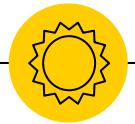
Renewable Energy Production



By Team Sunshine:
Maya Miller-Vedam, Laura Williams
Geoff Stirling & Boris Kletser

W205-2, Fall 2016

Our Hypothetical Challenge

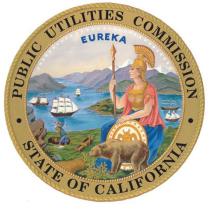
How might (near) real-time weather data better inform decision making around Solar Energy Production?

OUR GOALS:

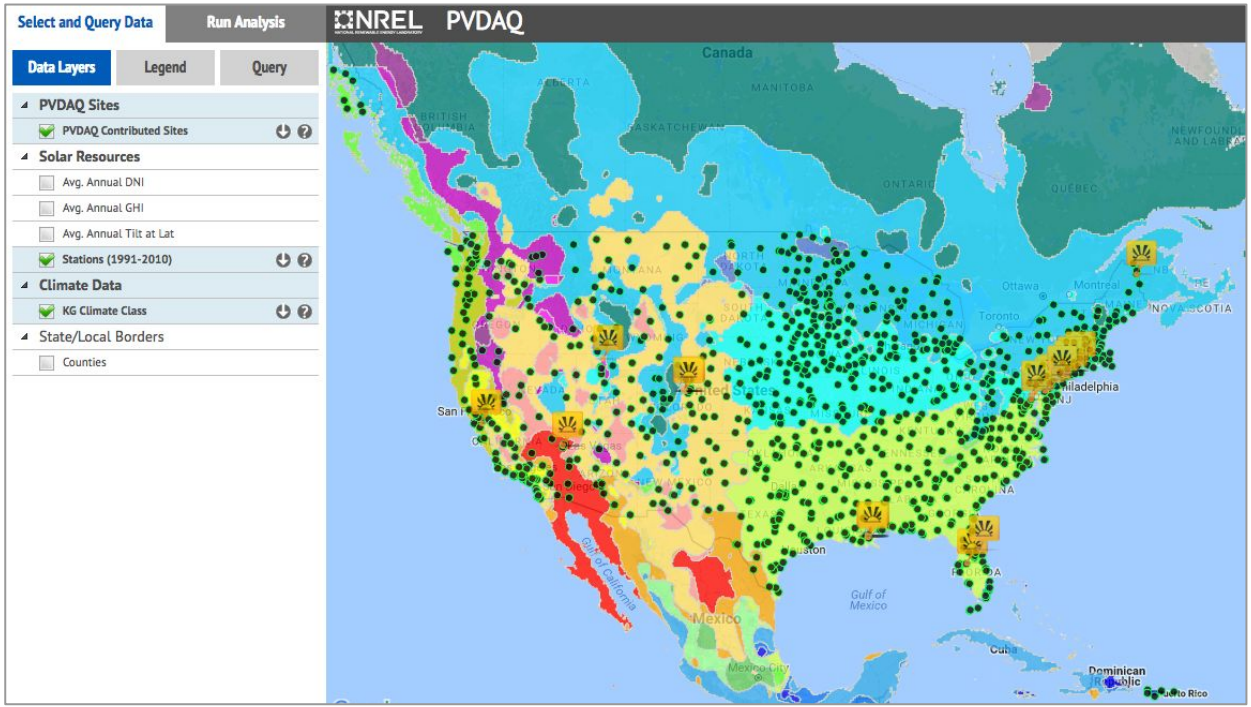
- Gain experience with open data in this problem space.
- Develop a decision relevant prototype of a data storage/retrieval system for solar energy stakeholders.
- Understand how this prototype might be extended to a more realistic industry setting.

Potential Clients and Stakeholders

- United States Dept. of Energy
- United States Environmental Protection Agency (EPA)
- Power companies and project developers
- Renewable energy research centers
- Regional power administration agencies
- Public utility companies
- State public utility regulatory agencies
- Environmental protection funding organizations
- Energy investors and traders



Existing Solar Energy Analysis



NREL ArcMap Data Analysis:

Multiple climate factors are mapped to energy production from photovoltaic panel installations.

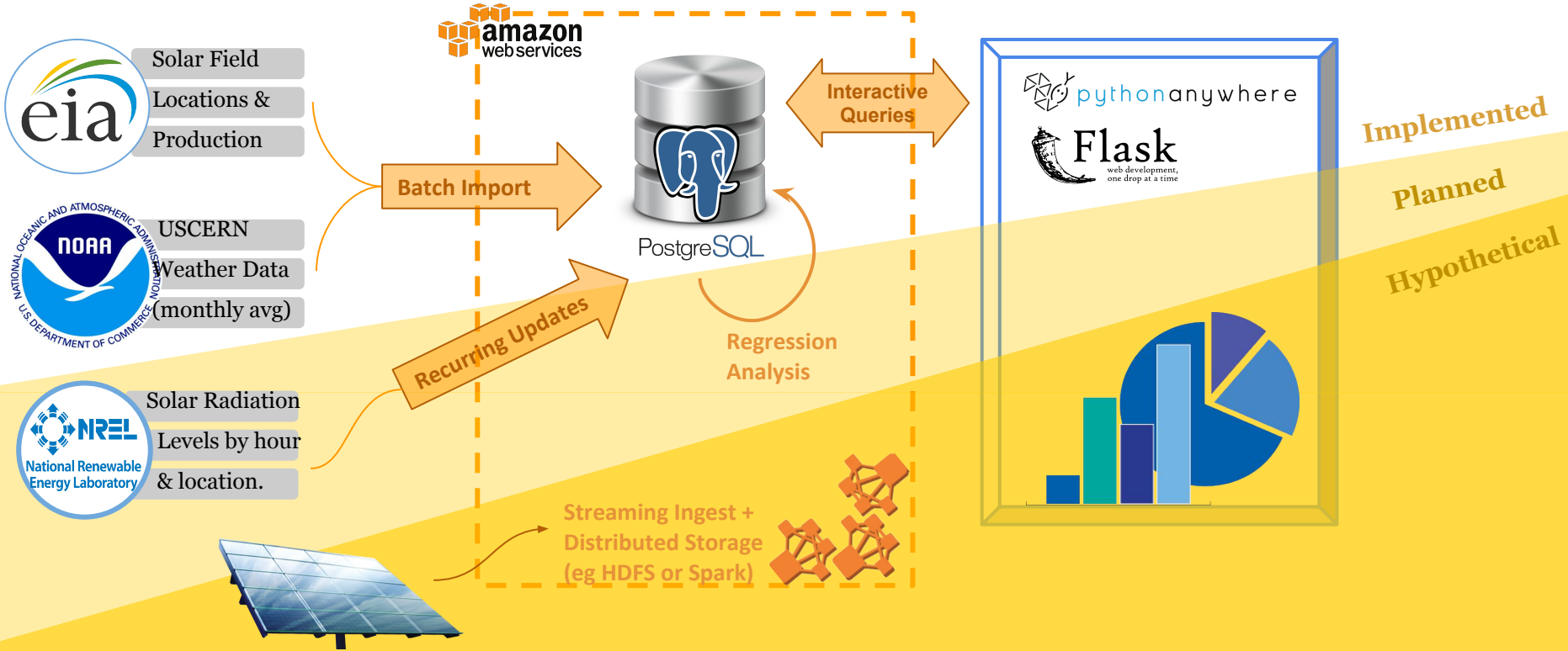
Data Sources - Overview

Source	Content Used	Variables
National Oceanic and Atmospheric Administration (NOAA) in collaboration with National Renewable Energy Laboratory (NREL)	Weather station locations and monthly weather data, in particular solar radiation	- Weather station ID, latitude and longitude - Monthly precipitation, solar radiation, and minimum, maximum and mean temperatures.
Energy Information Administration (EIA)	Monthly electricity generation per power plant	- Plant name, ID, latitude, longitude - Dates in year, month format - Net generation in megawatt hours

Data Sources - Details

Source	Format and Rate Limits	Frequency & Volume
National Oceanic and Atmospheric Administration (NOAA) in collaboration with National Renewable Energy Laboratory (NREL)	TXT and CSV files via FTP server Limit 1000 requests/day	- Frequency range: 5 mins. to yearly - 20 columns of numeric data - One month of weather data ~1.6MB
Energy Information Administration (EIA)	JSON via RESTful API No rate limit for non-robot applications.	- Monthly data aggregations - 10 columns of numeric data - All historic data - 7.5MB

Our Architecture

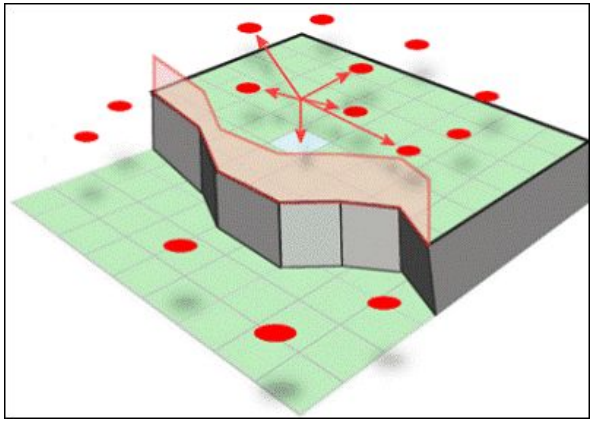


Technologies We Used

	Description	Rationale
Ingest	We wrote python scripts to access and clean government data ingested from RestfulAPIs and an FTP server. *Packages*: ***requests, pandas.***	An efficient, reliable method of pulling data on regular intervals in an automated fashion.
Processing & Storage	**Postgres DB on AWS**: We store our data in a SQL database housed on an Amazon virtual machine. *Packages:* ***psycopg2, scikit-learn, geopy.***	At the current stage of this application the data are quite small and our use case is very narrow, making a scale-up DB (Postgres) appropriate. Using AWS allows for the opportunity to build out more computationally costly analysis.
Serving	**Pythonanywhere + Flask:** We're pushing our data into pre-structured tables on a universally accessible website. *Packages:* ***flask.***	Flexible, easy-to-use format. Simple solution that requires little maintenance or overhead. Allows for API/DB queries with options to scale into more complex features

Technical Challenges

- **Data Cleaning:** *Tables broken up into multiple tables in some of the DBs. Solar Radiation numbers not available for all Weather Stations.*
- **Data-Linkage:** *Given the constraints in openly available data, our results will be less accurate than using private solar radiation data collected onsite at solar array installations.*



We are currently using the Nearest Neighbor to link weather and energy production. Ideally we'd use Inverse Distance Interpolation.

- **Missing Variables for Analysis:** *Newer solar cells are more efficient than older ones which may confound the model since we lack up to date information about the technology at various sites. A data scientist working in the industry would also have access to real time information about the angle of the sun and orientation of the solar cells.*

Project Results (In progress)

Team Sunshine Solar Power and Weather

Predicted Output For EOY

Aberdee | Look Up

December Predictio

wban_id	location
54933	Aberdee
54933	Aberdee
54933	Aberdee
54933	Aberdee

Plants in Best Conditions

Candidates for Expansion

name
Greene
Yates D
Bartletts
E C Gas
H Neely

Plants with Good Output vs Radiation

Worth Exploring, Investigating Technology

name	mwh	solar radiation	m/s
Bankhead Dam	467	6.4	73.0
Harris Dam	40	5.2	7.7
Barry	56	8.8	6.4
Bellefonte	27	4.5	6.0
Lay Dam	33	5.7	5.8

Future Opportunities

Wider Applications

This project could scale to other renewable energy systems, such as wind or hydroelectric, or to energy grids for integrated power management. Private industry data will yield more accurate and usable results.

Improved Predictions

Adding predictive sophistication using machine learning and other data mining techniques may leading to integration of additional data and improved accuracy of projections.

Storage & Serving

As we learn more about how stakeholders use this project, we may need to explore distributed file system and processing options, more efficient data ingest algorithms and/or different serving technologies.

Thanks!

Any questions ?