Young youth explore geospatial data for citizenship project: A case study.

Dr. Daniel Zalles
SRI International
333 Ravenswood Ave.
Menlo Park, CA

One can theorize that when youth become aware that raw data sets can help them carry out appealing activities, their self-efficacy increases because they realize that if they explore the data themselves, they can draw their own conclusions. The following case study provides one example of this theory's enactment. Yet, Whether in formal or informal science youths' engagement with data, particularly raw data, is threatened when supports are insufficient for drawing meaning from the data. For example, there may be insufficient explanation of how the data are structured and how to make sense of the statistics that are displayed as forms of data reduction (e.g., mean, median, variance). These challenges are compounded when the youth need to also (1) become sufficiently proficient in using the technology for data collection, querying, and analysis; (2) recognize errors in the data from, for example, inadequate instrument and protocol calibration; (3) persist with their inquiry when their initial findings do not yield simple answers; and (4) learn enough about the focal scientific phenomena to draw meaning from the data in the first place (Feldman et al., 2009; Lehrer & Schauble, 2000; Zalles & Vahey 2005). With data sets collected by and for scientists, these challenges are often exacerbated by lack of accompanying information about how the data were collected, how they were measured, and how they are visually represented (Zalles, 2013a, 2013b; Radinsky et al., in review).

The following case study describes how a group of 7th-, 8th-, and 10th-grade youth in the informal environmental science program Green Kids have explored and used data from STORE and DICCE. sets that SRI International education projects have made more readily available and comprehendible to non-scientists. In so doing, they are an example of how youth can persist with challenging data tasks when the data are presented in transparent comprehendible ways. The two projects they examined for data are Studying Topography, Orographic Rainfall, and Ecosystems with Geospatial Information Technology (STORE) (NSF DRL-1019645) (store.sri.com) and Data-enhanced Investigations for Climate Change Education (DICCE) (NASA GCCE NNX10AT54A) (dicce.sri.com). First, an overview of STORE and DICCE is provided.

Overview of STORE
STORE is developing and piloting classroom uses of GIS-based interactive data files displaying climatological, topographical, and biological data about (1) an especially ecologically and topographically diverse section of mid-California and (2) a section of western New York State, plus projected climate change outcomes in 2050 and 2099 from an IPCC climate change model. The participating students and teachers live in those areas, hence the place-based focus of the project. To help teachers make curricular decisions about how to use these data with their students, the project has, with input from six design partner teachers, produced a curriculum module exemplar consisting of six lessons. The lessons start with basic meteorological concepts about the relationship between weather systems and topography, then focus on recent climatological and land cover data. The last two lessons focus on IPCC-sanctioned climate change projections in relation to possible fates of different regional species. Technology light versions of these lessons send students directly to map layers displaying the data for scientific analysis. Technology-heavy versions address the additional goal of building students' capacities to manipulate features of geographic information systems (GIS). Hence, the technology-heavy versions require use of the ARC GIS Explorer Desktop software, whereas the technology light versions are available in both the ARC software and in Google Earth. Google Earth makes possible some student interactivity such as drawing transects and studying elevation profiles, but does not support more advanced use of geographic information system technology such as queries of data-containing shape files or customization of basemaps and data

representational symbology. Answer keys are provided for each lesson. Teachers have in addition access to geospatial data files that display some storm systems that moved over California in the winter of 2010-2001 so that students can study relationships between actual data about storm behavior, topography, and climate averages. This provides the student the opportunity to explore differences between weather and climate. To increase the likelihood of successful classroom implementation and impact on student learning, the professional development process provides the conditions for teachers to make good adaptability decisions for successful follow-through. Teachers can implement the six lessons, adapt them, or design their own from scratch. The project requires that they choose from these options, explain their rationales for those decisions, and provide assessment information about student learning from their implementations. Teachers are given STORE assessment items aligned to each of the six lessons, plus some items that test how well students can interpret the STORE GIS data layers. To summarize, the STORE resources include (1) tutorials about how to use the two GIS applications, (2) sufficiently adaptive geospatial data available in free easily transportable software applications, (3) lessons that they can implement as is, adapt, or discard if they want to make up their own (as long as they use the data, and (4) supportive resources to build their content knowledge (such as overview documents about their states' climates and information about the characteristics of each data layer and each data set available to them). Figures 1 and 2
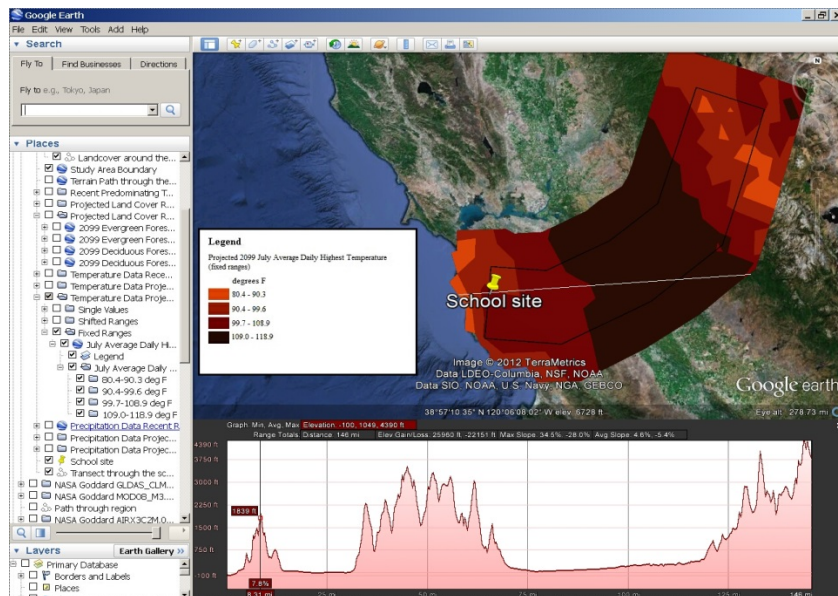

Figure 1. STORE data example – overlay of cross-section along a transect and projected 2099 temperatures in the California study area

Overview of DICCE

DICCE is making it much easier and more technologically feasible for middle and high school teachers and students to study climate change and related Earth system phenomena using data products from the Goddard Interactive Online Visualization and Analysis Infrastructure (GIOVANNI). GIOVANNI is a powerful NASA portal to Earth observation data that provides access to numerous data products on Earth system phenomena covering land biosphere, physical land, ocean biosphere, physical ocean, physical atmosphere, atmospheric gases, and energy and radiation system. These daily and monthly data products are derived from remote-sensing instruments on satellites, ground stations, and data assimilation models. DICCE is creating high school teacher and student access to some of these data to enable students to investigate their local climates. Teachers and students can query the GIOVANNI data archive and then save the results as map images, time series plots, vertical profiles, and data tables. The map images can

2

also be imported to appear as layers in Google Earth and other geographic information systems. DICCE-GIOVANNI (DICCE-G) is the name of this high school user-friendly access to the data. Figure 2 displays the forms of data representation that DICCE enables, including choropleth maps portable to Google Earth, time series plots, and ASCII test-delimited tables portable into Microsoft Excel and other spreadsheet programs.

The project has also produced DICCE-Learning Environment (DICCE-LE), a tool for teachers to author and adapt student data investigation activities and presentations on visualizations for their students via DICCE-G. The project strongly encourages teachers to focus on their local region and compare it either with other regions or with global data. Supports are provided to students and teachers about how to interpret trends in data products of their choice at the regional level, and a schema has been developed to help them understand how those data products fit into current scientific thinking about the certainties and uncertainties of global warming. DICCE-LE also supports teacher selection and adaptation of student assessment items for various data types such as amounts of carbon dioxide in the atmosphere, air temperatures, and precipitation.
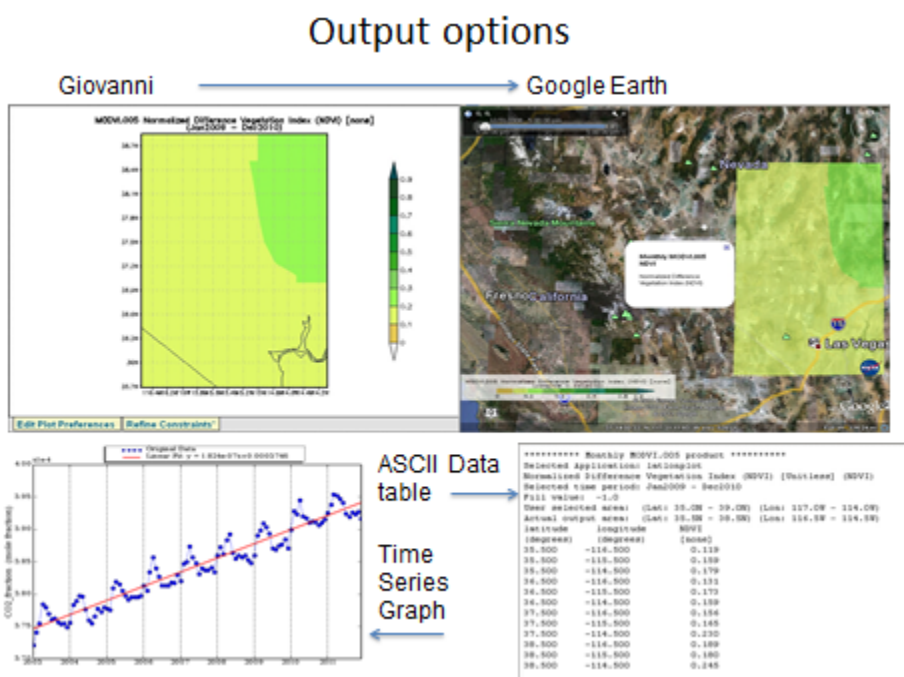

Figure 2. DICCE output options.

Case Study
The youth selected data about regional air temperature, cloud cover, and rainfall rate from DICCE to see if they can construct a viable model for predicting flash floods. In the process they looked critically at what other Web-based publicly available data sets might be helpful for informing the prediction model they were trying to develop. The names cited below are pseudonyms.

The flash flood-prediction project originated with one of the Green Kids youths, Ruben, a 14 year old 8th grader. In a June 2013 Green Kids Junior Scientist competition, Ruben presented his science research on and experimentation with a flotation device prototype that could be used during flood emergencies. He had been interested in super-storms and flash floods for more than 2 years and was educating himself about them. Then, with help from a Green Kids coordinator, Ruben contacted SPUR, a Bay Area nonprofit devoted to increasing public interest in finding strategies to cope with climate change in

California to see whether SPUR would be interested in helping him create a model predicting flash floods and then use the model in considering climate change adaptation strategies. The Green Kids staff had not had any previous connections with SPUR, but had heard of it, and thought that it would be good to tie their research to solutions that SPUR members are seeking. SPUR representatives responded that they were not currently interested in focusing on flash floods. Rather, based on their grantors' choices, SPUR was more interested in focusing on how to help Bay Area residents be prepared for sea level rises. The youth decided to proceed with the flash flood prediction work without SPUR.

Ruben and his teammate Olivia then turned to SRI's DICCE and STORE projects to see whether they provide data useful for their flash flood prediction goal. They studied the information on the STORE website about the data and how to examine it. Ruben and Olivia wanted to understand how data can help them make predictions, but they came to realize that, although useful for other science tasks, the STORE data sets were not helpful for testing the flash flood prediction model with data about prior flash flood events. They then turned to DICCE data, as well as data from the National Weather Service and other data sources agencies within the National Oceanographic and Atmospheric Administration. They then came up with the idea of considering flood events in relation to data parameters they chose to examine and then analyzing the data to see whether the data suggest predictable trends. This is work is still in progress. Ruben and Olivia have collected data for seven points of interest in California (e.g., the Bay Area, Sacramento, and the Mojave Desert) and are moving forward with formulating and testing a hypothesis about flash flood predictability. They are applying their analysis and conclusions to seven California locations of interests to understand the flash flooding phenomenon better. In an online space, Ruben wrote instructions for the other youth on his team to follow (see Figure).

The youth then designed a data representation (see Figure 2), which consisted of a time series plot showing flash flood events in relation to data from DICCE. The data show the average fraction of the sky covered with clouds over the particular geographical area and time range that the youth are investigating in the DICCE tool, plus over that same area and time the average monthly intensity of rainfall, and the temperature near the surface.

Steps for overlaying parameters:

1. Do an internet search and find out some actual dates of flash floods for your point of interest.
   a. Save the dates and the links from where you got that data into a separate word document. We would need to cite the references when we write the research paper.
2. Create a "Final" sheet in your workbook.
3. Copy the data from "Temperature of Land" which is in unit K (Kelvin) onto Column A in the Final sheet
4. Copy "Cloud Fraction" to Column B in the final sheet
5. Copy "RainFall rate" to Column C in the final sheet
6. From "Temperature of Land" sheet, copy the "month" to column D in the final sheet
7. Column E in the final sheet, we will be converting temperature to Farenheit, using the formula:
   a. $F = K * 9/5 - 459.67$
   b. =A4*9/5 - 459.67 – this is the spreadsheet formula
   c. Fill "down" so it is calculated for all the rows
8. Column F in the final sheet, we will be multiplying "cloud fraction" by 100
   a. =B4*100 – this is the spreadsheet formula
   b. The reason we are doing this is to overlay on the same Y-axis. It still keeps to the same trend of original data
9. Column G in the final sheet, we will be multiplying "RainFall rate" by 10000000
   a. =C4*10000000 – this is the spreadsheet formula
   b. The reason we are doing this is to overlay on the same Y-axis. It still keeps to the same trend of original data

The columns in the "Final" sheet should look like this:

	A	B	C	D	E	F	G
1							
2							
3	[K] Temp in Kelvin	Cloud Fraction - Actuals	Rainfall Rate actuals	Month	Temperature in Farenheit	Cloud Fraction	Rainfall Rate
4	316.292	0.28	0.000003	August 20, 2003	109.6556	28	30
5	314.89	0.205	0.000002	September 20, 2003	107.132	20.5	20
6	308.465	0.221	0	October 20, 2003	95.567	22.1	0
7	291.305	0.464	0.000005	November 20, 2003	64.679	46.4	50
8	286.862	0.465	0.000004	December 20, 2003	56.6816	46.5	40
9	288.227	0.414	0.000001	January 20, 2004	59.1386	41.4	10
10	288.769	0.509	0.000011	February 20, 2004	60.1142	50.9	110
11	302.08	0.28	0.000002	March 20, 2004	84.074	28	20
12	306.529	0.273	0.000004	April 20, 2004	92.0822	27.3	40
13	313.326	0.199	0	May 20, 2004	104.3168	19.9	0
14	318.64	0.208	0	June 20, 2004	113.882	20.8	0
15	319.913	0.225	0	July 20, 2004	116.1734	22.5	0
16	317.566	0.256	0.000001	August 20, 2004	111.9488	25.6	10

10. Now, to create the graph, select all the data on columns D, E, F, G including the headings

11. From the "Insert" menu, select the 2D line graph as marked below:
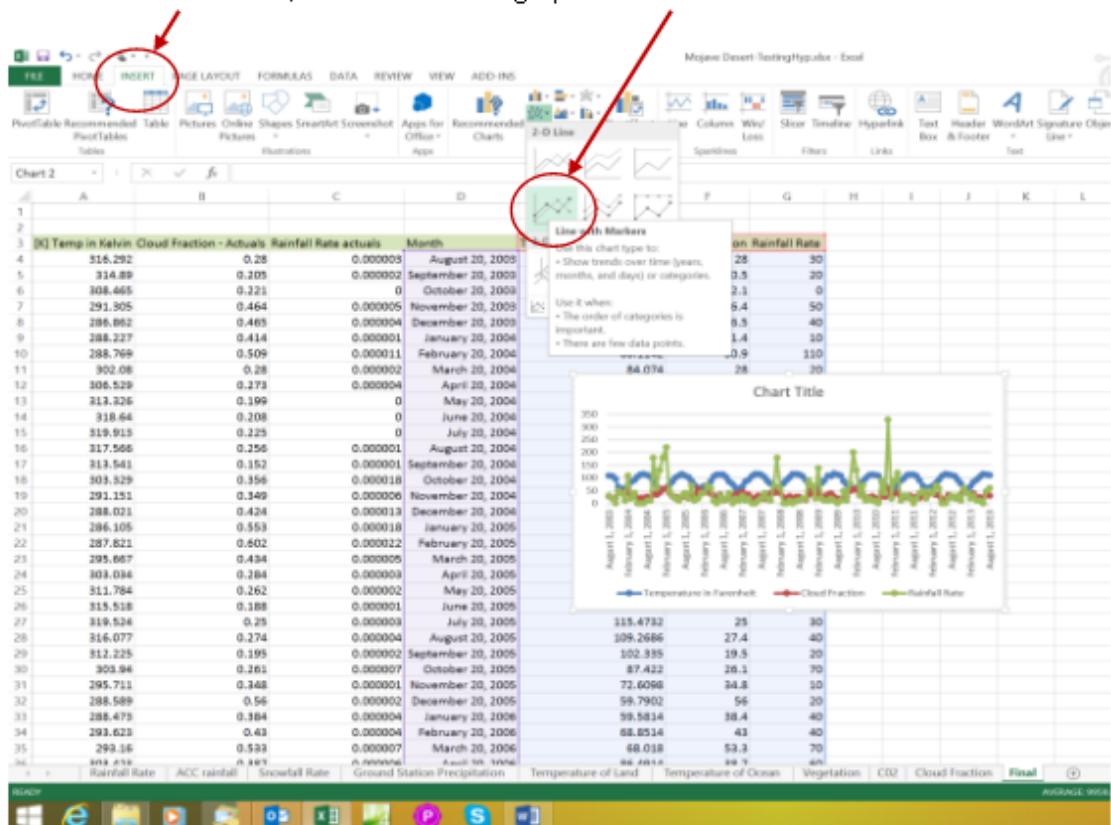
12. The graph produced is the overlay graph of all the three required parameters.
13. Stretch the graph and make it bigger, so you could see the details clearly
14. Now, add the dates that you obtained from Step 1 above on the "RainFall Rate" as shown in the key legend as "Flash Flood". I just used the cloud shape to mark this (you can find it from the Insert menu –Shapes)
15. You can also add the dates as shown below on the graph, it just makes it easy visually, pointing to the approximate temperature for that flash flood date.
16. We need to do all of these steps for all the Points of Interest

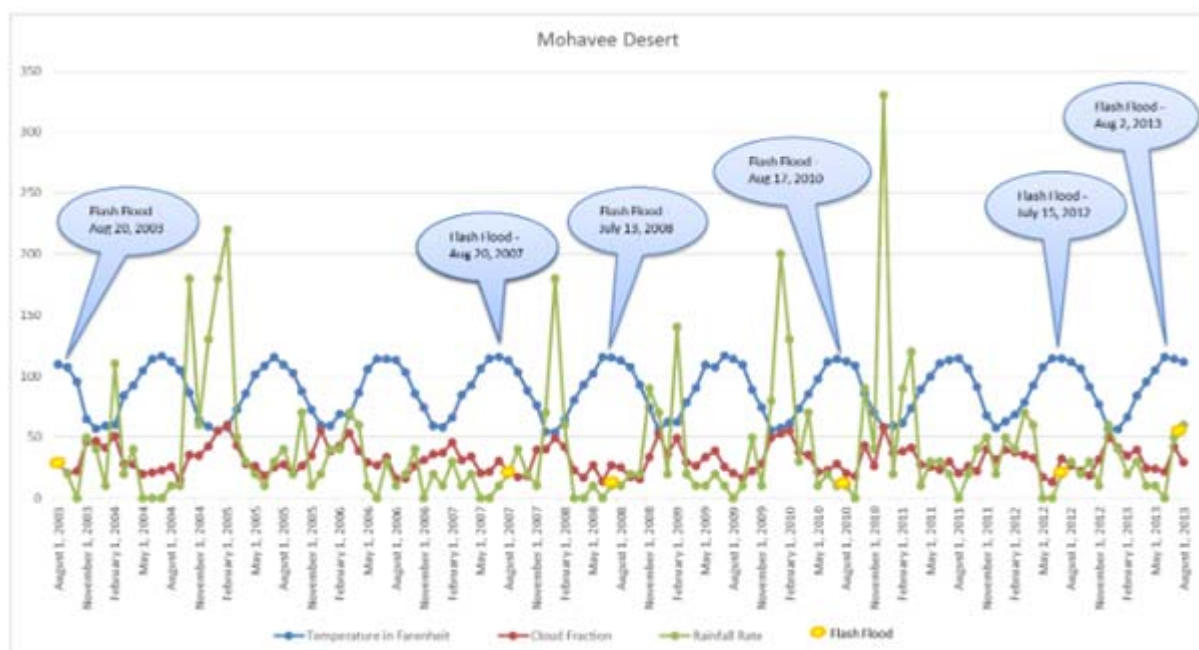Figure 2. Ruben's instructions to the other youth in his group

Figure 3. The youths' resultant data representation

Discussion
More research is needed to determine to what extend other youth who may be less motivated than these can be stimulated to be engaged with data from portals that, like STORE and DICCE, aim to make data more transparent and hence more accessible to non-scientists. Teachers need supports for how to facilitate the engagement, as do the leaders of informal out of school time programs.

Citations

Feldman, A., Konold, C. Coulter, B., & Conroy, B. (2009). *Network science, a decade later: The Internet and classroom learning*. Taylor & Francis e-Library edition. (Original work published by Lawrence Erlbaum Associates, Mahwah N.J., 2000)

Lehrer, R., & Schauble, L. (2000). Inventing data structures for representational purposes: Elementary grade students' classification models. *Mathematical Thinking and Learning,* (2), 51-74.

Radinsky, J., Hospelhorn, E., Melendez, J. W., Riel, J., & Washington, S. (in review). *Teaching American migrations with GIS census webmaps: A modified "backwards design" approach in middle-school and college classrooms*.

Zalles, D. (2013a). *Data-enhanced Investigations for Climate Change Education: Year 3 annual report*. MenloPark, CA: SRI International. http://dicce.sri.com

Zalles, D. (2013b). *Studying topography, orographic rainfall, and ecosystems with geospatial information technology: Year 3 annual report*. Menlo Park, CA: SRI International. http://store.sri.com

Zalles, D. R., & Vahey, P. (2005). *Teaching and assessing foundational data literacy.* Paper delivered at Annual Meeting of the American Educational Research Association, San Francisco, CA. http://www.sri.com/about/people/daniel-r-zalles