



Interactive Visualization of Global Long-Term Climate Data in a Web Application Using Python and Zarr in Amazon Web Services (AWS)

Matt Lammers, Craig Hoover, Christopher Cassidy, Robert Much

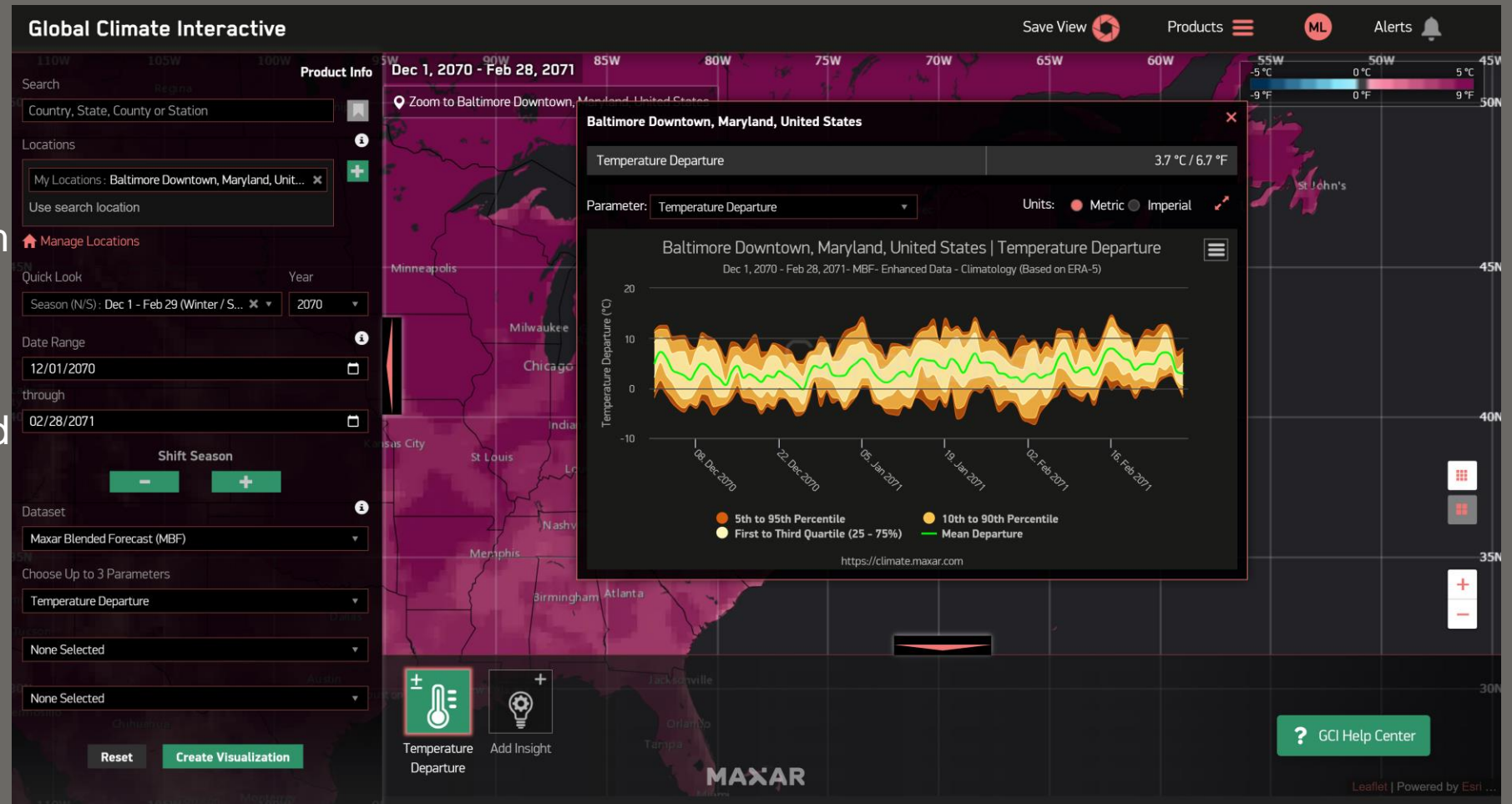
© 2024 Maxar Intelligence

MAXAR
INTELLIGENCE



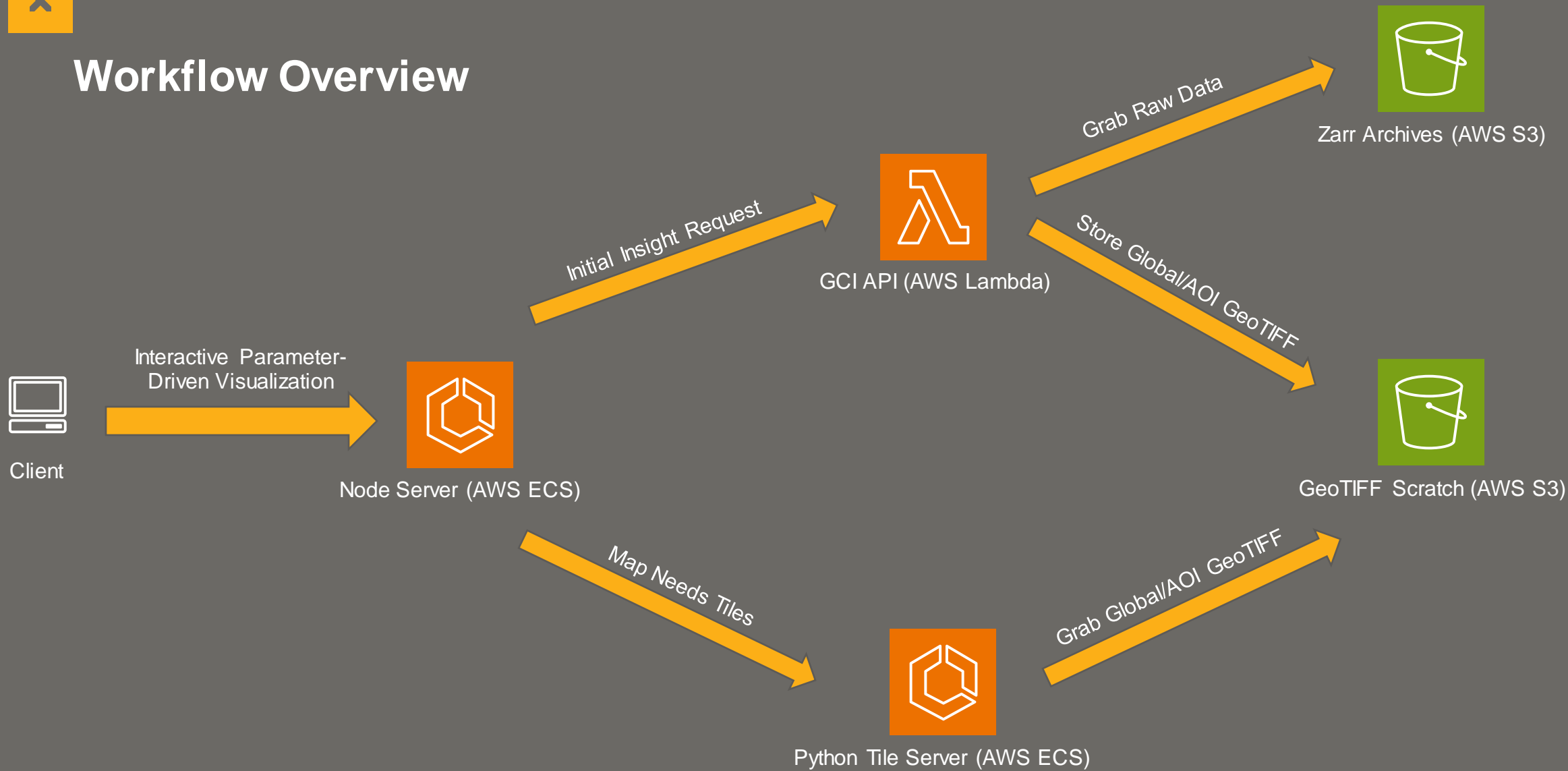
Maxar ClimateDesk Global Climate Interactive (GCI)

- Multiple models
- Global coverage at 25 km resolution
- Regional coverage at 1 km resolution
- Sub-seasonal to centennial time frames
- Proprietary Maxar Blended Forecast
- 18 TB of distributable raw data in easily consumable formats



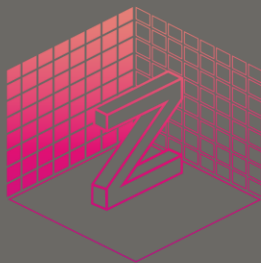


Workflow Overview





Storage



Zarr



AWS S3

- File format for descriptive n-dimensional arrays stored in a compressible, chunked way.
- Rapid access of random values or subsets from cloud storage using Python Xarray
- Data stored by model, variable, month/year

- Scalable object storage
- High speed throughput
- Reliable
- Durable (resistant to data loss)



Initial Requests

The screenshot shows the Maxar Intelligence web application interface. It includes a search bar for 'Country, State, County or Station', a 'Locations' section with a search input and a 'Manage Locations' link, a 'Quick Look' section with a date range selector (Quarterly: Jan 1 - Mar 31, 2036), a 'Date Range' section with start and end date pickers (01/01/2036 through 03/31/2036), a 'Dataset' section with a dropdown menu (Maxar Blended Forecast (MBF)), a 'Choose Up to 3 Parameters' section with three dropdown menus (% Normal Precipitation, Wind Departure, Minimum Temperature), and a 'Statistical Percentile (5 - 95) %:' section with a slider set to 75%. At the bottom, there are 'Reset' and 'Create Visualization' buttons. The 'Create Visualization' button is circled in orange.

Node Server (AWS ECS)

GCI API (AWS Lambda)

Zarr Archives (AWS S3)

GeoTIFF Scratch (AWS S3)

- Distributes API and tile server requests into asynchronous backend processes
- ECS enables auto-scaling based on load to rapidly spin up new servers
- Python function using Xarray to read Zarr data
- Transforms and merges raw data using Numpy
- Generates GeoTIFF output using OSGeo GDAL
- Each request is self-contained within Lambda, can run thousands simultaneously
- Global data selected for requested time range
- Climatology data also acquired for departure and “% normal” calculations
- Cloud Optimized GeoTIFFs (COGs) improve tiling performance
- Reduce costs with lifecycle policies that archive and/or delete older objects



Initial Requests

Product Info

Search
Country, State, County or Station

Locations
My Locations: Baltimore Downtown, Maryland, Unit...
Type to search

[Manage Locations](#)

Quick Look
Quarterly: Jan 1 - Mar 31 2036

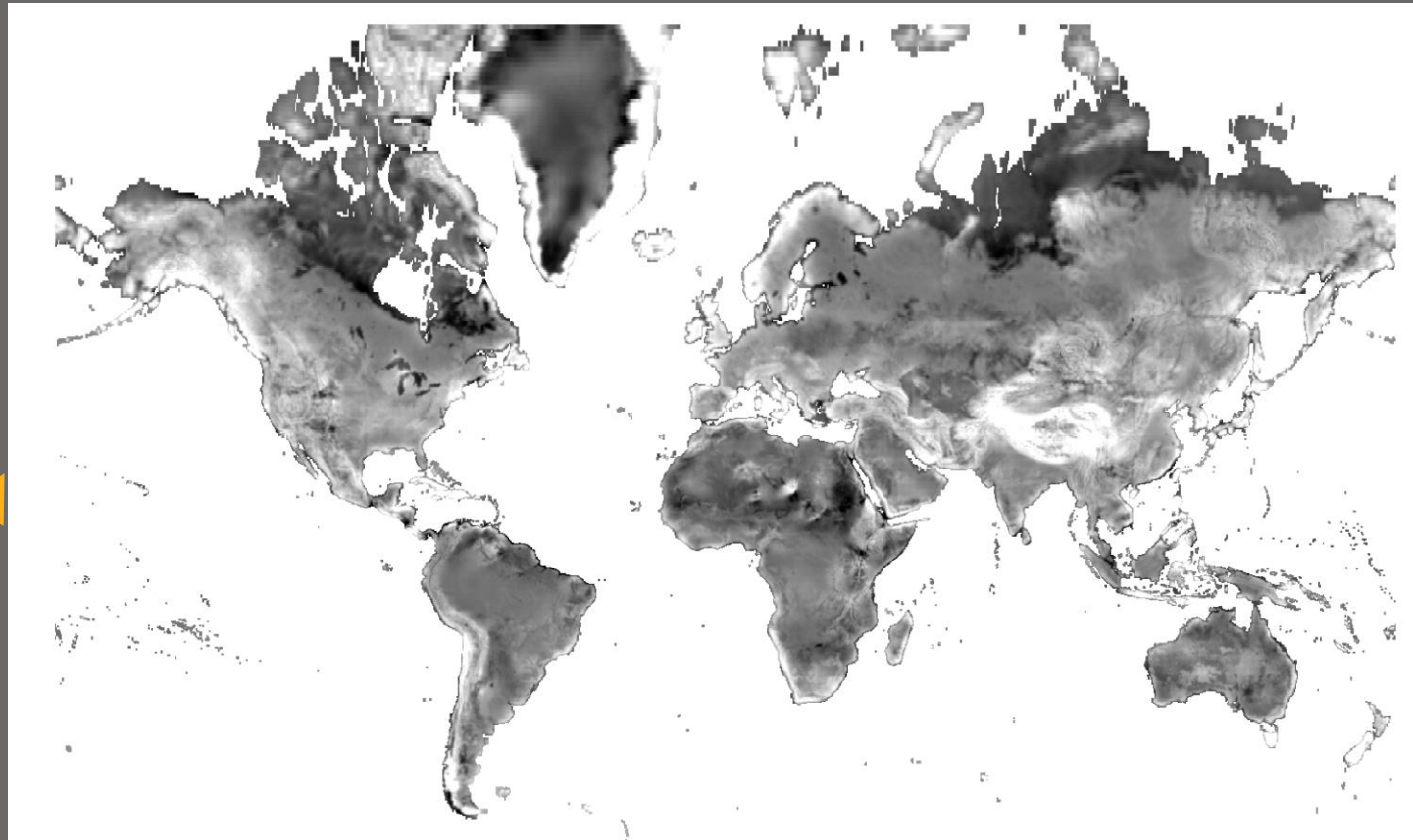
Date Range
01/01/2036 through 03/31/2036

Dataset
Maxar Blended Forecast (MBF)

Choose Up to 3 Parameters
% Normal Precipitation
Wind Departure
Minimum Temperature

Statistical Percentile (5 - 95) %: 75%

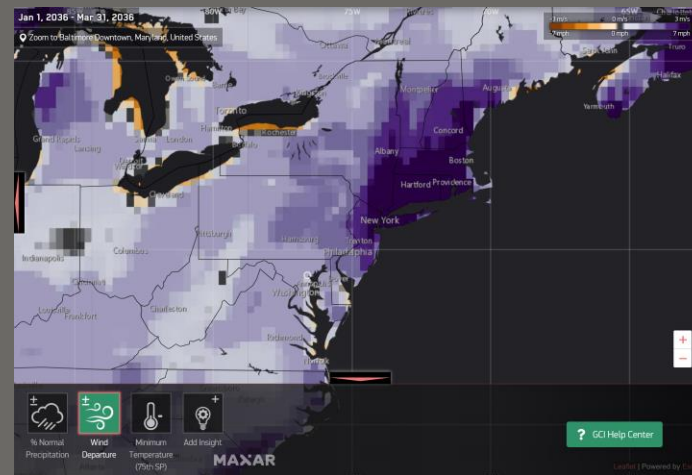
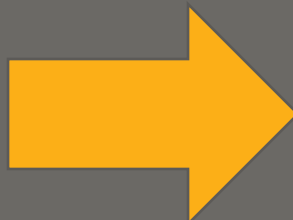
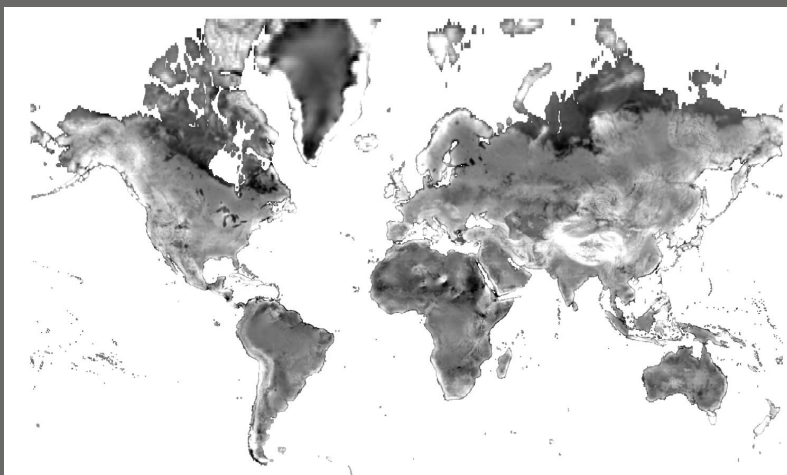
[Reset](#) [Create Visualization](#)



Resulting GeoTIFF



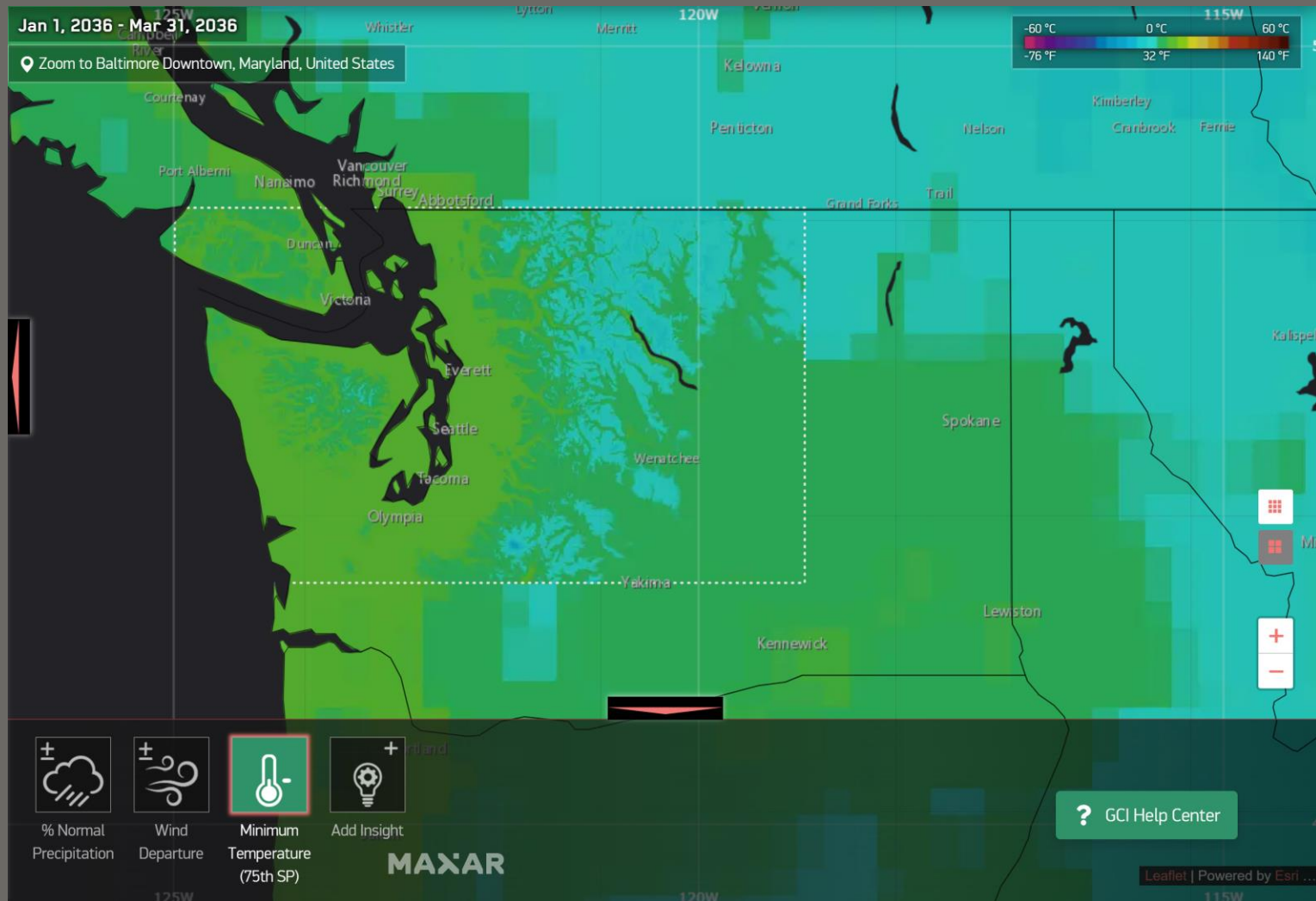
Image Processing and Handling



- Server requests TMS Z/X/Y image tile
- Python Tile Server (Tornado) running in AWS ECS captures request
- Rasterio and rio_tiler used to slice up GeoTIFF and apply a colormap
- Number of Tile Server ECS Instances scales based on demand



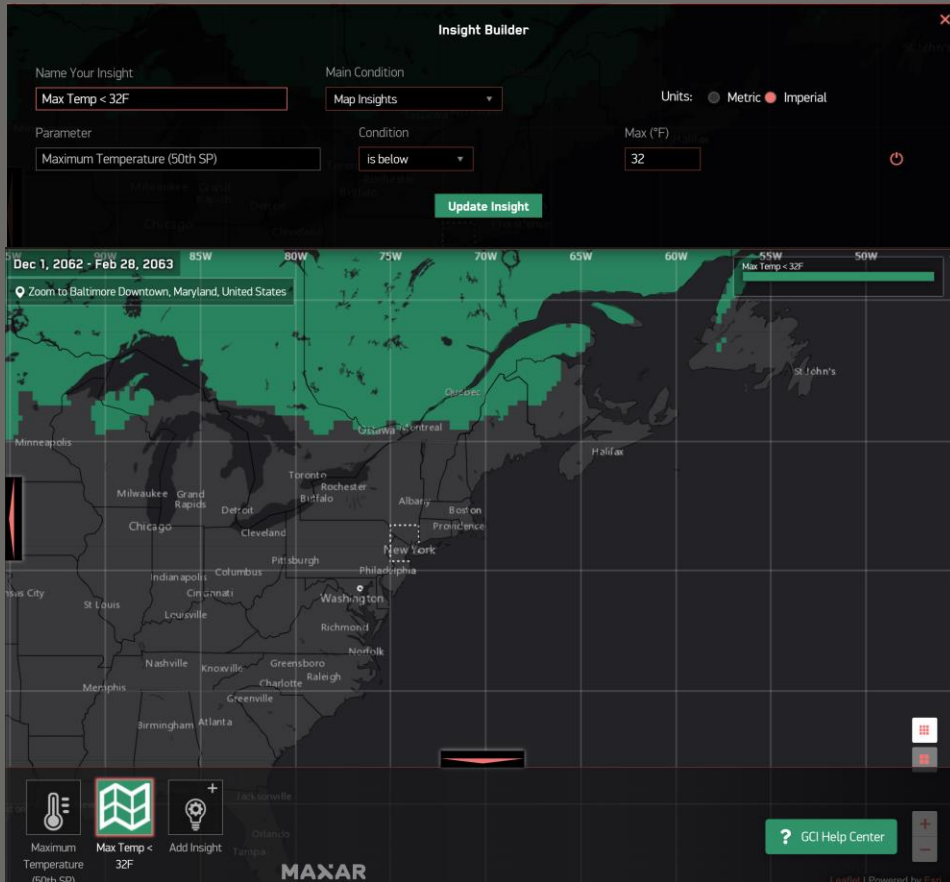
Adding Features: High-Resolution Areas of Interest



- 1 km data for customer AOIs
- Few modifications required to enable display



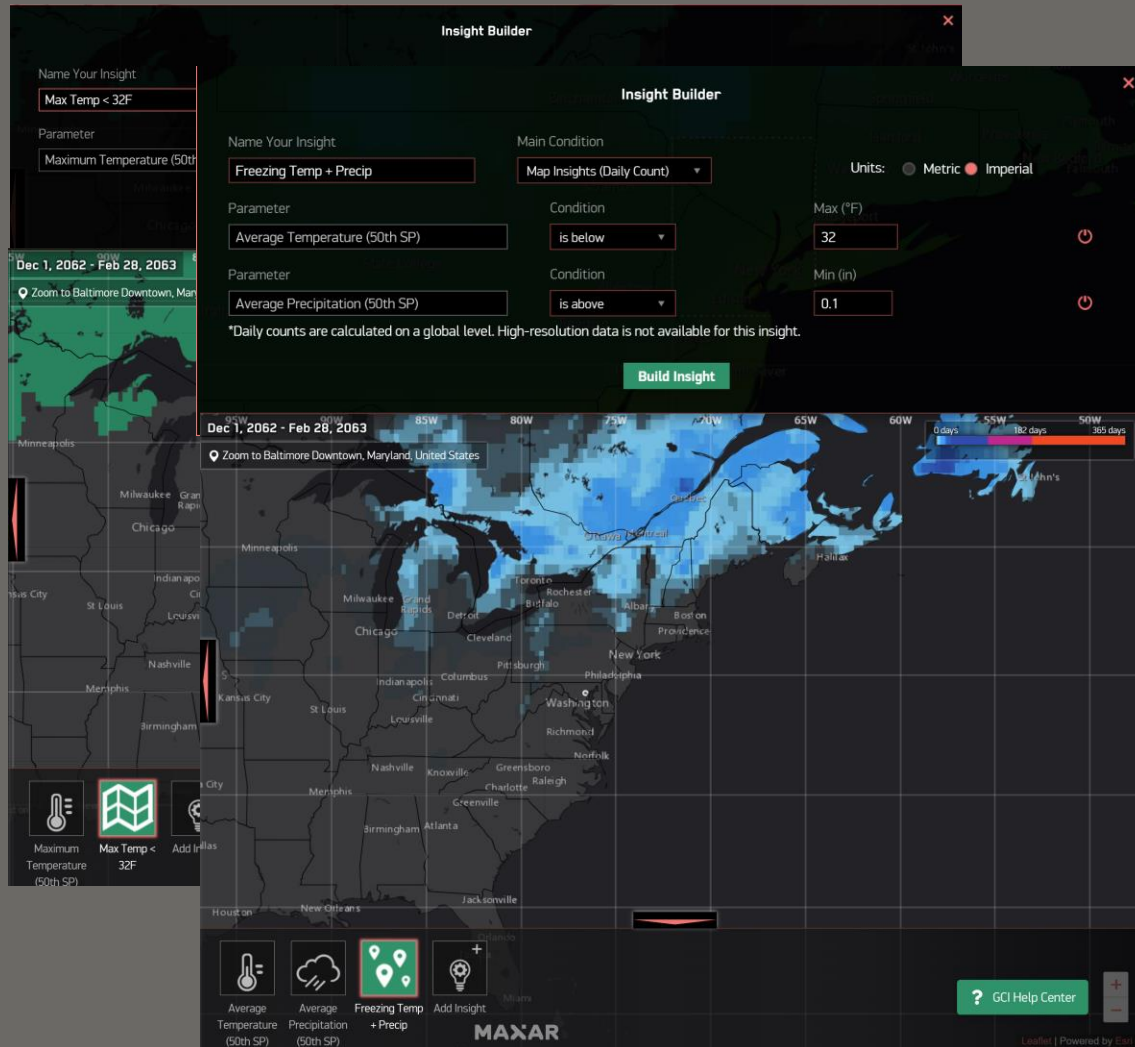
Adding Features: Custom Insights



- Areas exceeding threshold(s)
 - Average maximum temperature < 32F over a meteorological winter



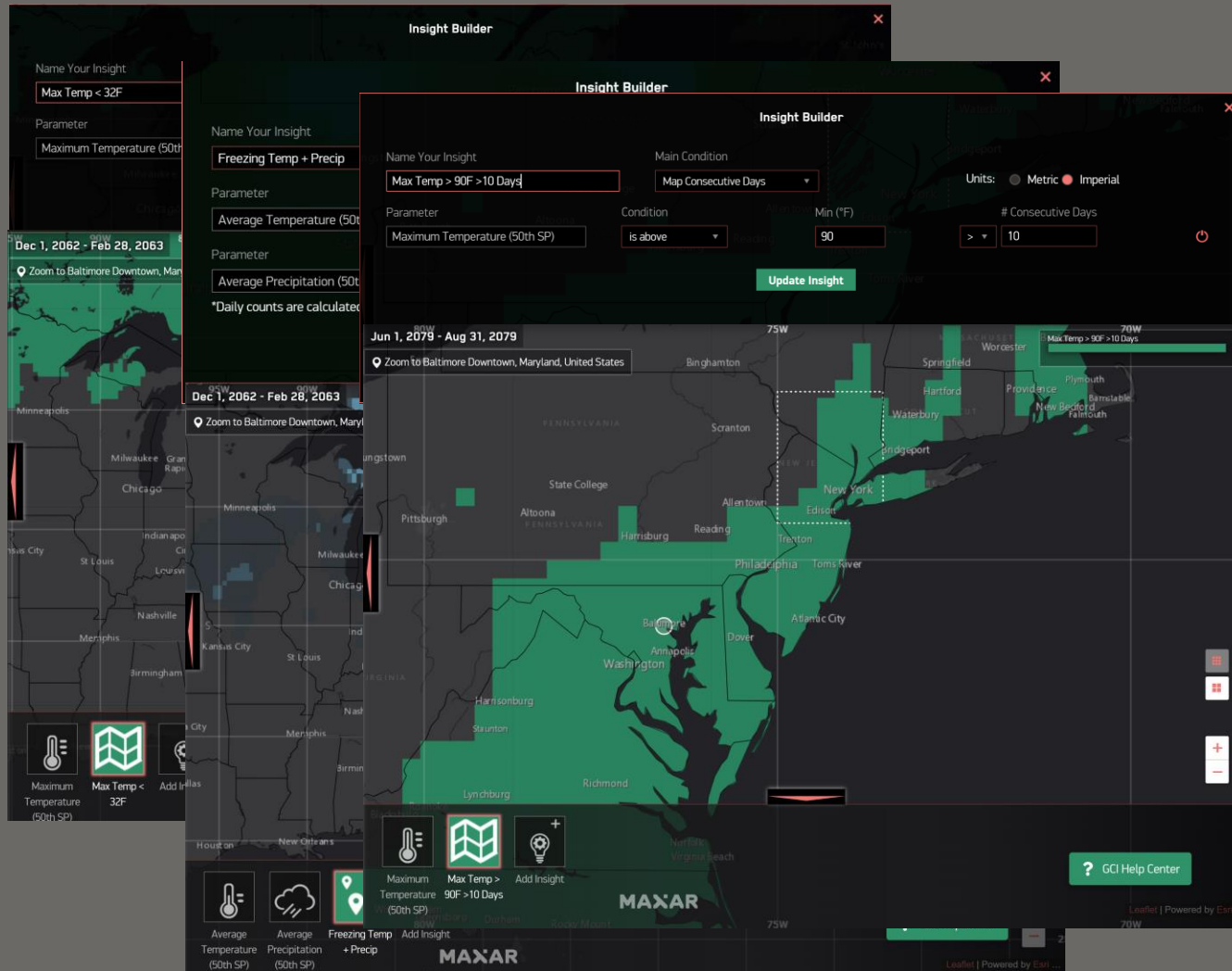
Adding Features: Custom Insights



- Areas exceeding threshold(s)
 - Average maximum temperature < 32F over a meteorological winter
- Days exceeding threshold(s)
 - Average temperature < 32F and >0.1" precipitation



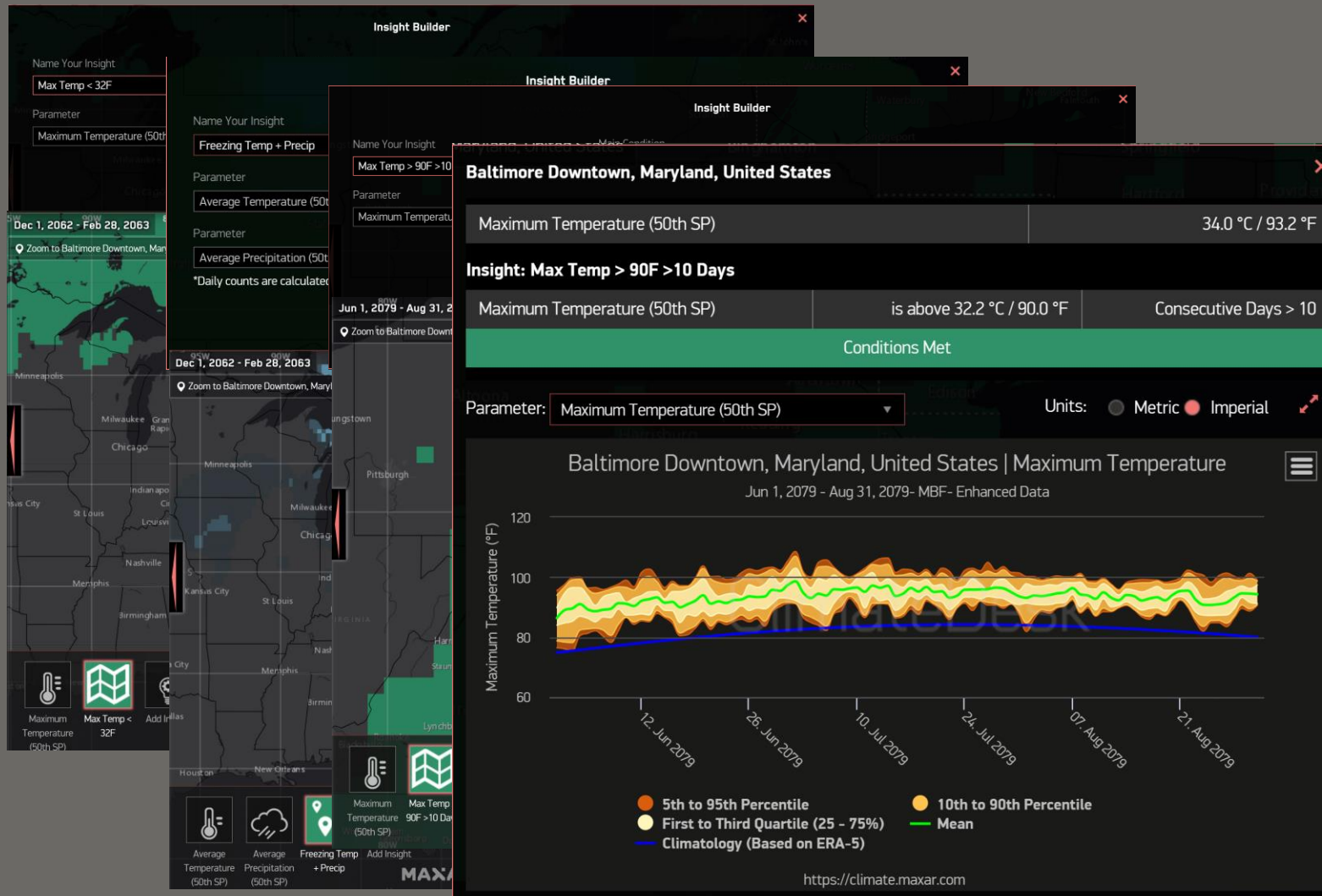
Adding Features: Custom Insights



- Areas exceeding threshold(s)
 - Average maximum temperature < 32F over a meteorological winter
- Days exceeding threshold(s)
 - Average temperature < 32F and >0.1" precipitation
- Consecutive Days exceeding threshold(s)
 - Maximum temperature > 90F for more than 10 consecutive days



Adding Features: Custom Insights



- Areas exceeding threshold(s)
 - Average maximum temperature < 32F over a meteorological winter
- Days exceeding threshold(s)
 - Average temperature < 32F and >0.1" precipitation
- Consecutive Days exceeding threshold(s)
 - Maximum temperature > 90F for more than 10 consecutive days
- Time Series
 - Python API can pull data and return data JSON just as easily as creating GeoTIFF



Summary

- Maxar's ClimateDesk Main Web Dashboard: Global Climate Interactive
 - Leveraging a variety of AWS tools to visualize efficient custom data requests
 - Open-source packages for data extraction, transformation, and image processing
 - Cons:
 - Internally-maintained
 - Relies on external support for a variety of libraries
 - Pros:
 - Flexible
 - Extensible
 - Customizable
 - Takes advantage of team's skills
-
- ClimateDesk Product Website: <https://www.maxar.com/products/climatedesk>



MAXAR.COM