Real-Time updates to GOES weather data processing

Justin Merz
Research Support Engineer
UC Davis Library
jrmerz@ucdavis.edu

Quinn Hart
Digital Applications
UC Davis Library
qjhart@ucdavis.edu

Outline

- What is the GOES-R (17/18)?
- Lightning Mapping
- Real-Time Data Processing
  - Architecture
- Thermal Anomalies
- Lessons Learned

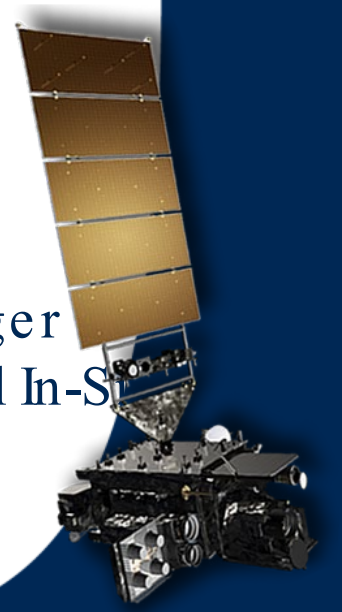
UC DAVIS LIBRARY

GOES-17/18 Weather Satellite

"The Geostationary Operational Environmental Satellite (GOES) – R Series is the nation's most advanced fleet of geostationary weather satellites. The GOES-R Series significantly improves the detection and observation of environmental phenomena that directly affect public safety, protection of property and our nation's economic health and prosperity."

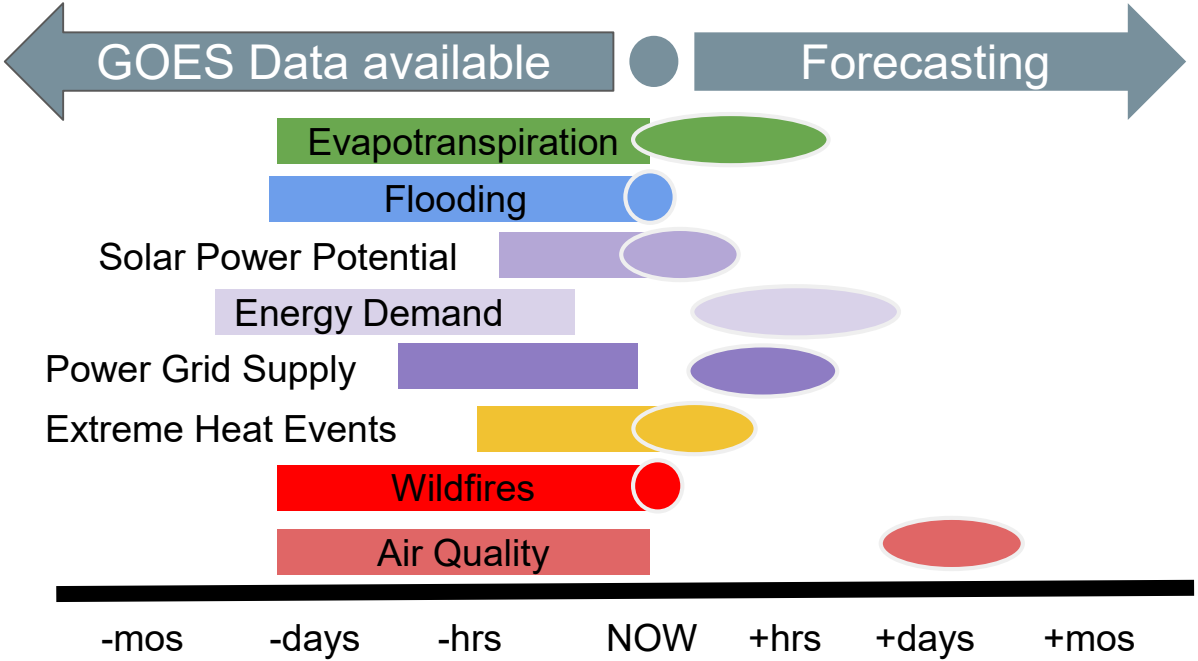
- Geostationary Lightning Mapper
- Advanced Baseline Imager
- Extreme Ultraviolet Ray Irradiance Sensor
- Magnetometer
- Solar Ultraviolet Imager
- Space Environmental In-S Suite

https://www.goes-r.gov/



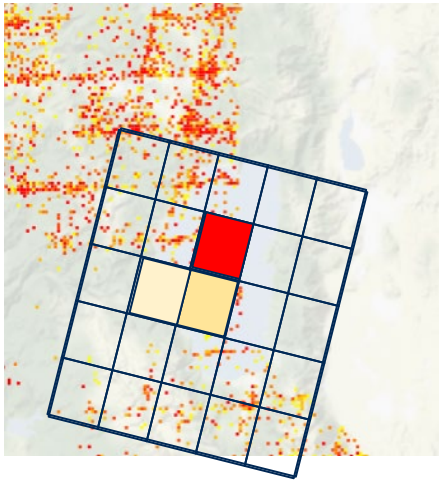
UC DAVIS LIBRARY

GOES Data- Example Use Cases



UC DAVIS LIBRARY

Geostationary Lightning Mapper (GLM)

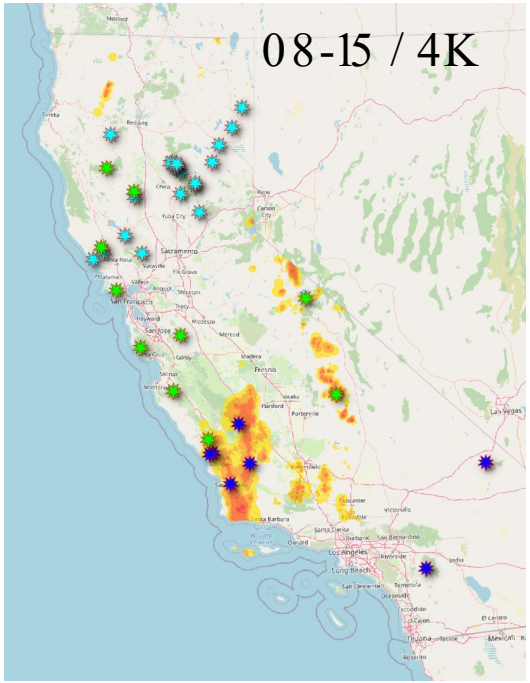


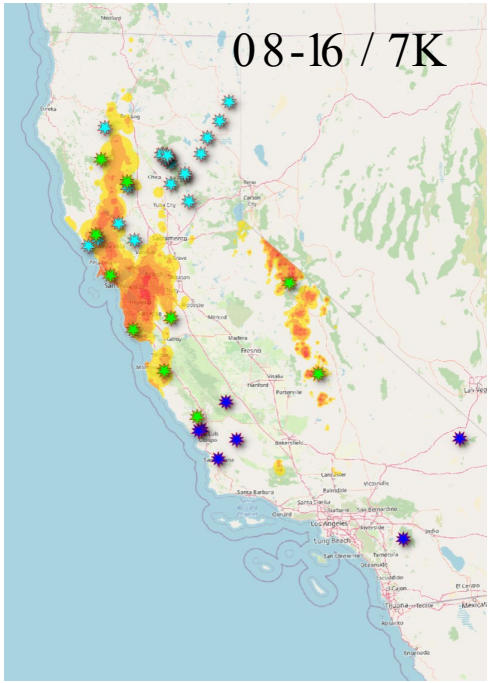
- Multiple strobes
- Combine location / time
- center, strength, duration

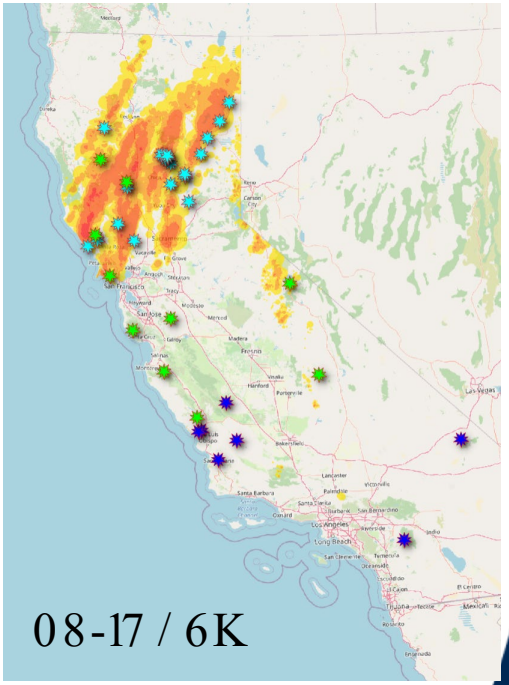
- Fetch NCL files (4000 / day)
- Simplify/Convert CSV (100M rows/yr)
- Gross subset (750K / yr in CA)
- Add to Postgres
  - PostGIS GIS
  - Quantization Functions
  - Summarize to a regular grid
  - Statistics

UCDAVIS LIBRARY

GLM Lightning - 2020-08 Siege

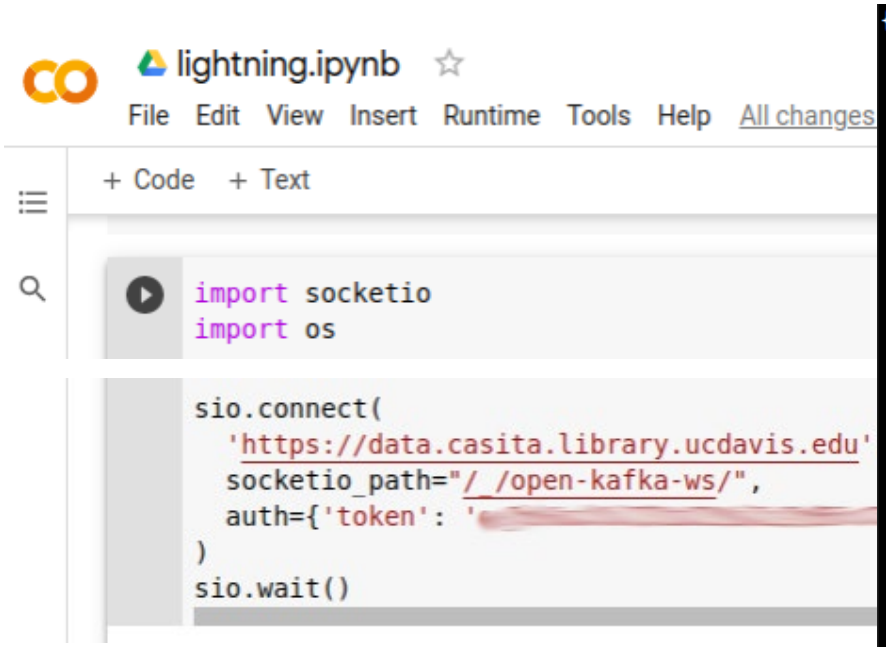






UCDAVIS LIBRARY

GLM - Real Time Processing

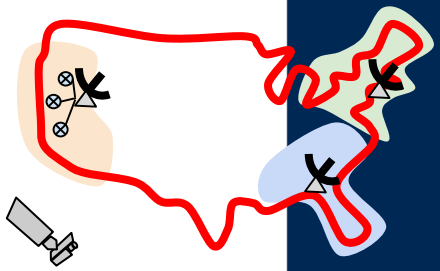



```
{
  "id": "lightning:2:1530086",
  "payload": {
    "properties": {
      "product": "lightning-detection-flash-data",
      "date": "2022-02-13",
      "hour": "15",
      "minsec": "01-00",
      "ms": "10867587",
      "apid": "302"
    },
    "data": [
      {
        "flash_id": 48998,
        "flash_time_offset_of_first_event": 24777,
        "flash_time_offset_of_last_event": 25590,
        "flash_frame_time_offset_of_first_event": 25097,
        "flash_frame_time_offset_of_last_event": 25910,
        "flash_lat": -19.067237854003906,
        "flash_lon": -133.26071166992188,
        "flash_area": 3226,
        "flash_energy": 158,
        "flash_quality_flag": 0,
        "flash_x": 11625,
        "flash_y": 14942
      },
```


UCDAVIS LIBRARY

CaSITA- GOES-R and Library

Speculative research project: goal is to; push the boundaries of the library's role in research data access, test out new technologies and service offerings.



Partnership: DWR and UC Davis
- Existing GOES Station

In-house domain knowledge
- Potential evapotranspiration

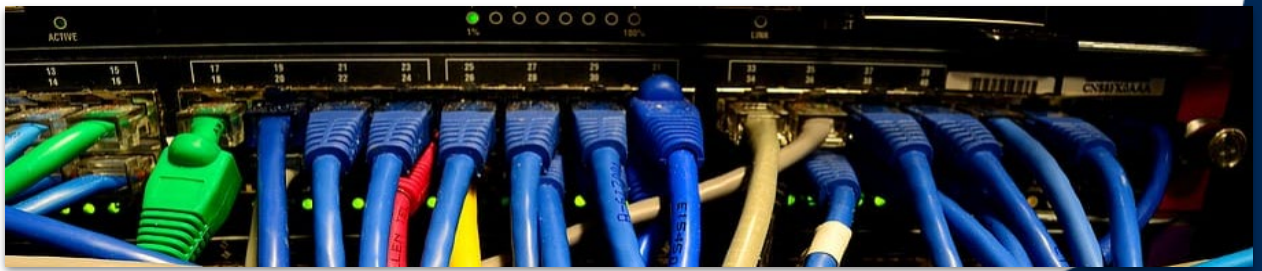
Data for multiple UC Davis departments / projects
- Share via library

- Exercise New technologies
- Scalable Cloud Compute
  - 100% containerized
  - Kubernetes (GKE) + Docker
- Realtime data access for campus researchers

UC DAVIS LIBRARY

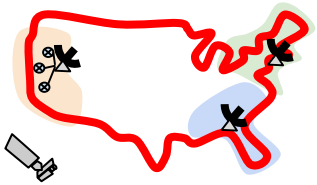
CaSITA Data Processing

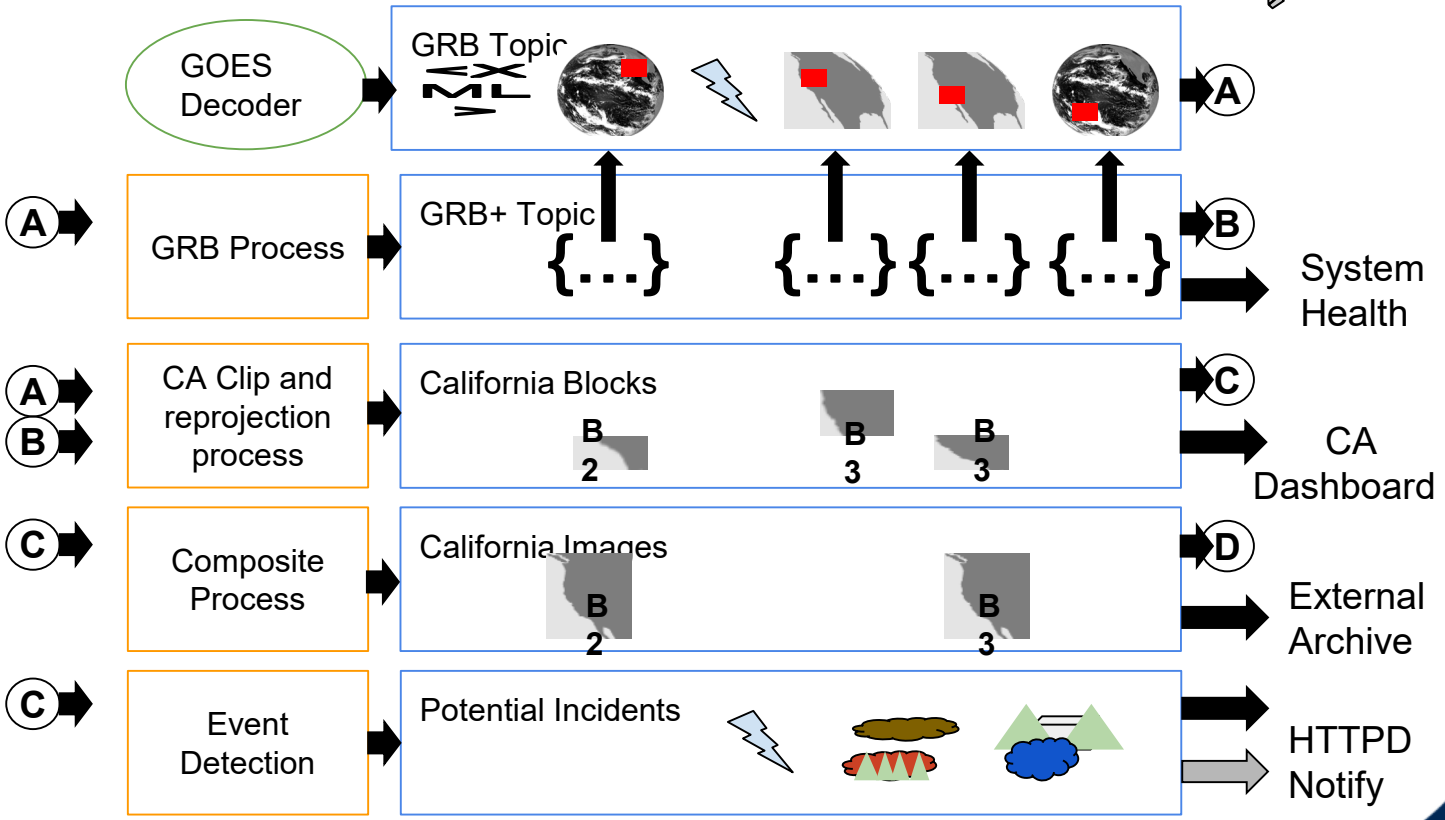
- ~**500GB** / day of data to disk
  - Library retains entire data product suite for 24 hours
- ~**30 CPU / 100GB memory** cluster
  - The cluster automatically scales depending on load
- ~**20 CPU cores** continuously decoding and compositing image products
- ~**42 processes** (pods/containers) running at a time.
  - Includes: workers, databases, message brokers and services.



UCDAVIS LIBRARY

Kafka Pipeline





Kafka Pipeline- File Access

~ / west / mesoscale / 2022-02-17 / 19 / 58-25 / 1 / d0 /

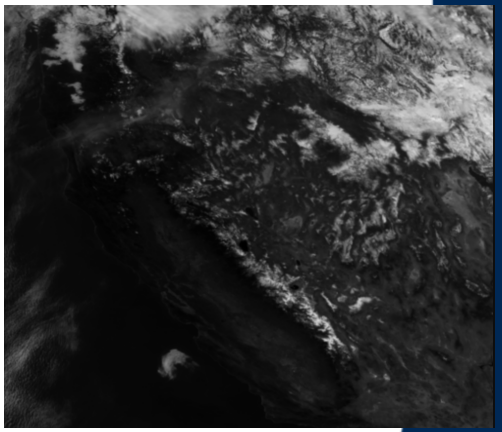
..	blocks	image.png
web-scaled.png	web.png	

~ / west / mesoscale / 2022-02-17 / 19 / 58-25 / 1 / d0 / blocks / 0-0 /

..	fragments	fragment-metadata.json
image.png	web-scaled.png	web.png

~ / west / mesoscale / 2022-02-17 / 19 / 58-25 / 1 / d0 / blocks / 0-0 / fragments / 0 /

..	image-fragment-metadata.json	image-fragment.jp2



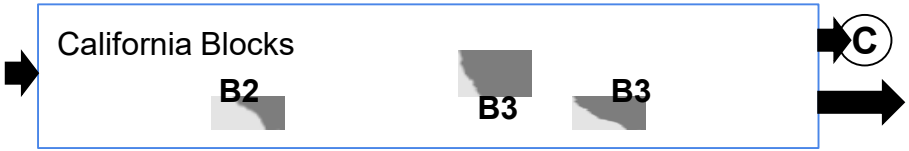
UC DAVIS LIBRARY

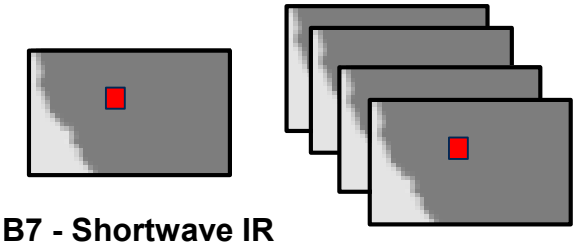
Kafka Pipeline- Events

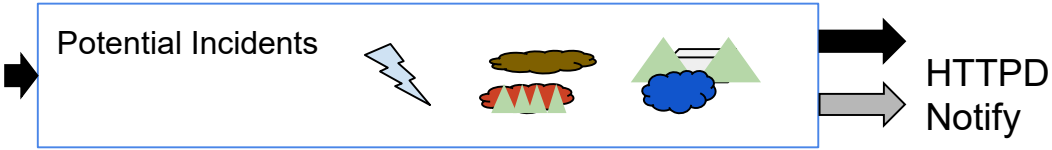
- Event notifications
  - **HTTP2** provides streams
  - **Websockets** via Socket.io library
- HTTP selection
  - https://data.casita.library.ucdavis.edu/_/h2/*
  - `/fulldisk/{{date}}/{{hh}}/{{mm-ss}}/2/91/blocks/{{b}}/image.png`
- Websockets
  - Authorized Access
  - Kafka Streams

UCDAVIS LIBRARY

Application Dev - Thermal Anomalies





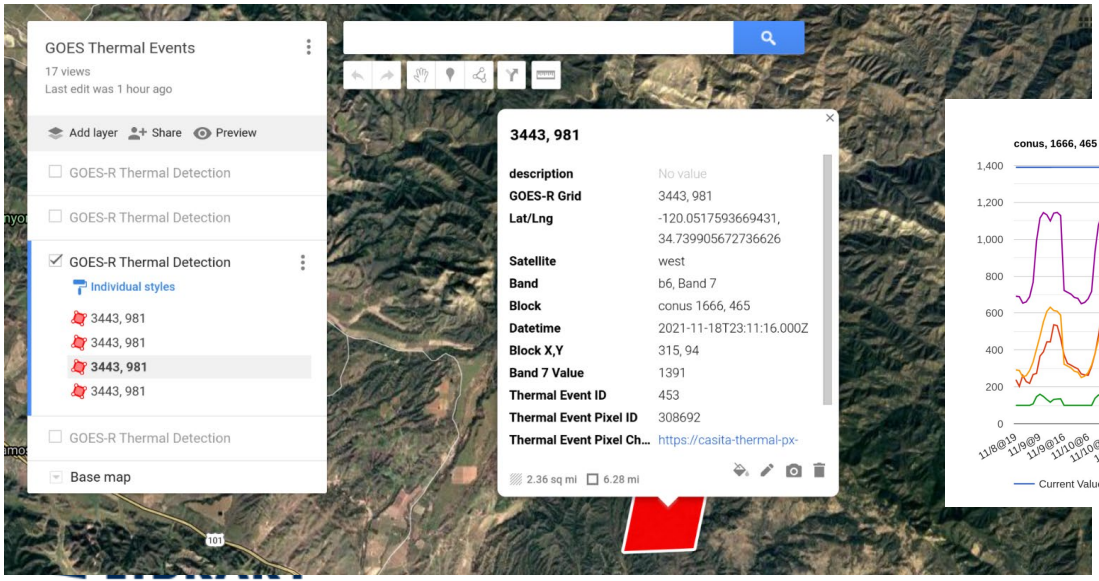


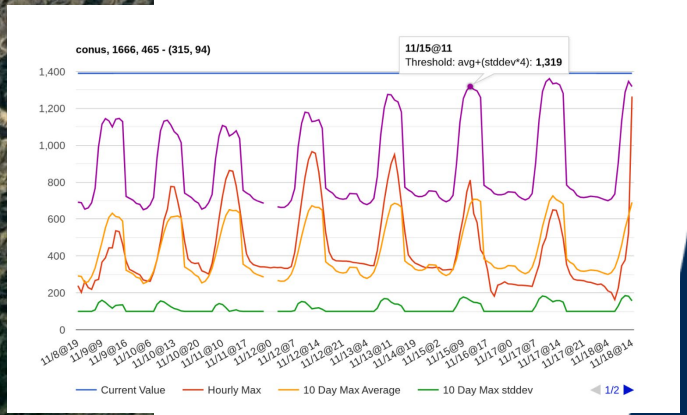
UC DAVIS LIBRARY

Application Dev - Thermal Anomalies

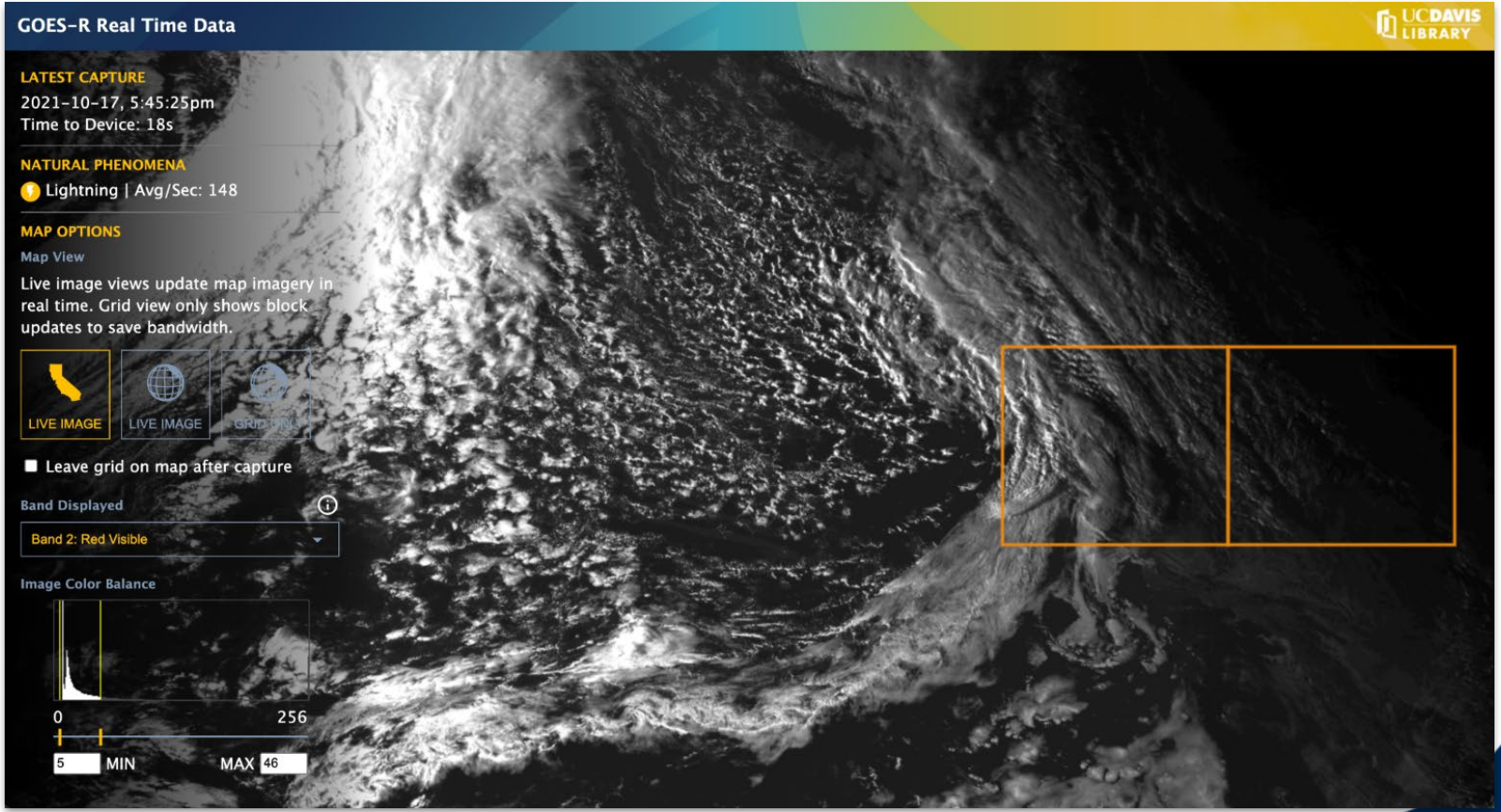
GOES-R - Thermal Event Notification APP 6:30 AM

🔥 New Thermal Event - {"thermal_event_id":822} - https://data.casita.library.ucdavis.edu/_/thermal-anomaly/kml/network?thermal_event_id=822





Demonstration App - CaSITA



Next Steps

- More CA Specifics
  - Band Ratios / Differences
  - Real-time cloud mask
  - Projections to CA?
- Standardize Alerts
- Just in Time Processing
- Machine Learning Partner
  - Solar Output predictions

UC DAVIS LIBRARY

Lessons Learned- Technical

- Process monitoring is critical
  - Logs find the needles in the haystack
- Scope the problem
  - Subset data to researchers needs
- Messaging systems: learn them, love them
  - Kafka, Websocket.io

- Alternatives now exist
  - Google Earth Engine

UC DAVIS LIBRARY

Lessons Learned- Science

- GOES Science data products
  - Complex
  - ATBDs not always complete
  - External data requirements
- Can Easier Interfaces jumpstart algorithms ?
  - Jupyter Notebooks
  - Simple Tech
    - HTTP Events
    - Standard image formats
    - JSON

UC DAVIS LIBRARY

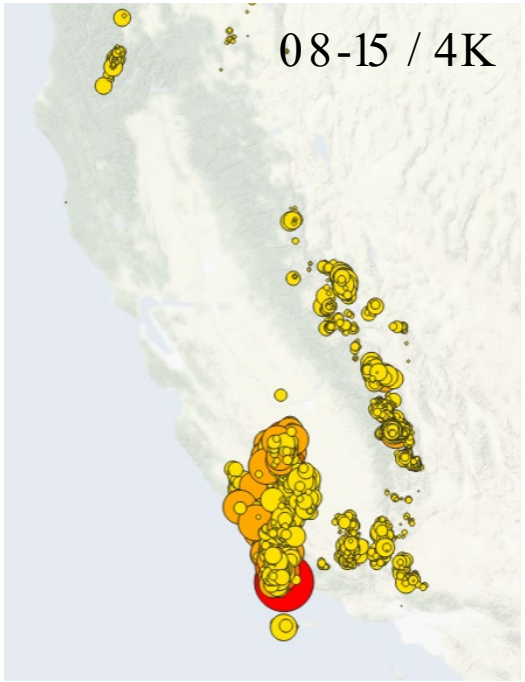
Thank You -

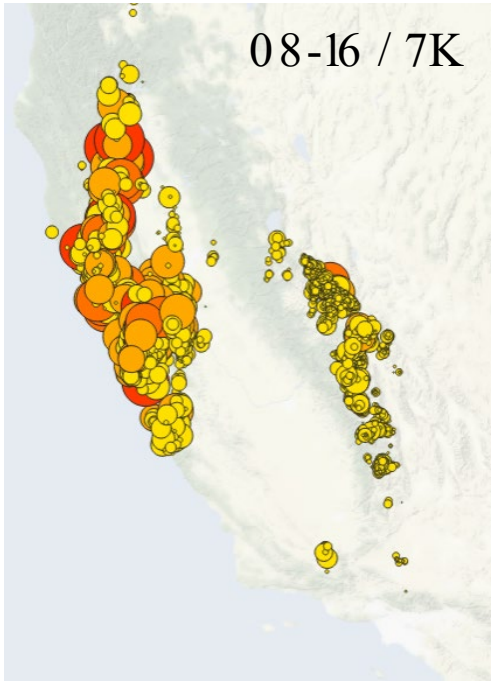
Real-Time updates to GOES weather data processing

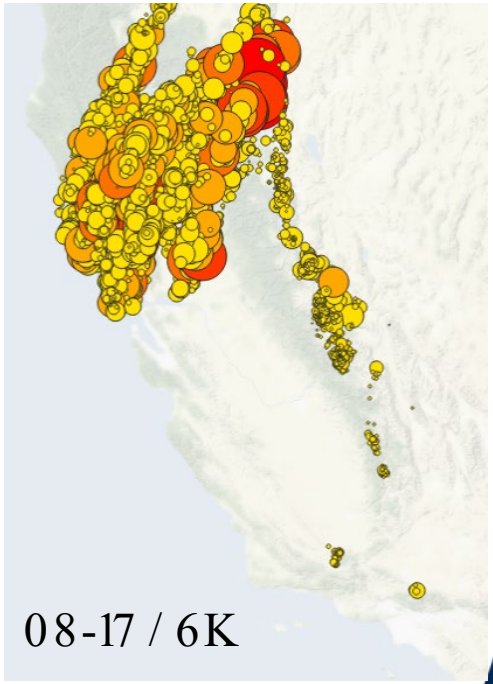
Justin Merz
Research Support Engineer
UC Davis Library
jrmerz@ucdavis.edu

Quinn Hart
Digital Applications
UC Davis Library
qjhart@ucdavis.edu

GLM Lightning - 2020-08 Siege

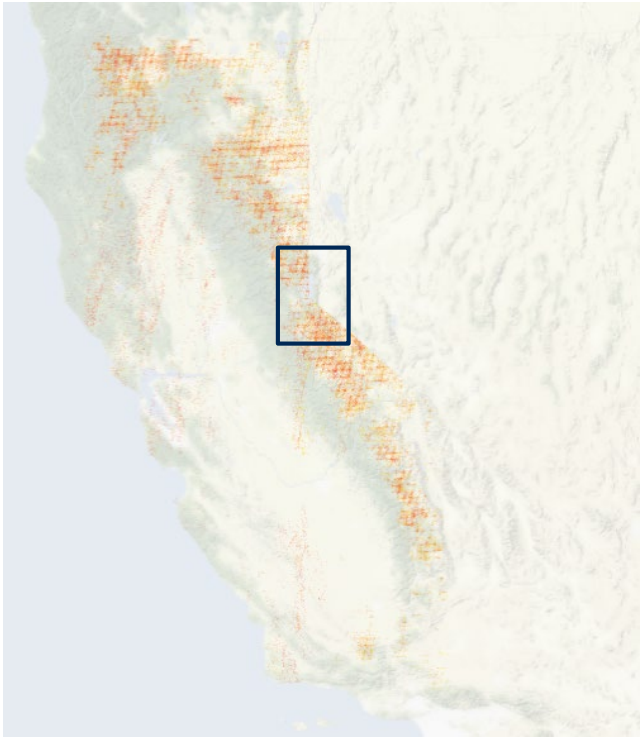


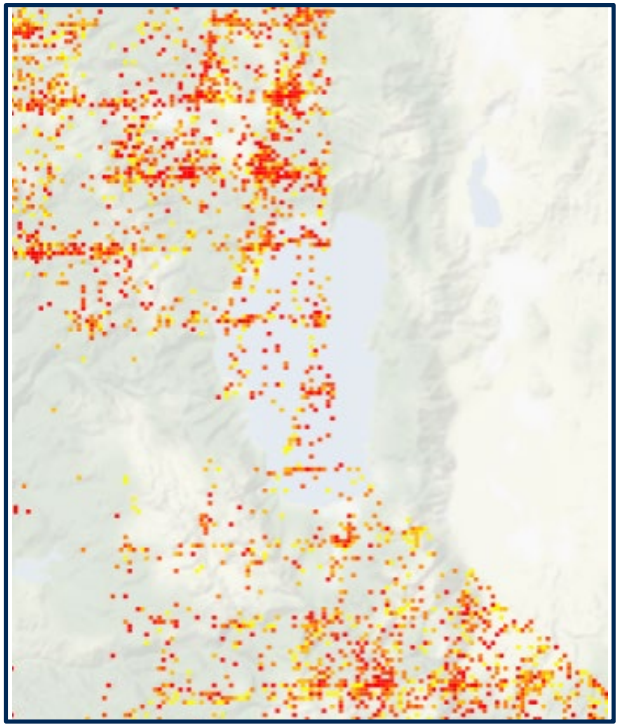




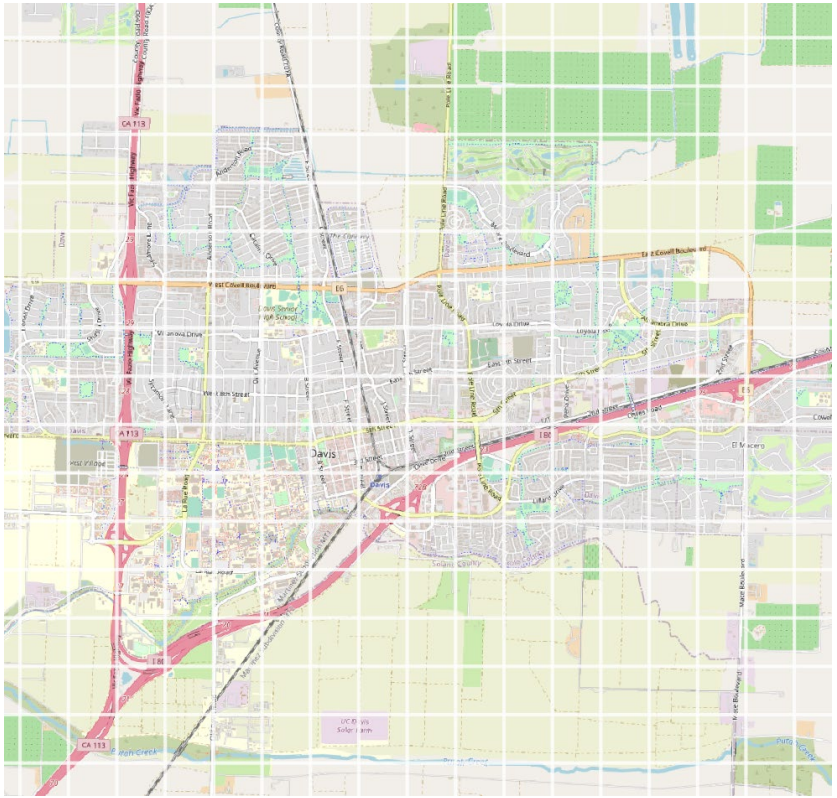
UCDAVIS LIBRARY

GLM Lightning - Summaries





UCDAVIS LIBRARY



UC DAVIS LIBRARY

GOES-R + UC Davis Library

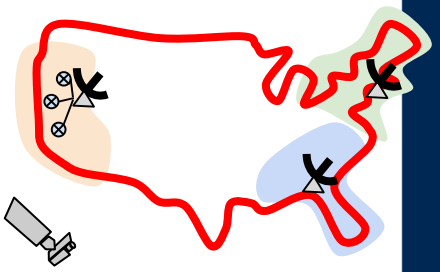
Why this project/weather data?

- Partnership: DWR and UC Davis
  - Existing GOES Receiving Station
- In-house domain knowledge
  - Potential evapotranspiration
- Data for multiple UC Davis departments / projects
  - Share via library

UC DAVIS LIBRARY

CALIFORNIA DEPARTMENT OF WATER RESOURCES

CaSITA- GOES-R and Library



- DWR and UC Davis
  - Existing GOES Station
- In-house domain knowledge
  - Evapotranspiration
- Data for multiple projects
  - Library managed

- Exercise New technologies
- Scalable Cloud Compute
  - 100% containerized
  - Kubernetes (GKE) + Docker
- Realtime data access for campus researchers

UC DAVIS LIBRARY