time-averaged data maps, area-averaged time-series, Latitude-Time and Longitude-Time Hovmöller diagrams, correlation maps, accumulation maps, and map animations (22 analysis options are available in total).



Wind speed and surface pressure for the high wind event ("bomb cyclone") of March 2, 2018. Giovanni's overlay map analysis was used for this plot.



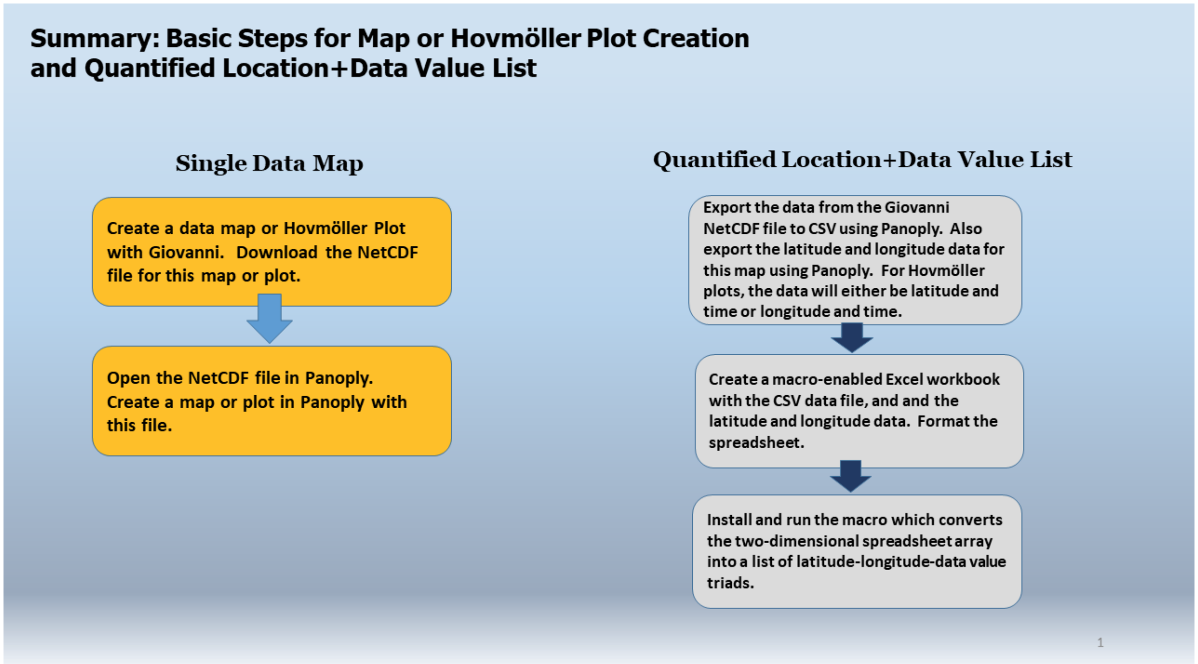
Instantaneous surface wind speed time-series for the high wind event of March 2, 2018.

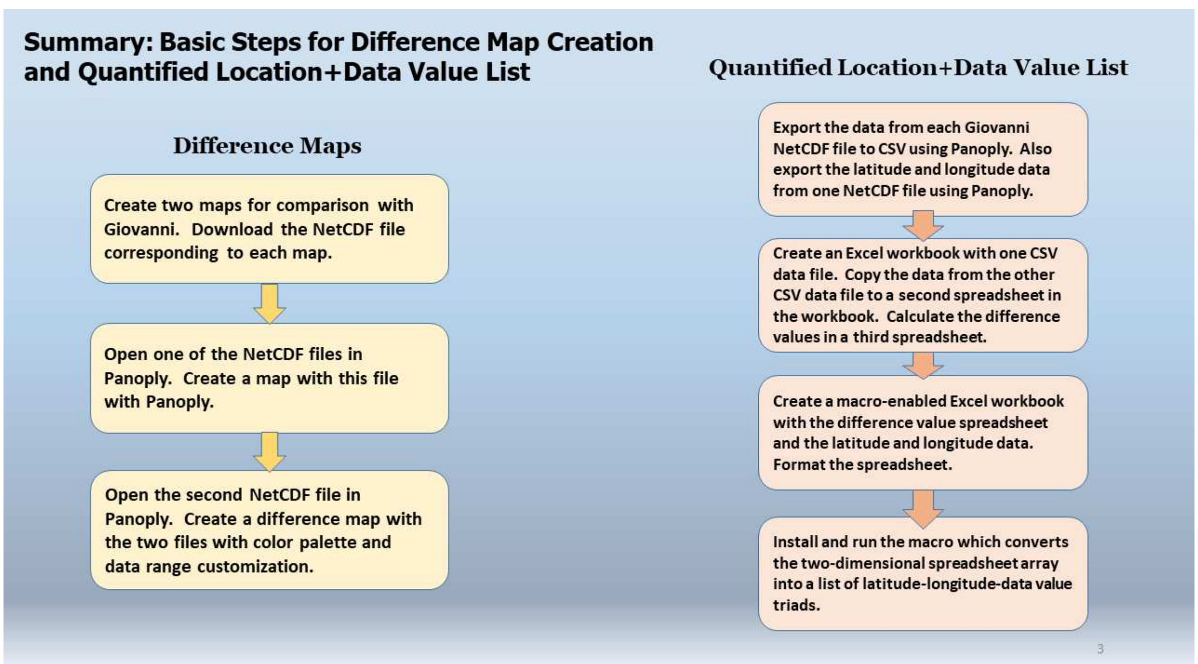
Giovanni also provides the capability to download the data and data files corresponding to a visualization. These data files can be imported into other software packages (such as NASA Panoply or Microsoft Excel) for further analysis and detailed result quantification.

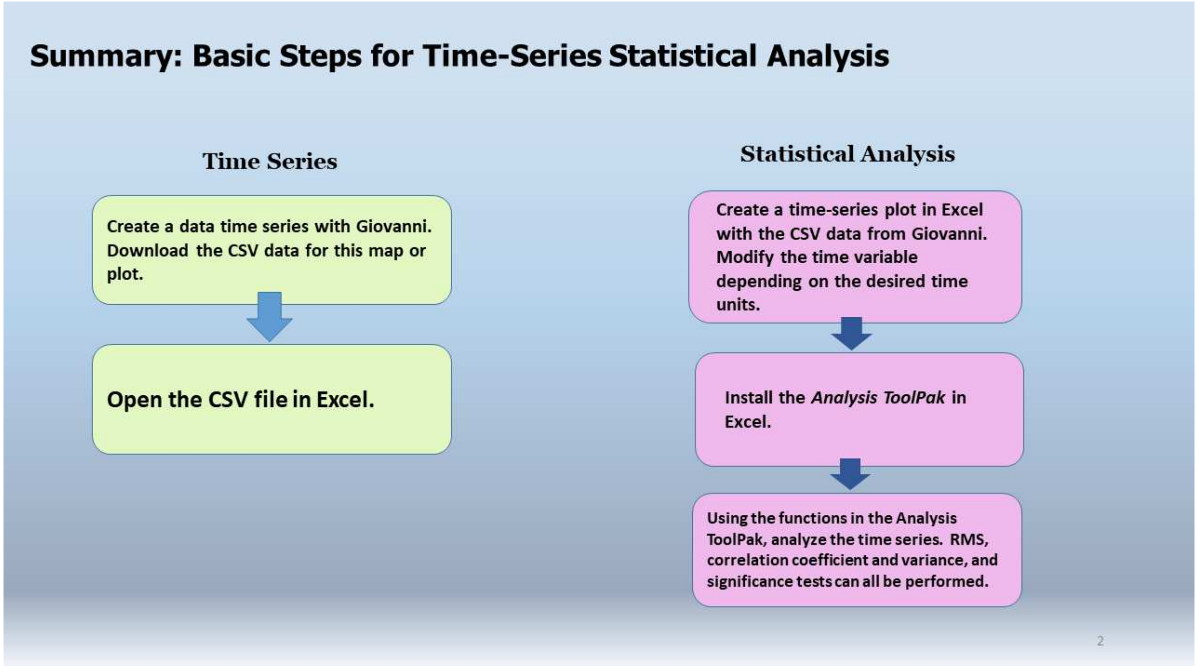
This presentation will demonstrate download and analysis procedures using Giovanni, Panoply, and Excel, showing how the combination of these tools provides synergistic research benefits using the wide variety of NASA data sets available in Giovanni.

BASIC TECHNIQUE

The basic technique for using Giovanni, Panoply, and Excel together is similar for data maps, difference maps, and Hovmöller plots. Time-series analysis only requires Giovanni and Excel.



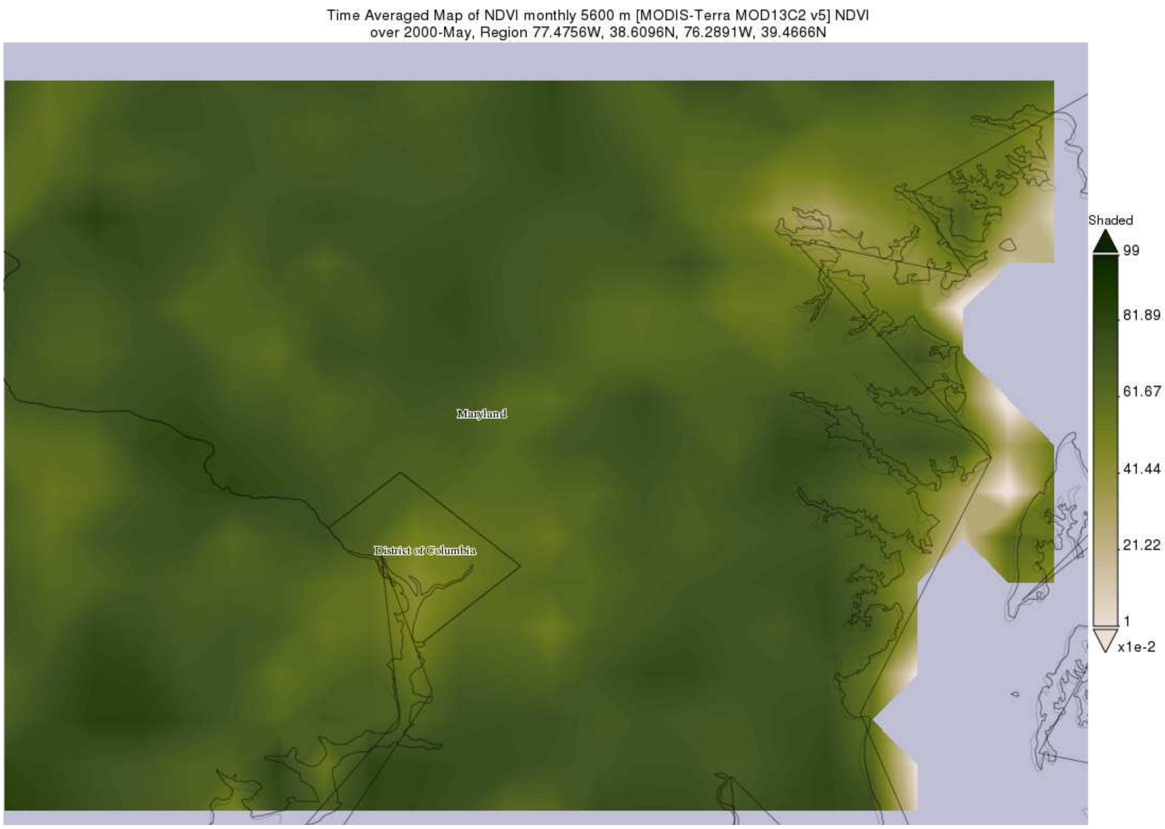



Summary: Basic Steps for Time-Series Statistical Analysis
Time Series
Create a data time series with Giovanni. Download the CSV data for this map or plot.
Open the CSV file in Excel.
Statistical Analysis
Create a time-series plot in Excel with the CSV data from Giovanni. Modify the time variable depending on the desired time units.
Install the Analysis ToolPak in Excel.
Using the functions in the Analysis ToolPak, analyze the time series. RMS, correlation coefficient and variance, and significance tests can all be performed.
2

DATA MAPS

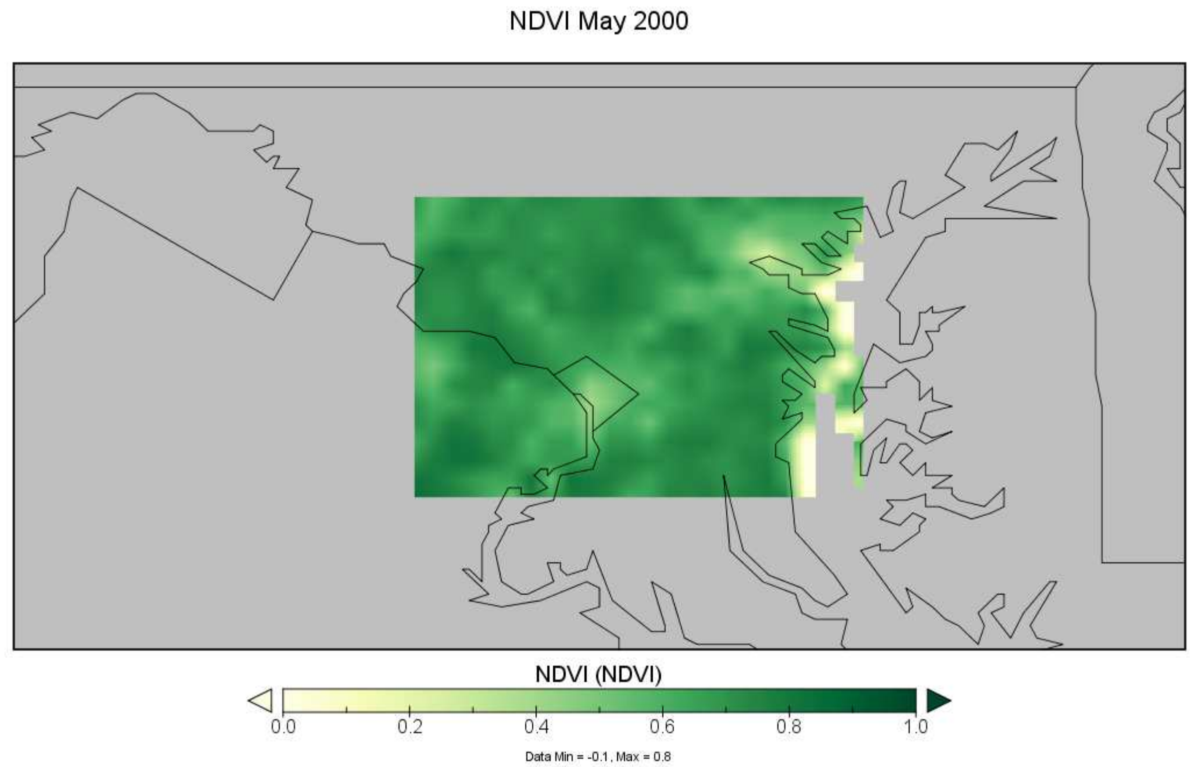
This demonstration will examine the Normalized Difference Vegetation Index (NDVI) in the Baltimore-Washington metropolitan area over 16 years, using NDVI data from the Moderate Resolution Imaging Spectroradiometer (MODIS).

First, Giovanni was used to create an NDVI map for March 2000 for this region.



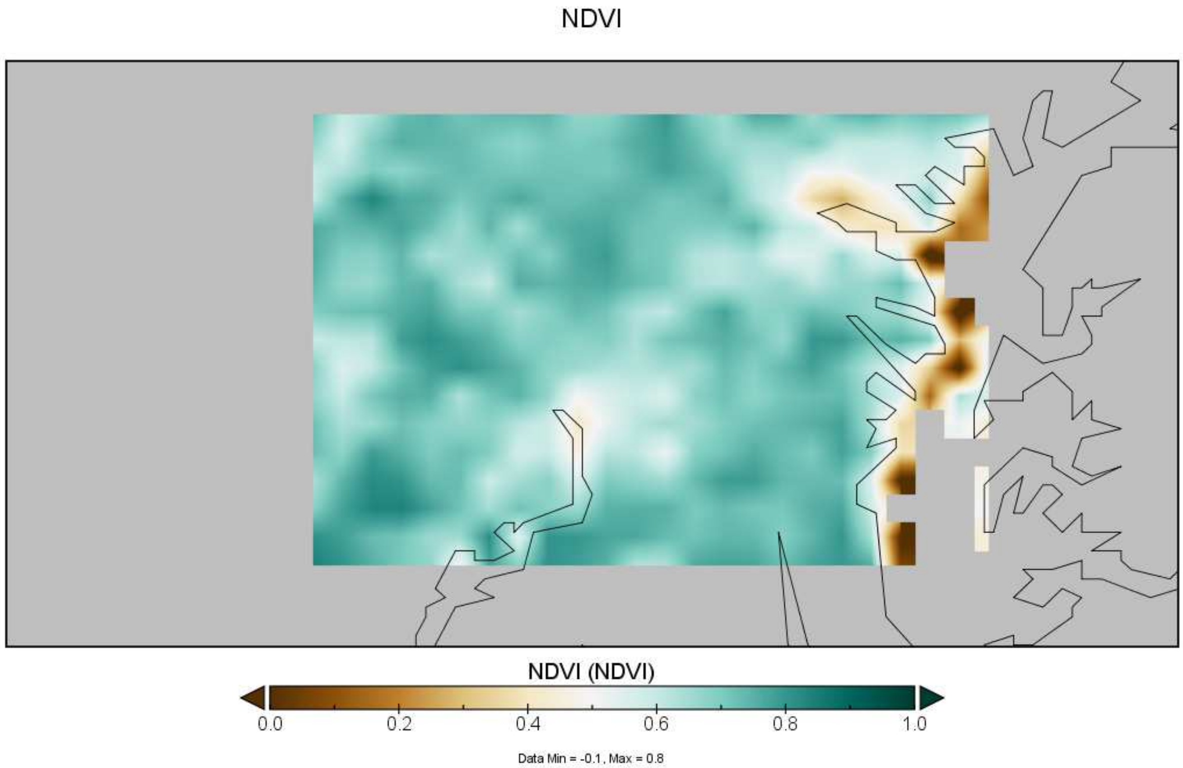
NDVI map of the Baltimore-Washington region, May 2000, created with Giovanni.

Next, the data was downloaded as a NetCDF file and exported to NASA Panoply, and a map similar to the Giovanni map was generated.



NDVI map of the Baltimore-Washington region, May 2000, created with Giovanni output data plotted in Panoply.

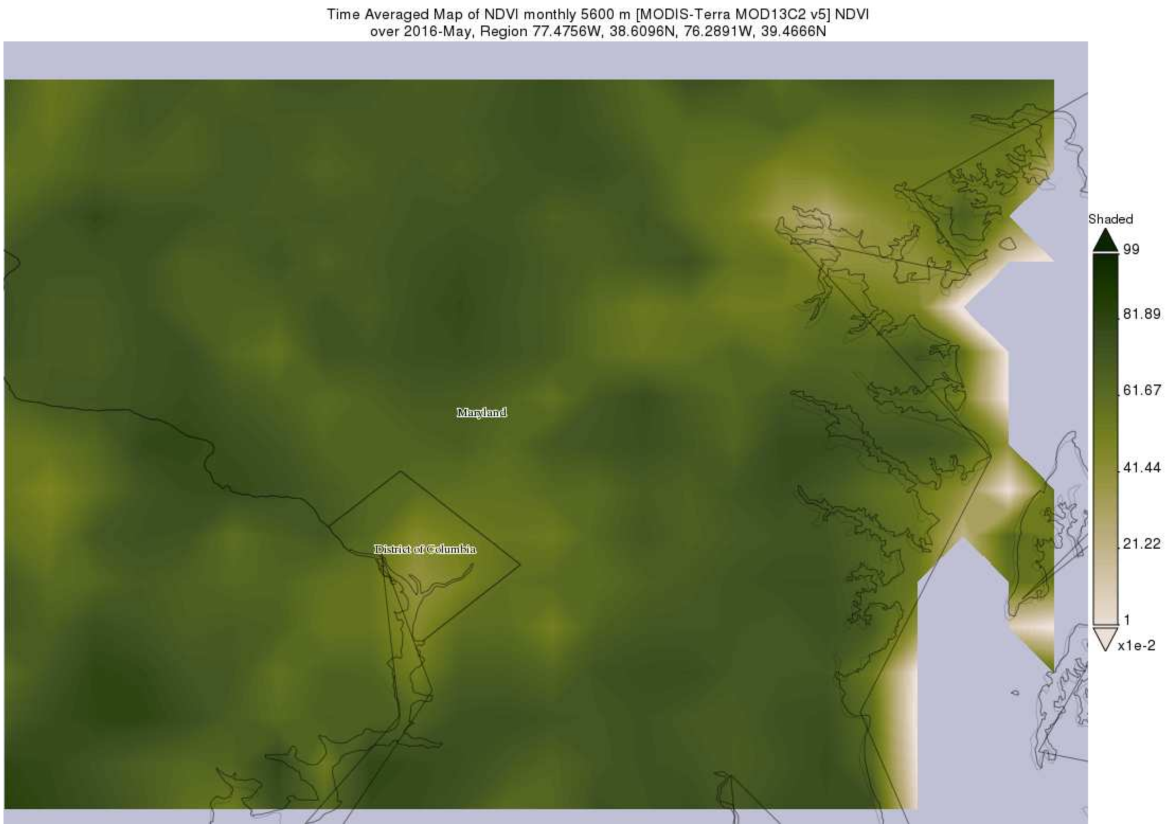
Both Giovanni and Panoply offer a variety of color palettes for data display. This Brown-Blue Green palette is better at showing contrast between vegetated and less-vegetated areas. The Chesapeake Bay waters do not have terrestrial vegetation, which is why they appear dark brown.



NDVI map of the Baltimore-Washington region, May 2000, created with Giovanni output data plotted in Panoply, using the Brown-Blue Green color palette.

Using the Excel spreadsheet, we can determine that the NDVI value for the pixel containing Goddard Space Flight Center is 0.7238 (GSFC is in a reasonably well-wooded area).

Now look at this map of NDVI for May 2016, plotted with Giovanni.



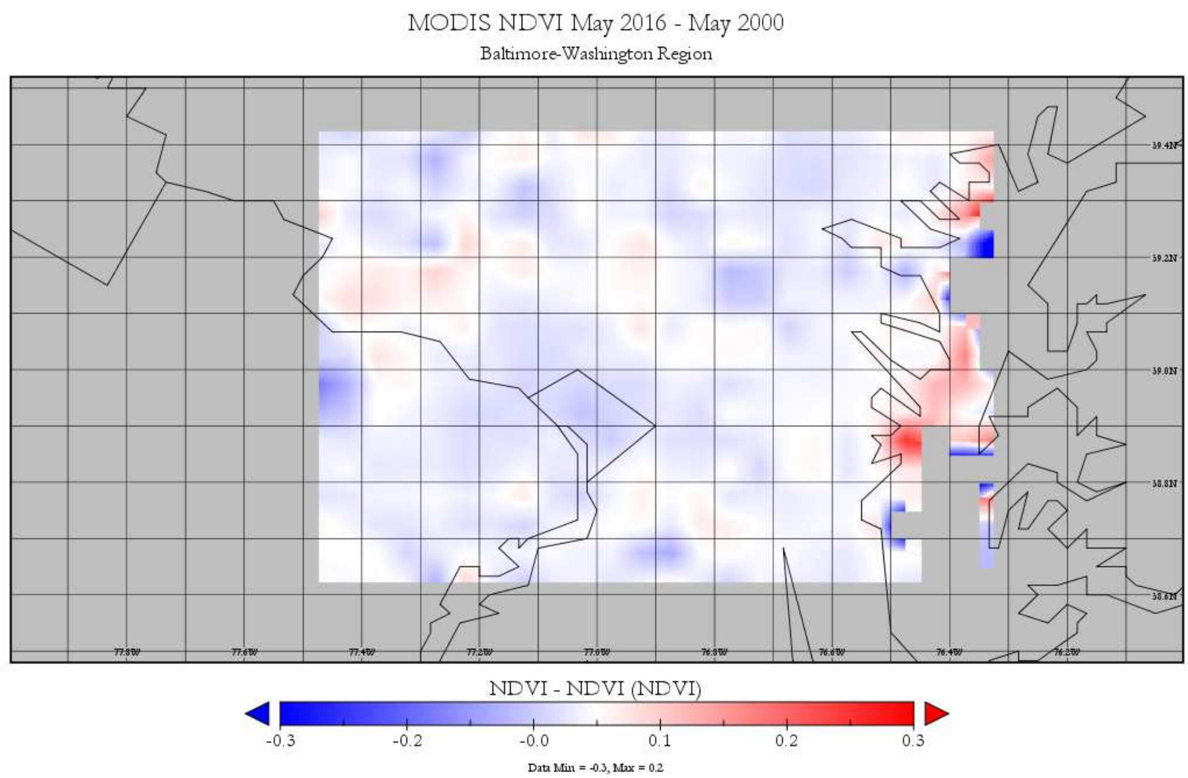
NDVI map of the Baltimore-Washington region, May 2016, created with Giovanni.

It is not easy to see any differences between the May 2000 map and the May 2016 map. So now we can create a difference map for these two time periods.

DIFFERENCE MAPS

By subtracting the data in one Giovanni NetCDF output file from the other NetCDF output file in Panoply, it is easy to create a difference map. In this case, we subtracted the May 2000 file from the May 2016 file, so the areas where vegetation cover was reduced would be negative, and areas where vegetation increased would be positive.

A blue-white-red color palette was used to display the differences clearly.



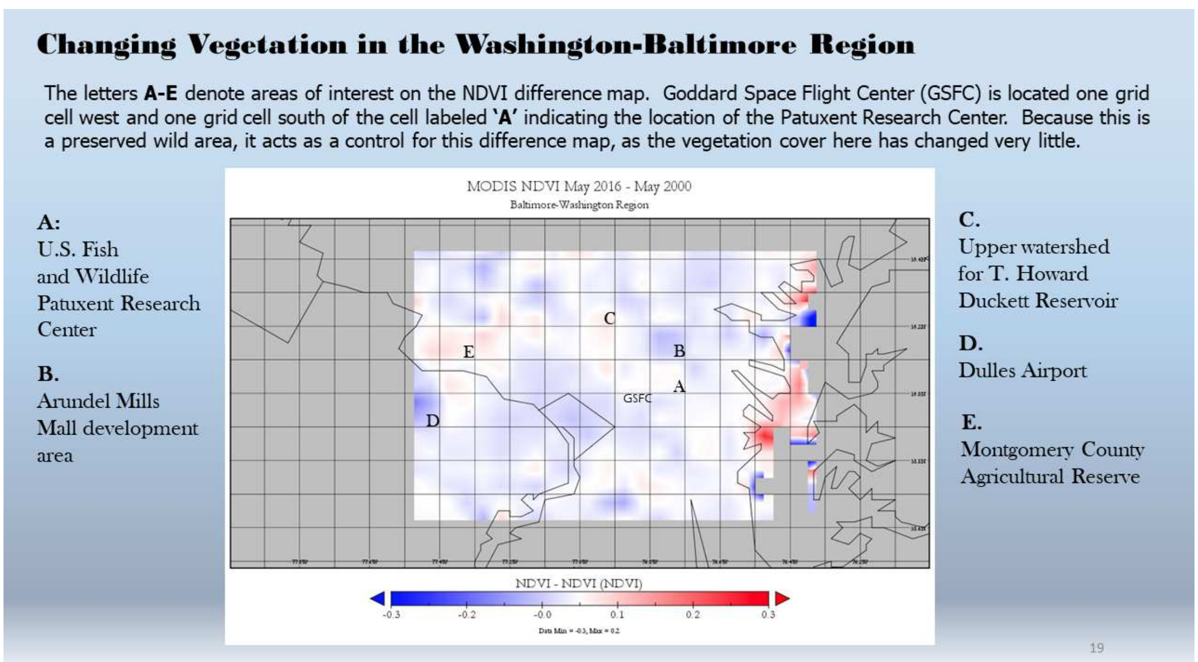
NDVI difference map, May 2016 - May 2000, using Giovanni output data plotted with NASA Panoply.

Using the Excel workbook and our difference map quantification technique, we can now determine that the NDVI difference between May 2000 and May 2016 near GSFC is 0.0048 - which is a very small change, so we've kept most of our trees.

Latitude	Longitude	NDVI difference

227	38.875	-76.825	-2.85E-02
228	38.925	-76.825	-3.27E-02
229	38.975	-76.825	-0.0177
230	39.025	-76.825	3.00E-04
231	39.075	-76.825	4.80E-03
232	39.125	-76.825	-0.0173
233	39.175	-76.825	-0.0144
234	39.225	-76.825	-0.0238

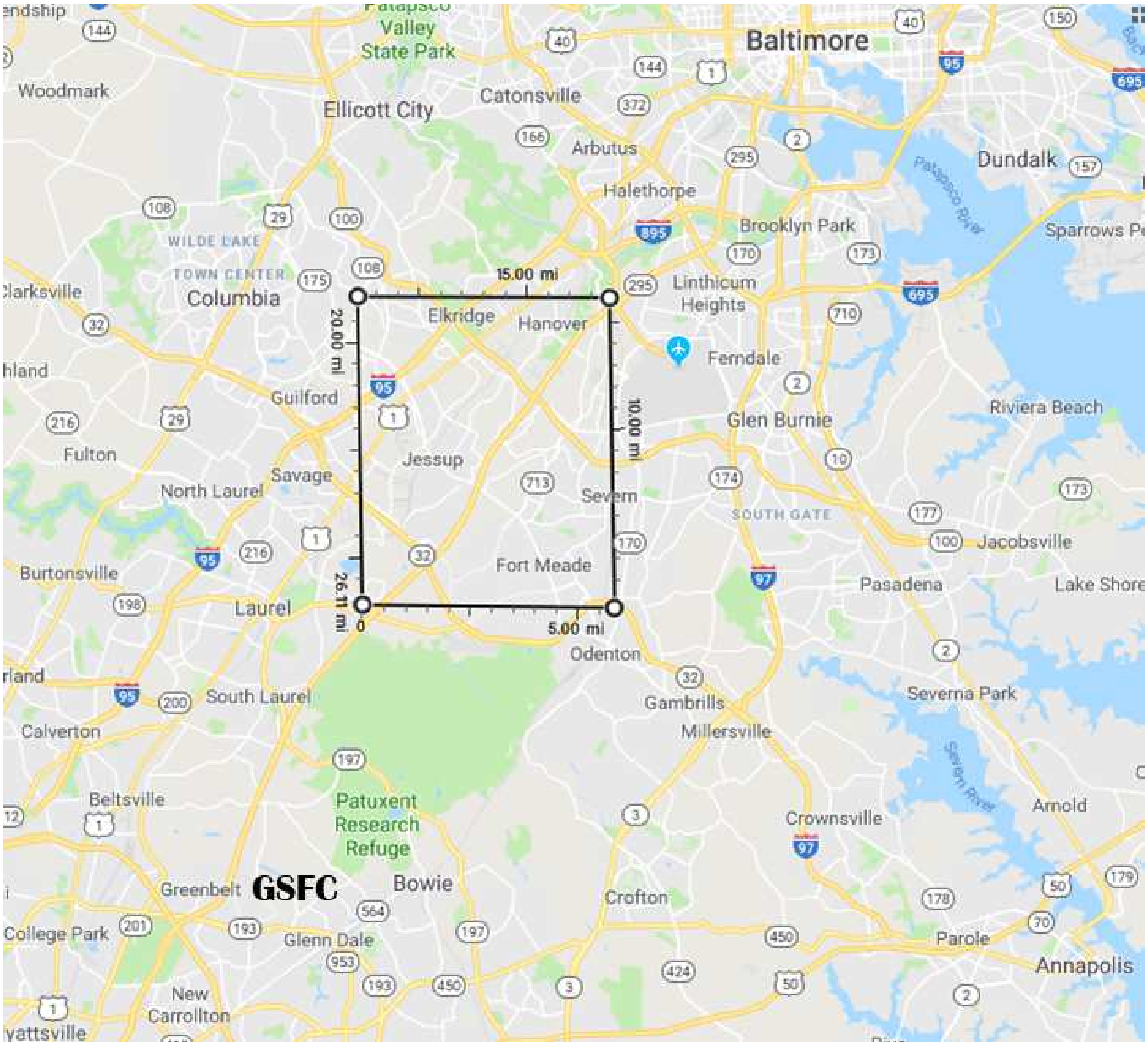
That isn't the case everywhere. The figure below shows some features in the region. Vegetation cover during May 2016, compared to May 2000, was significantly reduced near the Arundel Mills Mall development south of Baltimore, and around Dulles Airport in Virginia. However, vegetation cover actually increased (or was growing better) in the Montgomery County Agricultural Preserve area.

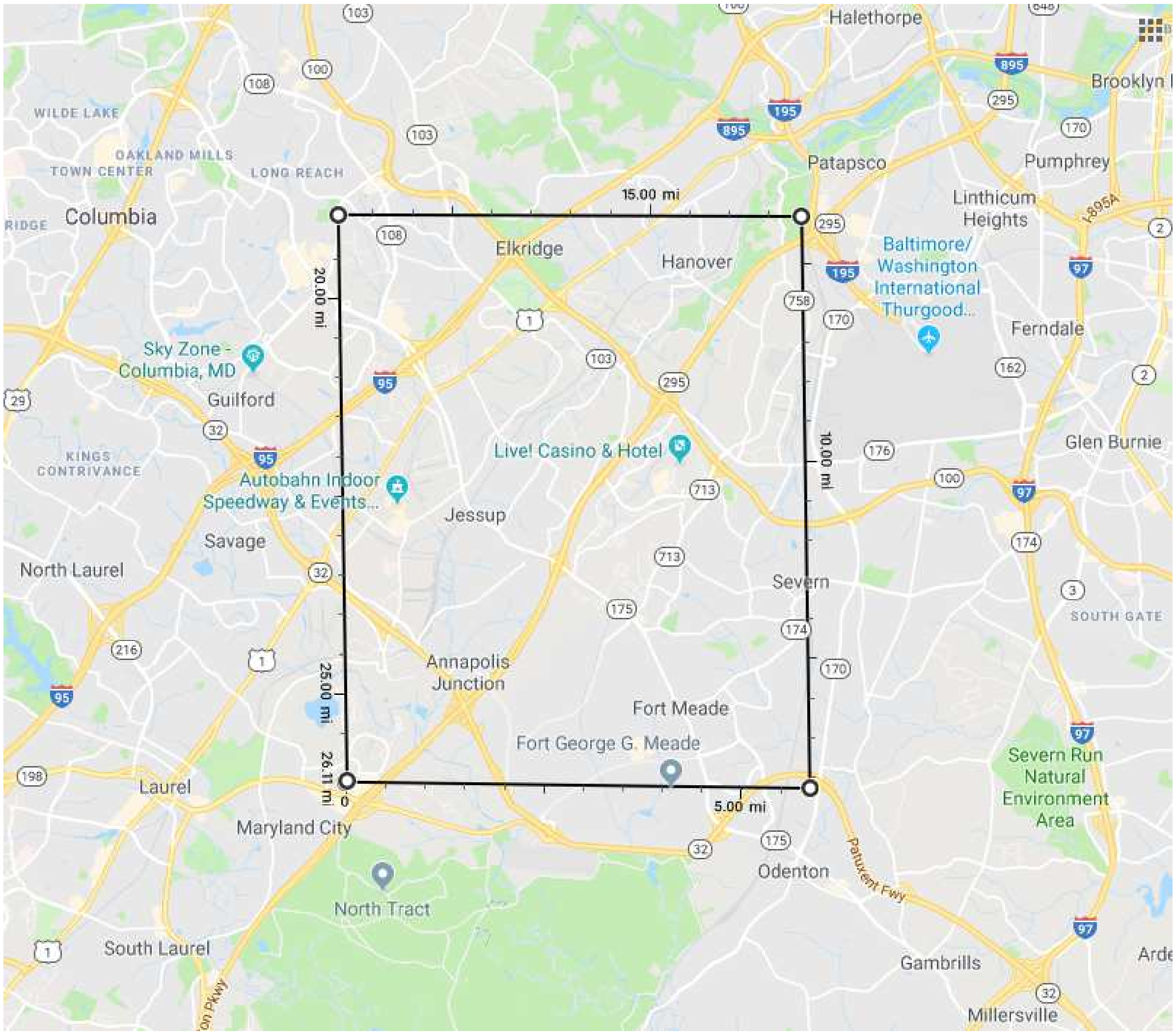


NDVI difference map, showing locations of interest in the Baltimore-Washington region.

Giovanni and Excel can also be used to investigate the possible influence of rainfall variability on the NDVI differences shown here. The following plots examine the Arundel Mills Mall grid cell ('B' in the plot above). This large outlet mall opened in November 2000, and both commercial and housing development around it occurred subsequently.

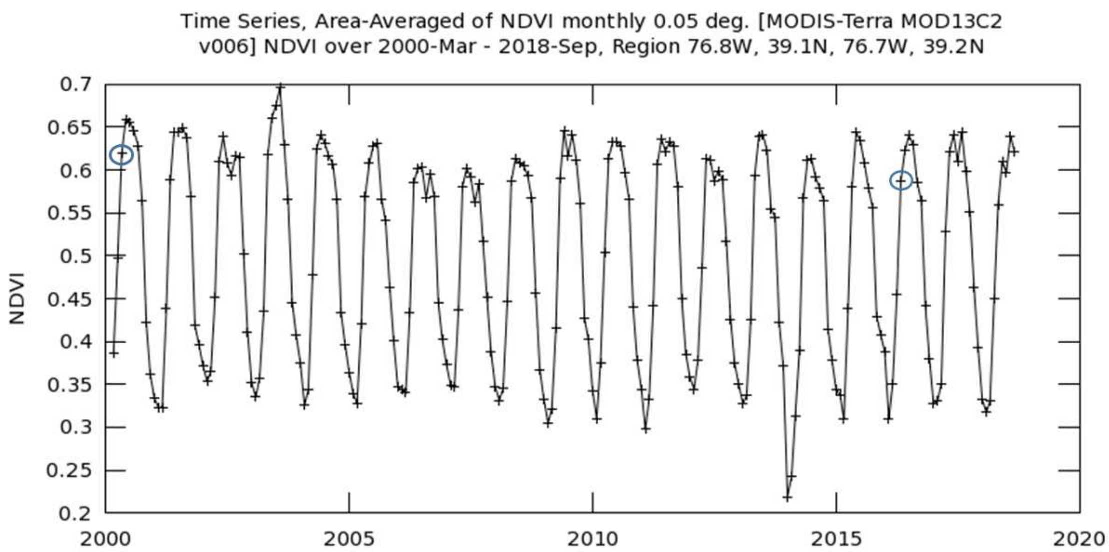
This is where the grid cell is located (shown on Google Maps).





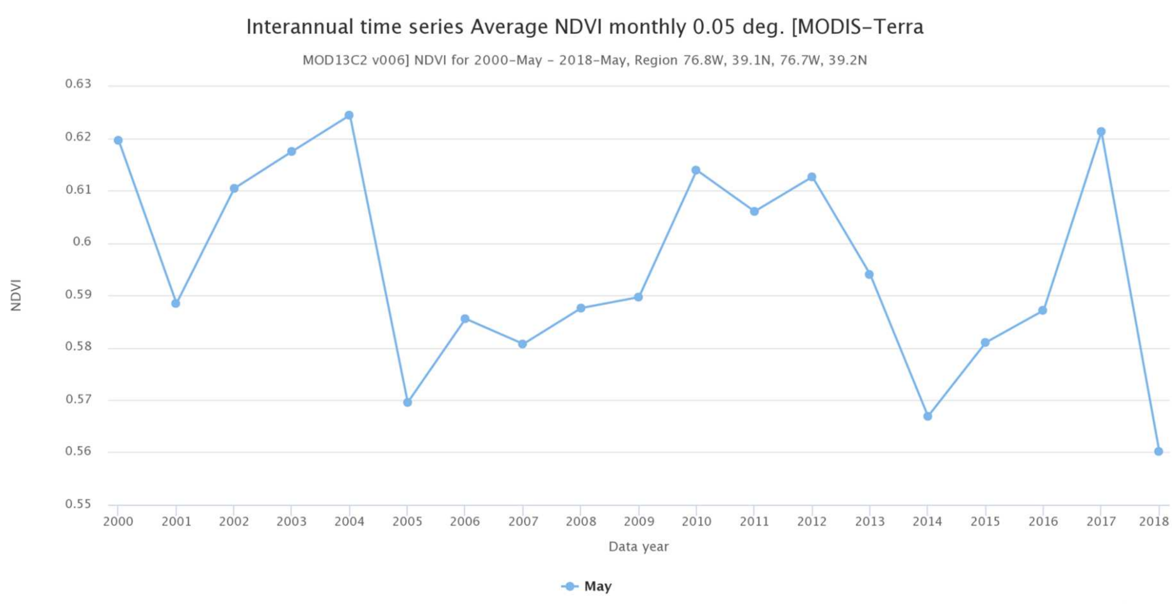
Location map for the Arundel Mills Mall grid cell. The *Live!* Casino is located adjacent to the shopping mall.

First, a time-series of NDVI for the Arundel Mills Mall grid cell was created. The latitude range is 39.1 to 39.2 degrees N, and the longitude range is 76.7 to 76.8 degrees W. In this time-series, the points for May 2000 and May 2016, the months used for the difference map, are circled.



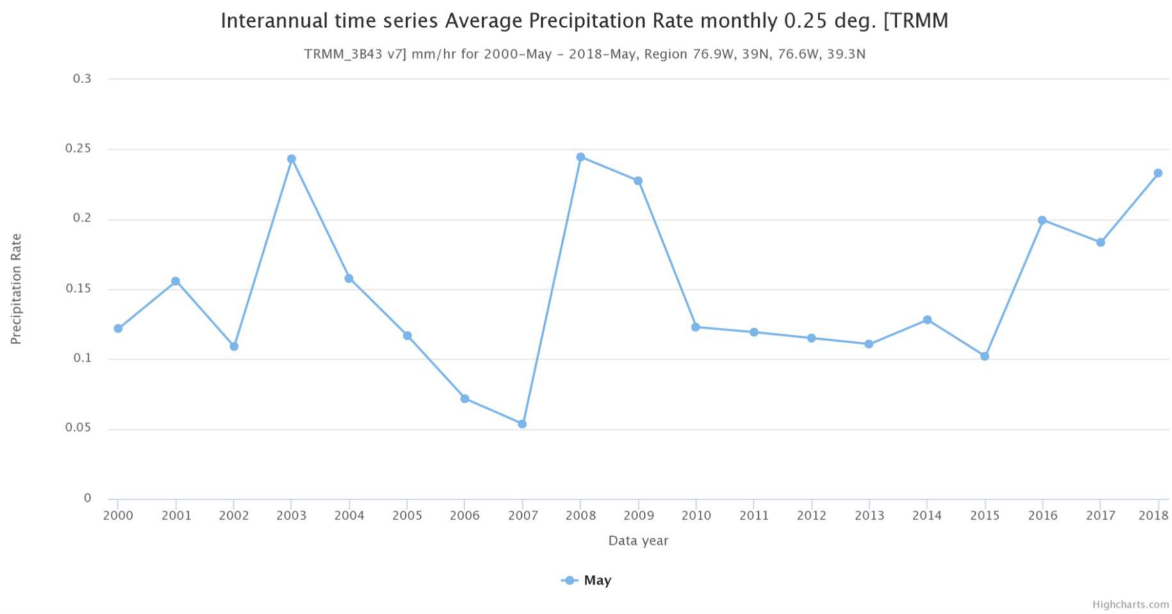
NDVI time-series, March 2000 - September 2018, for the Arundel Mills Mall grid cell. May 2000 and May 2016 are circled.

This time-series is difficult to interpret, other than clearly indicating the seasonal pattern. Giovanni can be used to create time-series for specific months or seasons, so we can make a time-series for only the month of May.



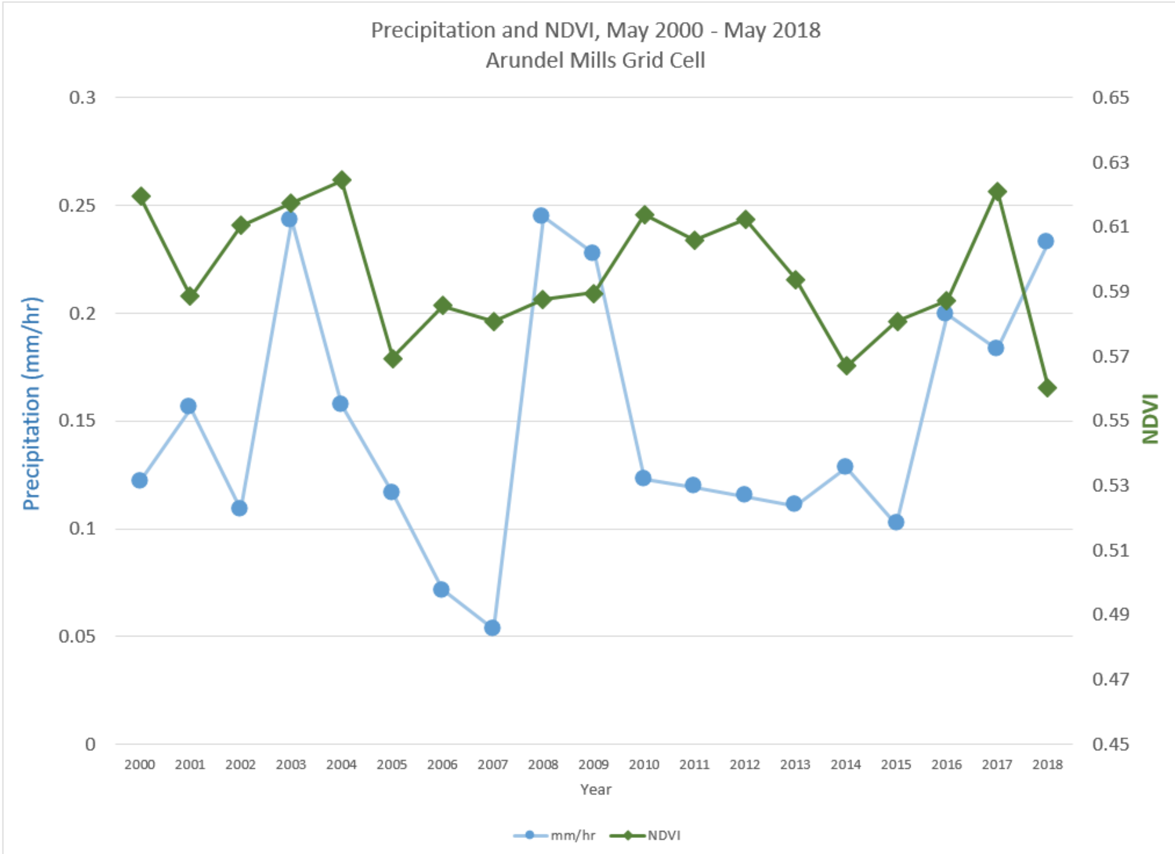
NDVI time series, May 2000-May 2018.

To examine possible relationships to the amount of rainfall in May, the TRMM 3B43_v7 data product can be used. This data product originated with the Tropical Rainfall Measuring Mission, but continues to present by including other precipitation data sources, such as "high-quality microwave data, infrared data, and analyses of rain gauges". Due to grid cell size, the area used was enlarged congruently around the Arundel Mills Mall grid cell to create this plot of regional rainfall in May.



TRMM 3B43_v7 Precipitation Rate time series, May 2000 - May 2018.

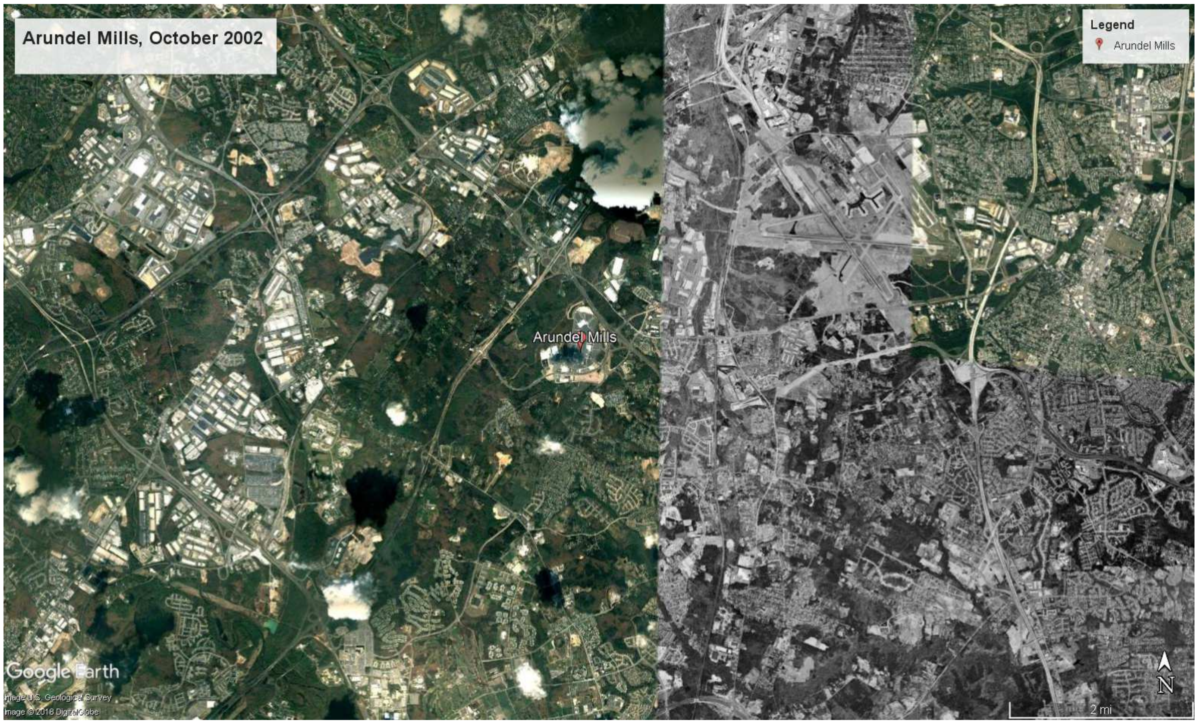
The data for both of these plots can be downloaded directly in CSV format. The downloaded data was imported into an Excel spreadsheet and used to create a two-axis, two-variable plot.

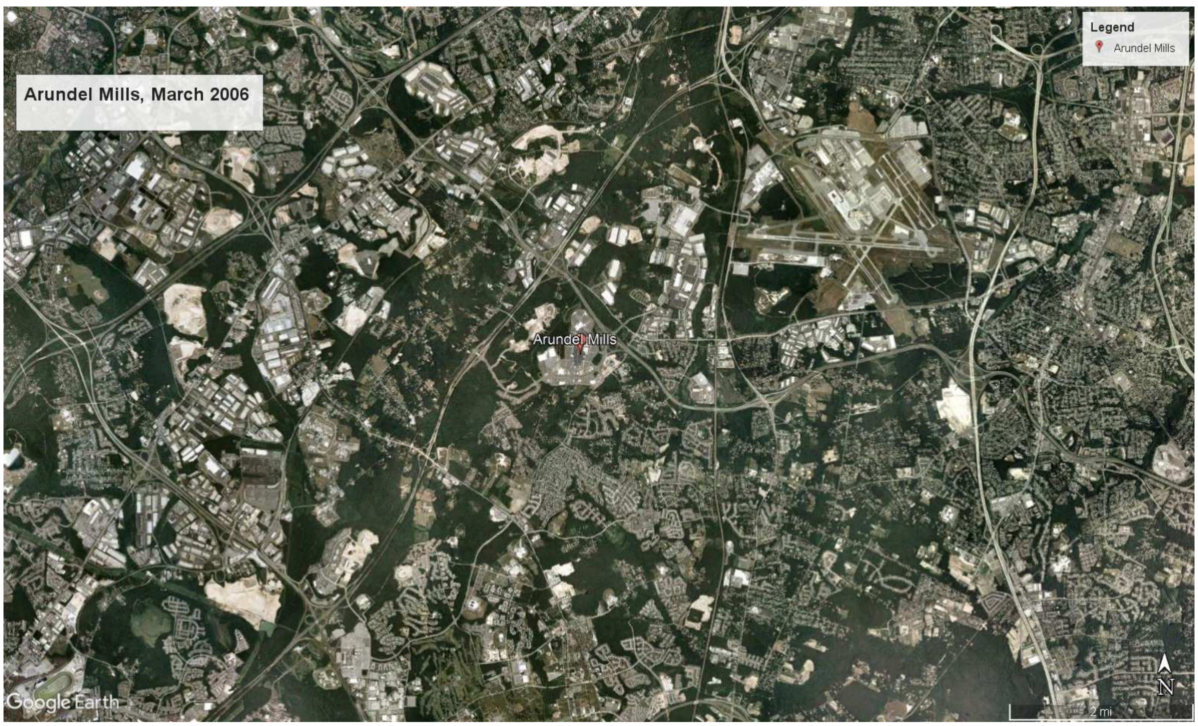


Excel time-series plot of precipitation rate and NDVI, May 2000 - May 2018.

This time-series plot does not indicate a significant correlation of May rainfall with the May NDVI value for the Arundel Mills Mall grid cell. In fact, there was more rain in May 2016 than in May 2000. However, because the May 2017 NDVI value is nearly the same as the May 2000 value, other factors may influence the value.

To visually validate this analysis, historical satellite images were found for the area with Google Earth. Three images of the region for October 2002, March 2006, and the most recent image available are shown below.


Arundel Mills, October 2002
Legend
Arundel Mills
Arundel Mills
Google Earth
Image U.S. Geological Survey
Image © 2018 DigitalGlobe
2 mi


Arundel Mills, March 2006
Legend
Arundel Mills
Arundel Mills
Google Earth
2 mi

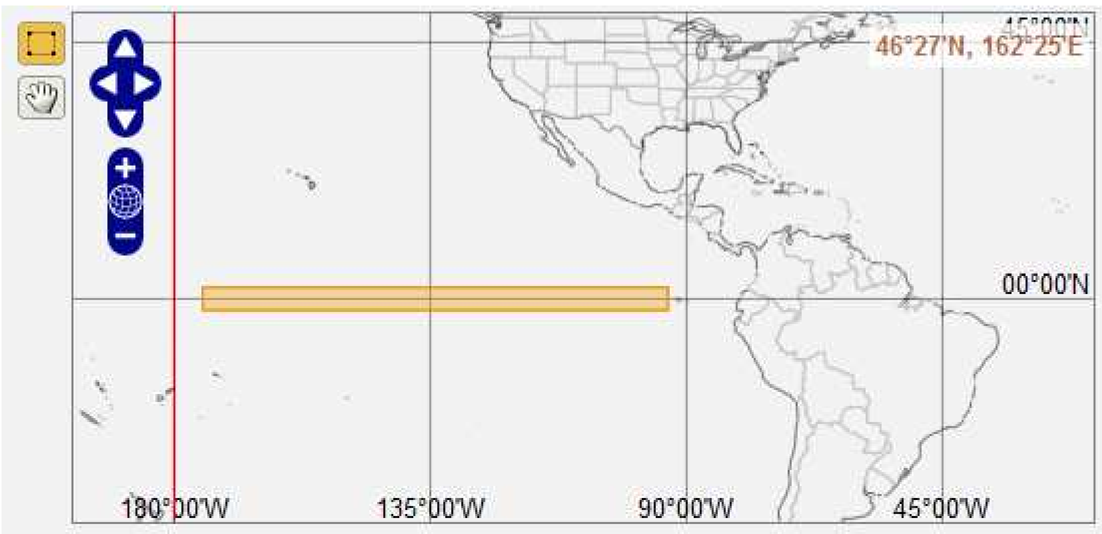
Another example of the use of NDVI similar to this analysis can be found in Witte et al. (2011). In this paper, an increase in NDVI was observed in Beijing, China, prior to the 2008 Olympic Games. This was believed to be due to the planting of thousands of trees to beautify the city for the Games. The paper also showed that NDVI declined after the Games, likely due to the lack of capability to care for the newly-planted trees.

HOVMÖLLER PLOTS

Hovmöller plots (also called Hovmöller diagrams) are a form of time-series presentation in which data values are plotted in either a longitude vs. time plot (where the data are averaged over latitude) or latitude vs. time plot (where the data are averaged over longitude). These plots are very useful for providing a depiction of spatial variability in the data over time.

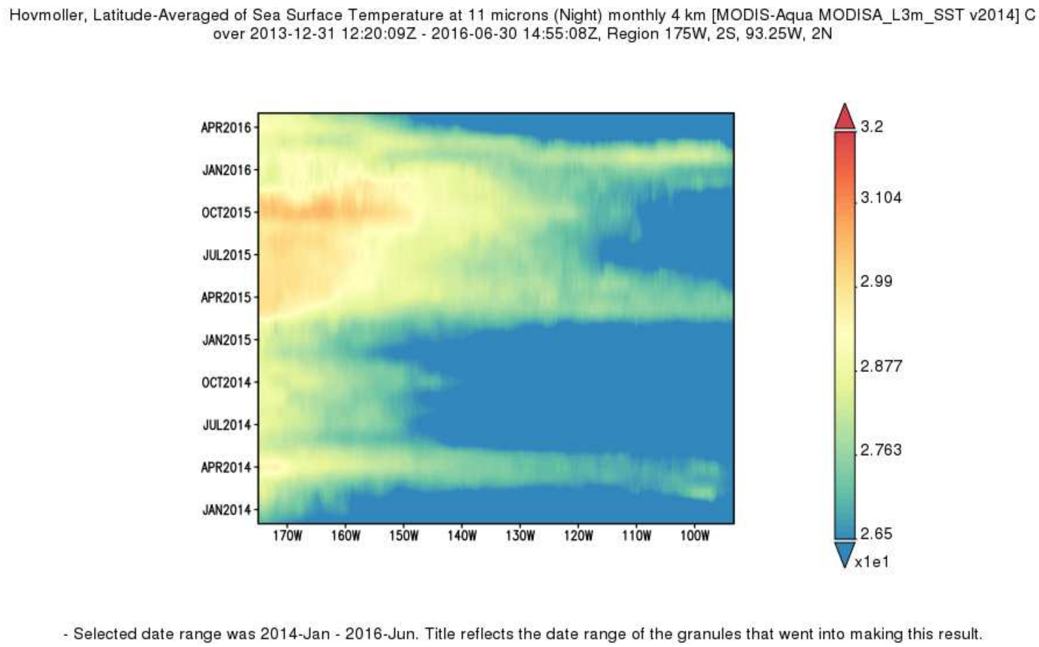
In 2015-2016, a strong El Niño event occurred in the equatorial Pacific Ocean. One of the classic oceanic hallmarks of El Niño events is increasing sea surface temperature (SST) in equatorial waters, generally progressing from west to east. Because the events evolve longitudinally, they provide excellent examples of the use of Hovmöller plots.

For this example, the region of interest shown below in orange was used.



Region of interest used for the examination of SST variability during the 2015-2016 El Niño event.

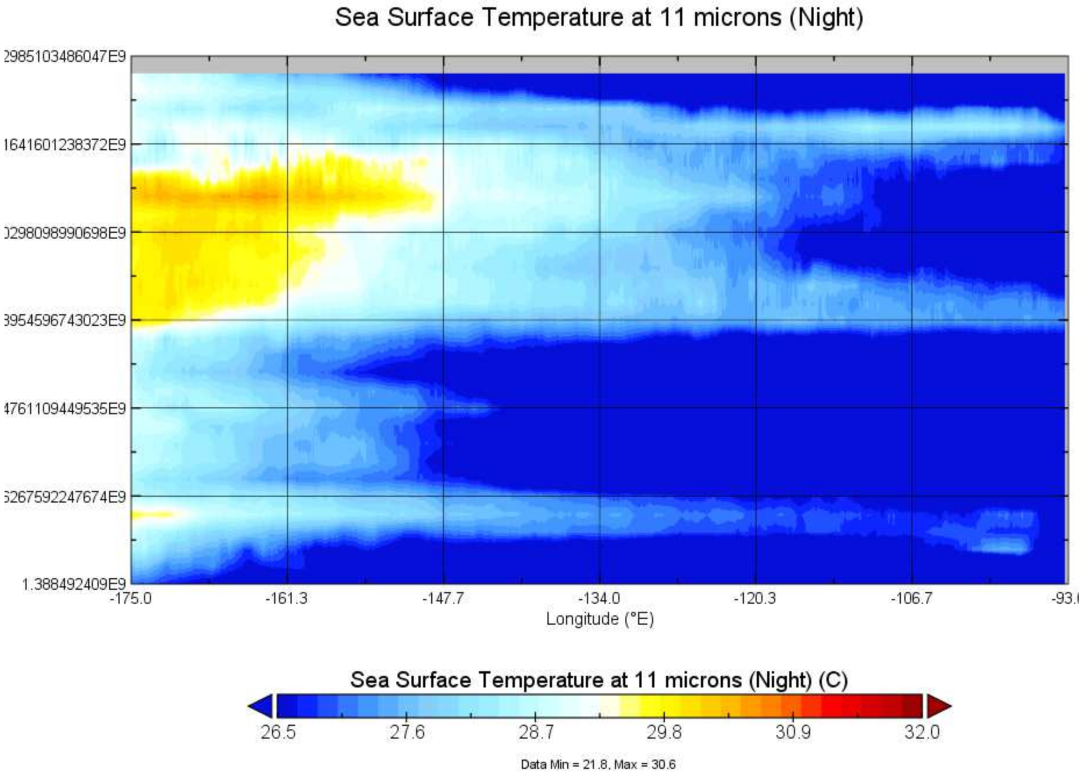
SST data from MODIS-Aqua was used to create the Hovmöller plot. The Hovmöller plot generated with Giovanni is shown below.

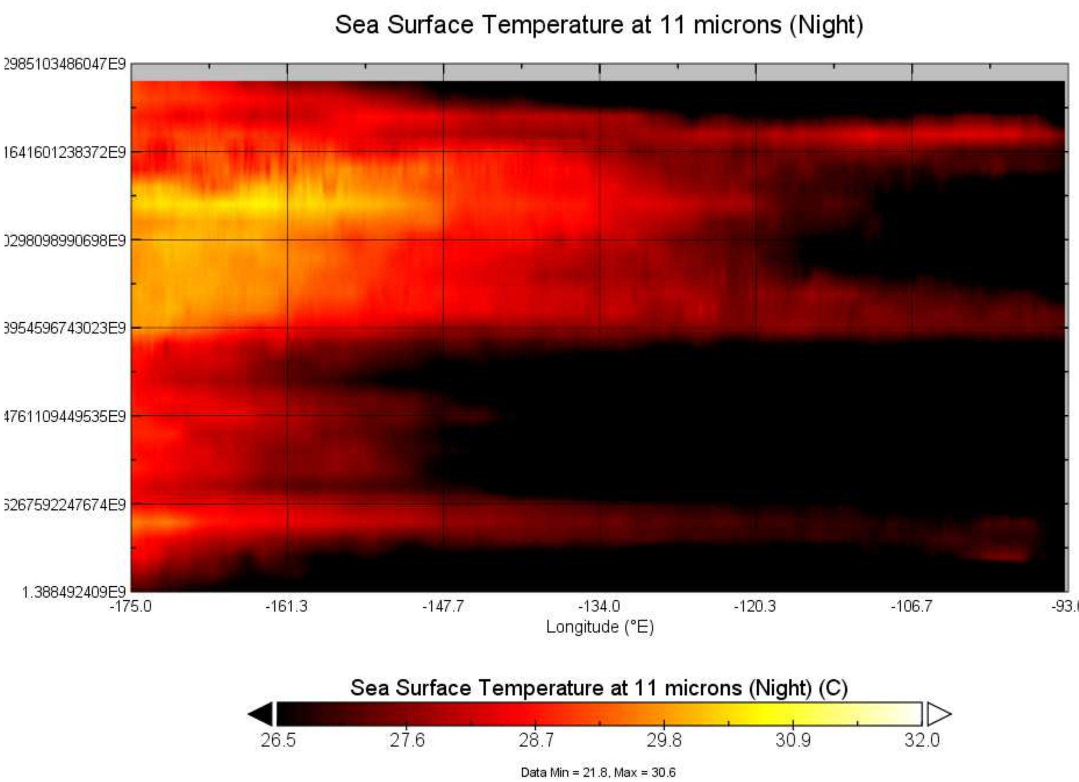


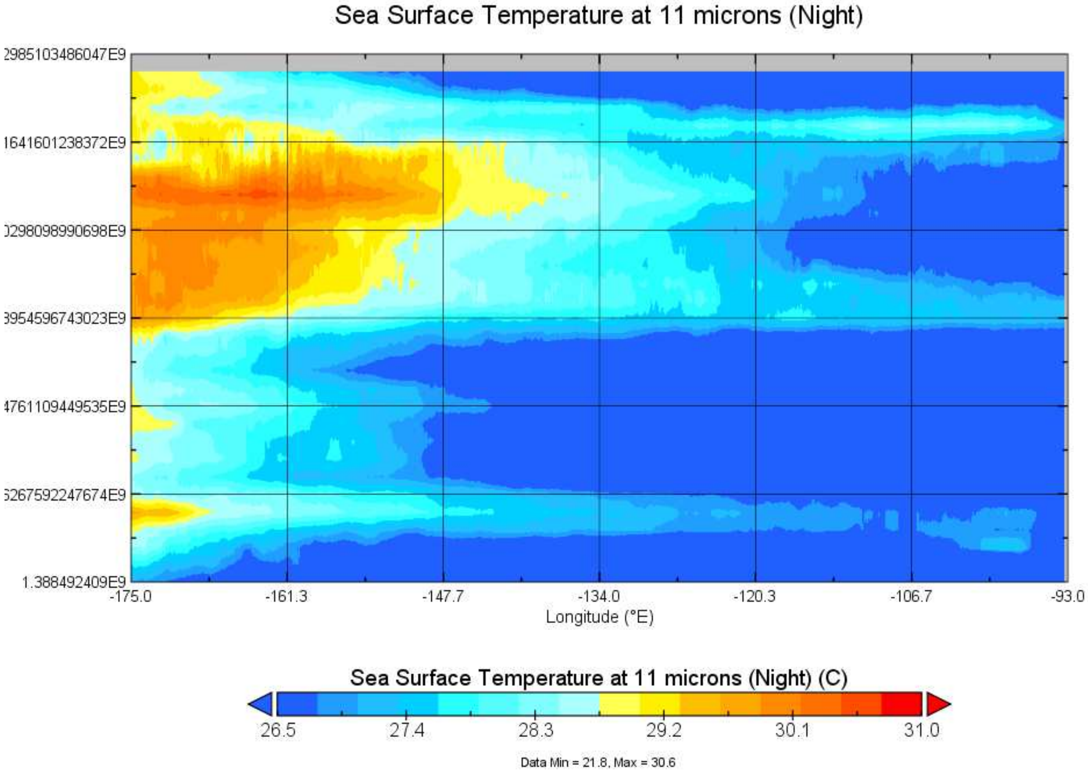
Longitude vs. time Hovmöller plot of SST in the equatorial Pacific Ocean, January 2014-June 2016.

This plot shows both the "almost" El Niño that occurred in 2014, and the full El Niño conditions occurring in 2015-2016. The plot also clearly demonstrates that this was a central equatorial Pacific El Niño, as the region of warmest waters did not extend much past 150 degrees W, so that the Galapagos archipelago and the western coasts of Central and South America were less impacted by the warmer surface waters. The size and strength of the event, however, did cause significant and observable meteorological phenomena, although it differed significantly from the strong 1997-1998 El Niño event due to the difference in where the warmest SSTs occurred (Paek, Yu, and Qian 2017).

The data file for this plot was exported to Panoply, which allowed the creation of the plots shown below, using different color palettes available in Panoply. The data was exported as a CSV file.


Sea Surface Temperature at 11 microns (Night)
2985103486047E9
1641601238372E9
0298098990698E9
3954596743023E9
4761109449535E9
5267592247674E9
1.388492409E9
-175.0
-161.3
-147.7
-134.0
-120.3
-106.7
-93.0
Longitude (°E)
Sea Surface Temperature at 11 microns (Night) (C)
26.5
27.6
28.7
29.8
30.9
32.0
Data Min = 21.8, Max = 30.6


Sea Surface Temperature at 11 microns (Night)
2985103486047E9
1641601238372E9
0298098990698E9
3954596743023E9
4761109449535E9
5267592247674E9
1.388492409E9
-175.0
-161.3
-147.7
-134.0
-120.3
-106.7
-93.0
Longitude (°E)
Sea Surface Temperature at 11 microns (Night) (C)
26.5
27.6
28.7
29.8
30.9
32.0
Data Min = 21.8, Max = 30.6



Longitude vs. time Hovmöller plots of SST in the equatorial Pacific Ocean created with Panoply.

The CSV file was opened with Excel. Utilizing the method developed by summer intern Melanie Ventura, which was adapted from the Difference Map method, the data for the plot was converted in Excel to an array of time (month) - longitude - data value triads. For this plot, which consists of 4 km resolution SST data, the data array contained 58,860 data values. A section of the array is shown below, showing that the warmest waters at 160 degrees W occurred in November and December 2015, and just exceeded an average value of 30 degrees C.

1/1/2014	-160.021	26.09777
2/1/2014	-160.021	26.71388
3/1/2014	-160.021	26.67101
4/1/2014	-160.021	27.42765
5/1/2014	-160.021	28.43884
6/1/2014	-160.021	28.02124
7/1/2014	-160.021	27.17178
8/1/2014	-160.021	27.57877
9/1/2014	-160.021	27.4864
10/1/2014	-160.021	27.69232
11/1/2014	-160.021	27.83864
12/1/2014	-160.021	27.54765
1/1/2015	-160.021	27.12189
2/1/2015	-160.021	27.16251
3/1/2015	-160.021	27.65065
4/1/2015	-160.021	28.47677
5/1/2015	-160.021	28.88593
6/1/2015	-160.021	29.36092
7/1/2015	-160.021	29.55812
8/1/2015	-160.021	29.60698
9/1/2015	-160.021	29.40014
10/1/2015	-160.021	29.796
11/1/2015	-160.021	30.17249
12/1/2015	-160.021	30.08168
1/1/2016	-160.021	29.47253
2/1/2016	-160.021	28.9997
3/1/2016	-160.021	28.68035
4/1/2016	-160.021	28.1911
5/1/2016	-160.021	28.27881
6/1/2016	-160.021	27.5836

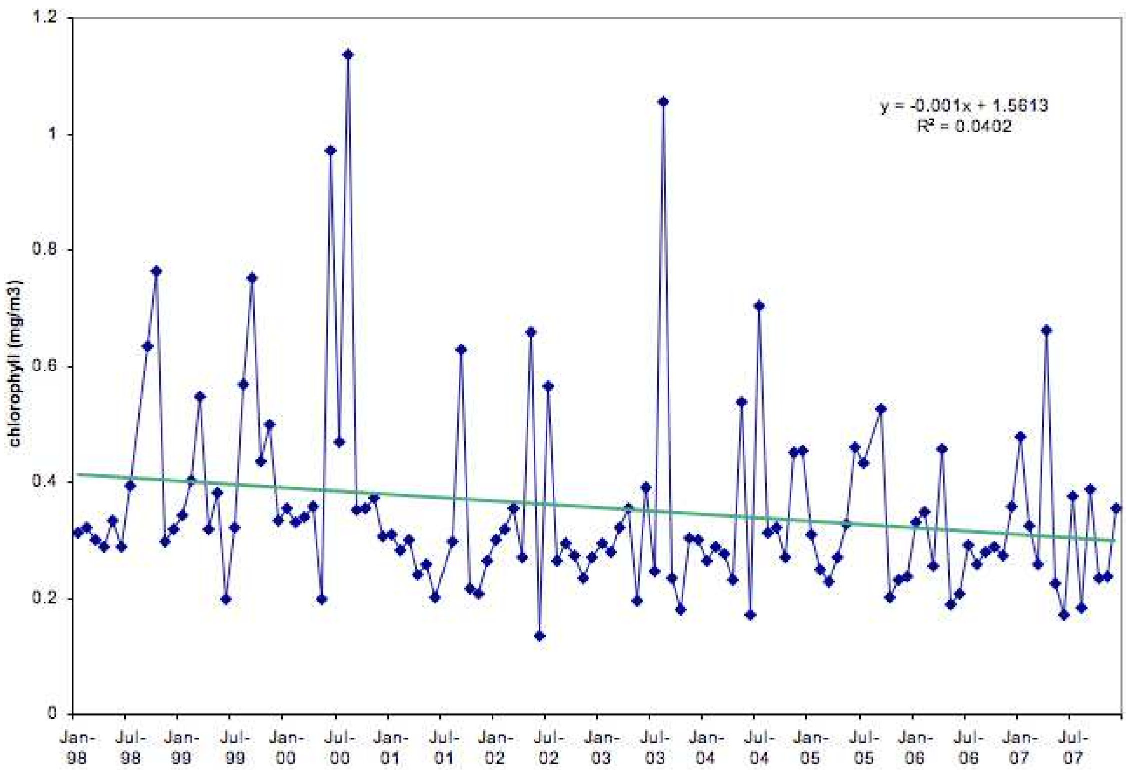
Excerpt from the output results in Excel for the longitude vs. time Hovmöller plot.

This example shows that the method can be used for many different data variables in Giovanni, and can be applied to large data arrays.

TIME-SERIES ANALYSIS

Acker et al. (2009) used the method described for time-series analysis to examine trends in chlorophyll *a* concentration (chl *a*) acquired by the Sea-viewing Wide Field-of-view Sensor (SeaWiFS) at the outlets of several major rivers around the world. This work built on previous use of Giovanni to examine chl *a* at the mouth of the Chesapeake Bay in the years 2002 and 2003 (Acker et al. 2005).

The outflow region of the Ganges River of India exhibited a clear trend of declining chl *a* concentrations over the study period from 1998 - 2007. This decrease was believed to be primarily related to decreasing discharge from the river due to increasing water usage and reservoir impoundment.



Time series of SeaWiFS chl *a* at the mouth of the Ganges River, 1998-2007.

Results of the statistical analysis applied to this time series are shown below. The p-value < 0.05 from the significance *f* test indicates that the observed trend was statistically significant.

	Study area distance from river mouth	Significance *f*	*f*-value	Slope	r^2	Average chl *a*
Ganges	275 km	p = 0.03	4.8	-0.001	0.0402	0.35 mg m^{-3}

There are now many more data sets in Giovanni that are long enough to provide good statistical results in a time series analysis similar to the one described here. It is also possible to export the time series data to Matlab or R to perform similar statistical analyses. In this era of climate change, many observed and modeled variables may exhibit significant trends.

AUTHOR INFORMATION

James G. Acker[1], Gary T. Alcott[2], Melanie Ventura[3], Jennifer C. Wei[2], David J. Meyer[2]

[1] NASA Goddard Earth Sciences Data and Information Services Center (GES DISC) / Adnet Systems Inc.

[2] NASA GES DISC

[3] Fairmont Heights High School, Capitol Heights, Maryland

ABSTRACT

The NASA Giovanni data analysis system provides a multitude of basic analysis capabilities for numerous Earth science data products which are available in the NASA Goddard Earth Sciences Data and Information Services Center (GES DISC) archive, as well as for additional selected data products provided by other NASA Distributed Active Archive Center (DAAC) archives. In Giovanni, users can easily generate time-averaged data maps, area-averaged time-series, Latitude-Time and Longitude-Time Hovmöller diagrams, correlation maps, accumulation maps, and map animations (22 analysis options are available in total). While ASCII text output is available for time-series plots, it is not included as an option for data maps. In order to provide a quantitative, easy-to-use numerical output in ASCII text form, the NetCDF file output from a Giovanni visualization is downloaded and then opened with the free NASA visualization software package Panoply. Panoply provides the capability of translating the Giovanni file into comma-separated-variable (CSV) output. Panoply also provides additional visualization options, including the facile calculation of difference maps and quasi-anomaly maps using Giovanni output files. The CSV files from Panoply can then be imported into an Excel spreadsheet, where an Excel macro converts the CSV files. The output consists of latitude-longitude-data value triads in text form for maps, and either longitude-time-data value or latitude-time-data value triads in text form for Hovmöller diagrams. This presentation will explicate the basic procedure for the conversion, and then provide several examples where the procedure is applied to Giovanni output from different analysis options.

REFERENCES

JOURNAL ARTICLES

James G. Acker, Lawrence W. Harding, Gregory Leptoukh, Tong Zhu, and Suhung Shen (2005) Remotely-sensed chl a at the Chesapeake Bay mouth is correlated with annual freshwater flow to Chesapeake Bay. (http://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2004GL021852) *Geophysical Research Letters*, 32, L05601, doi:10.1029/2004GL021852.

James G. Acker, Erin McMahon, Suhung Shen, Thomas Hearty, and Nancy Casey (2009) Time-series analysis of remotely-sensed SeaWiFS chlorophyll in river-influenced coastal regions (http://pdfs.semanticscholar.org/e11d/5028cadfbeeb11e76dbca629393da789d3dd.pdf). *EARSeL eProceedings*, 8(2), 114-139.

Paek, H., J.-Y. Yu, and C. Qian (2017) Why were the 2015/2016 and 1997/1998 extreme El Niños different? (http://www.ess.uci.edu/~yu/PDF/Paek-Yu-Qian.GRL.2017.pdf) Geophysical Research Letters, 44, doi:10.1002/2016GL071515.

J.C. Witte, B.N. Duncan, A.R. Douglass, T.P. Kurosu, K. Chance, and C. Retscher (2011). The unique OMI HCHO/NO2 feature during the 2008 Beijing Olympics: Implications for ozone production sensitivity. (http://www.sciencedirect.com/science/article/pii/S1352231011002500) *Atmospheric Environment*, 45 (18), 3103-3111.

WEB PAGES

How to Create Difference Maps with Giovanni and Panoply: Creating Quantified ASCII Text Difference Map Output with Giovanni, Panoply, and Excel. (http://disc.gsfc.nasa.gov/information/howto?title=How%20to%20Create%20Difference%20Maps%20with%20Giovanni%20and%20Panoply:%20Creating%20Quantified%20ASCII%20T

How to Create Hovmöller Plots with Giovanni and Panoply and Quantified ASCII Text Hovmöller Output with Giovanni, Panoply, and Excel (http://disc.gsfc.nasa.gov/information/howto?title=How%20to%20Create%20Hovm%C3%B6ller%20Plots%20with%20Giovanni%20and%20Panoply%20and%20Quantified%20ASCII%20Text%20Hovm%C3%B6ller%20Output%20with%20Giovanni,%20Panoply,%20and%20Excel)