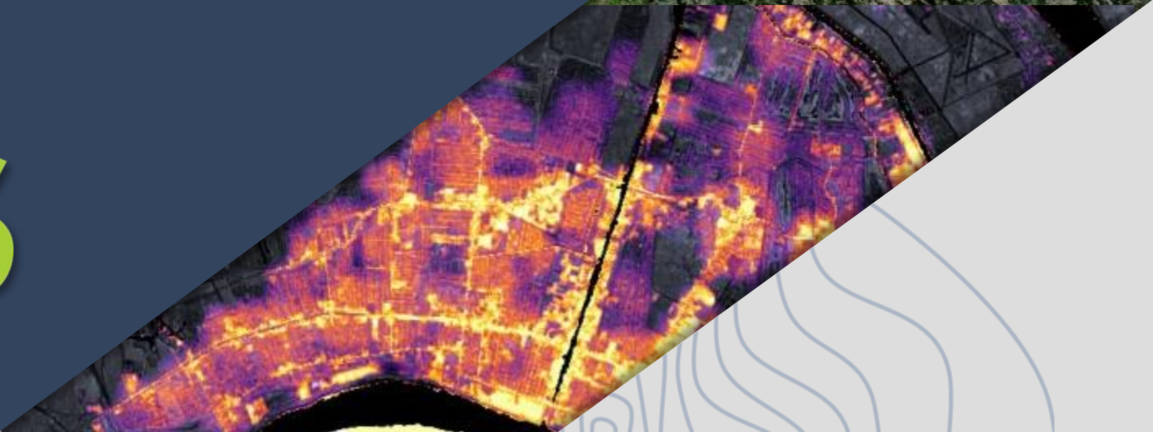
USGS AI/ML Use Cases



Tom Sohre, USGS EROS

Agenda ID: 2022.10.06_12.00

WGISS-54

Tokyo, Japan (JAXA)

3-7 October 2022

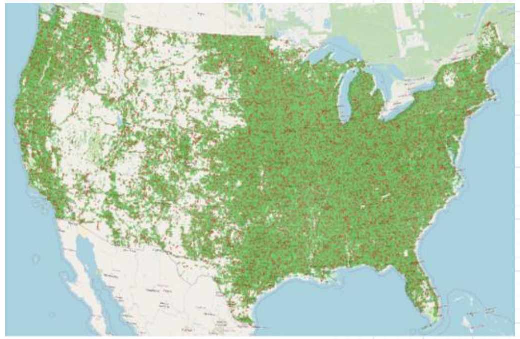
CEOS
Committee on Earth Observation Satellites

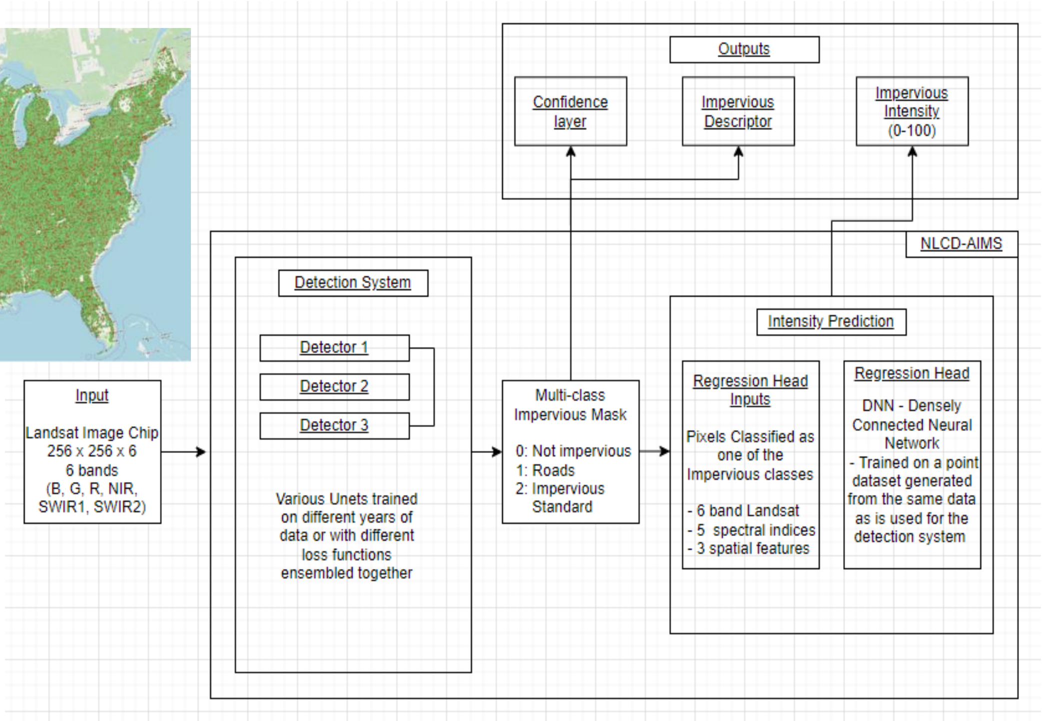
Summary: USGS AI/ML Workflows w/ Earth Observation (EO) Data

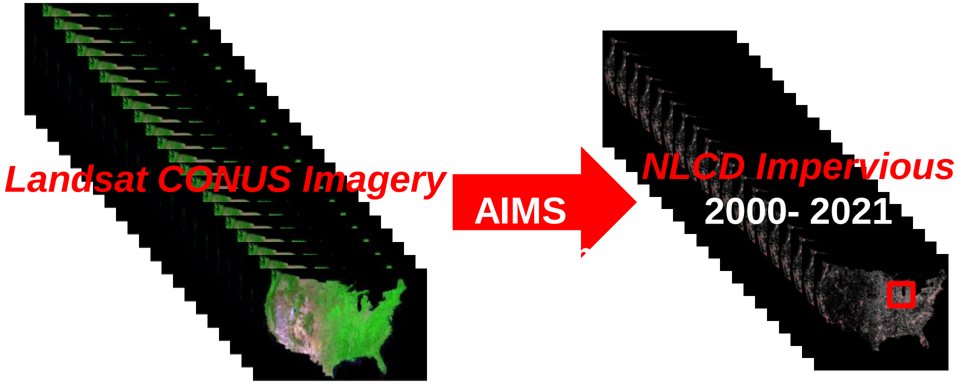
- USGS EROS AI/ML workflows generally make use of similar tools and data for informing our understanding of Earth surface and subsurface conditions, changes, and drivers.
  - Software libraries:
    - STAC, DASK, Xarray, GDAL, Pytorch, & Tensorflow
  - Compute platforms:
    - Amazon Web Services, Microsoft Azure, Google Cloud Platform, & High-performance computers (HPCs)
  - Data:
    - Active and passive remote sensing imagery (e.g., Landsat Collection 2, Harmonized Landsat & Sentinel-2, Sentinel-1 and 3, Planet, MAXAR, IceSAT-2)
- Targets of interest:
  - Land cover conditions (e.g., thematic, fractional cover), land surface phenology, species and lifeform level mapping.

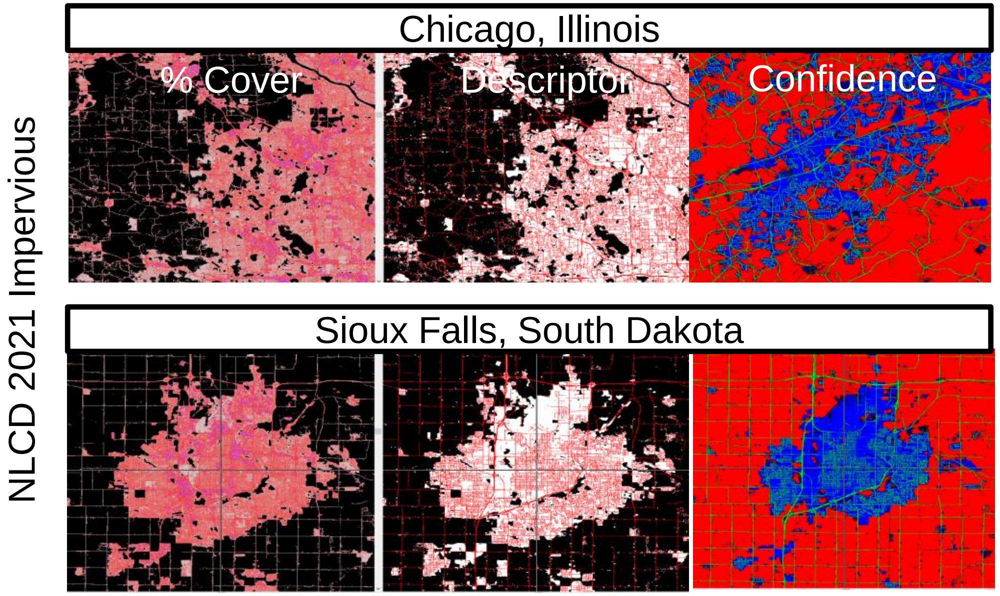
NLCD Impervious Surfaces

Artificial Impervious Mapping System (AIMS)





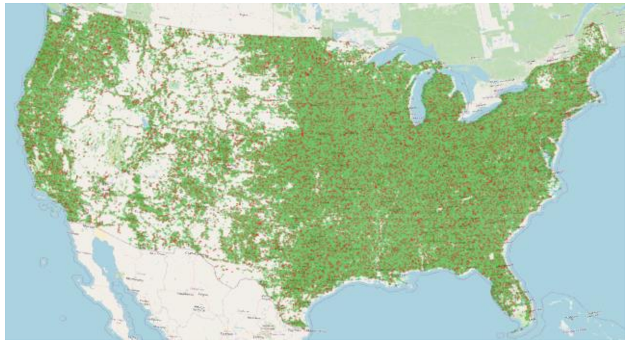




- Key methods and data:
  - **Inputs**: Field and high-resolution observations of land surface conditions and remotely-sensed data (e.g., Landsat).
  - **Compute**: AWS + USGS HPCs.
  - **Models**: Ensemble of U-Nets + DNNs.

Improvements in Impervious Surface mapping for National Land Cover DB 2021

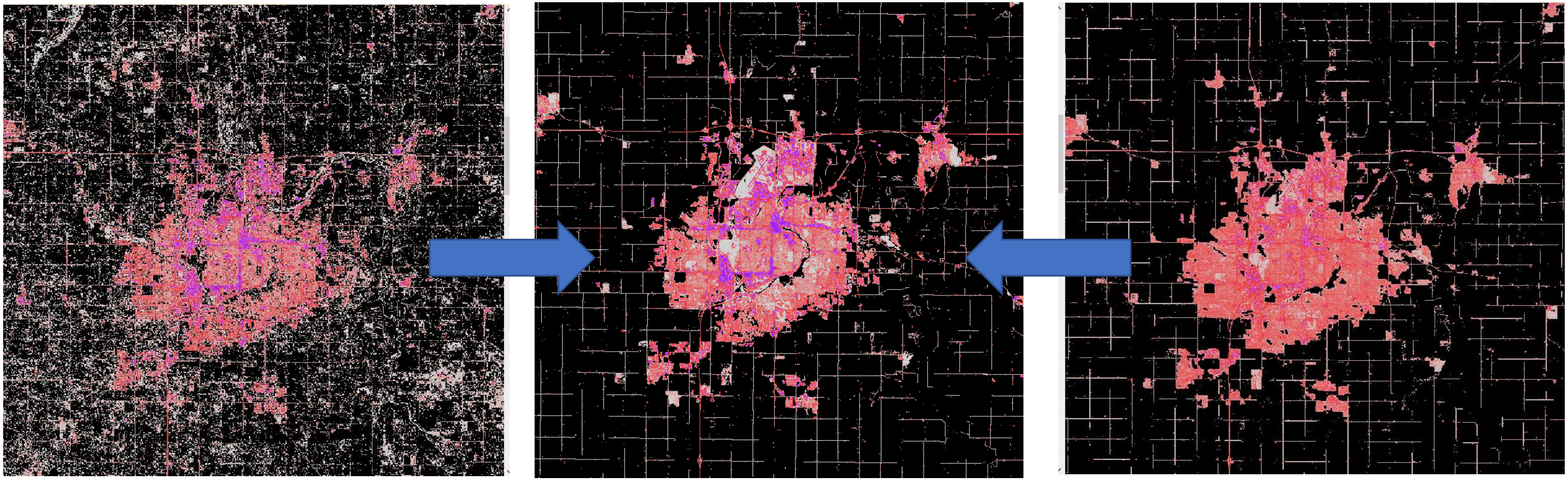
AI Approach Trained on entire CONUS



Cubist Model Outputs
(NLCD 2019)

NLCD Final Product Hand edited
from Cubist

Artificial Intelligence Outputs
(NLCD 2021)

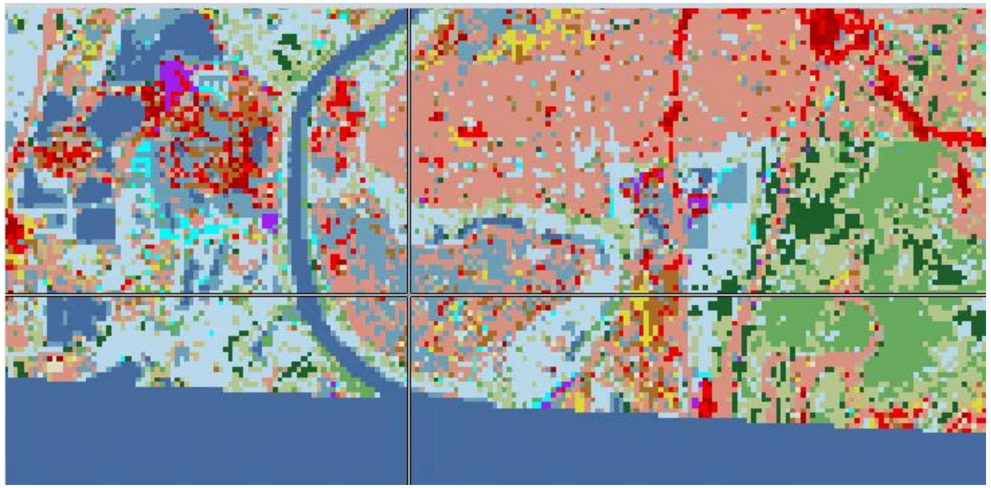


NLCD Land Cover

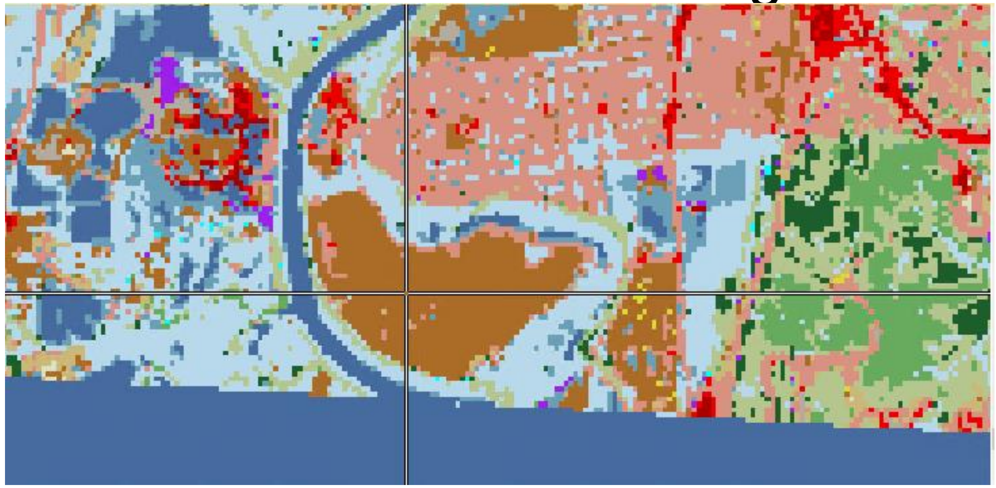
NLCD testing of AI/ML models to land cover classification

Google Earth

See5

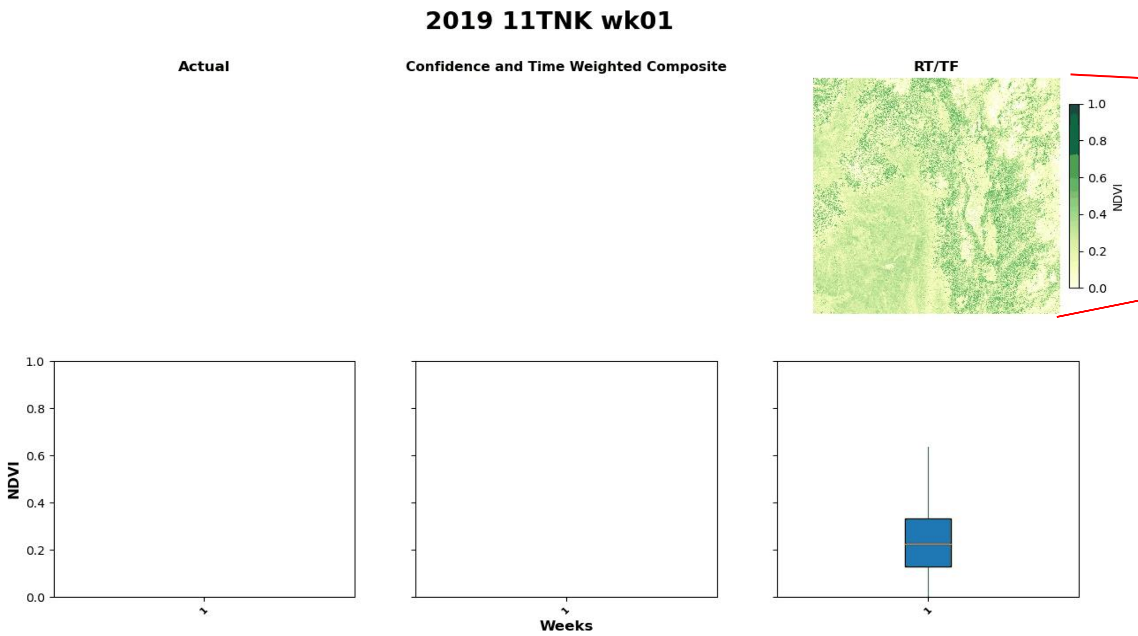


AI Regression

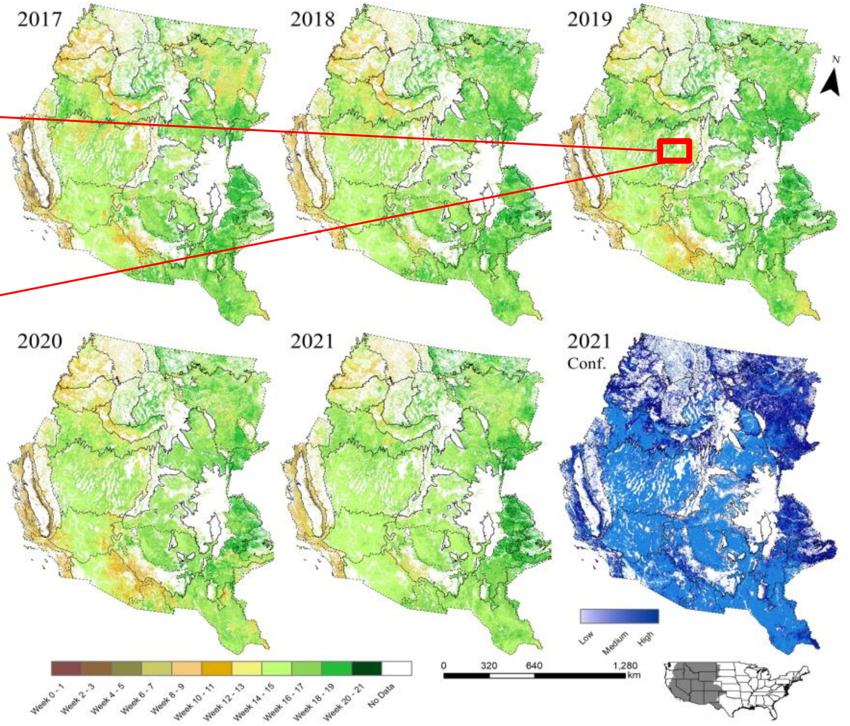


Land Surface Phenology

DL-enabled weekly estimates of spectral data and land surface phenology



DL-enabled weekly estimates of spectral data

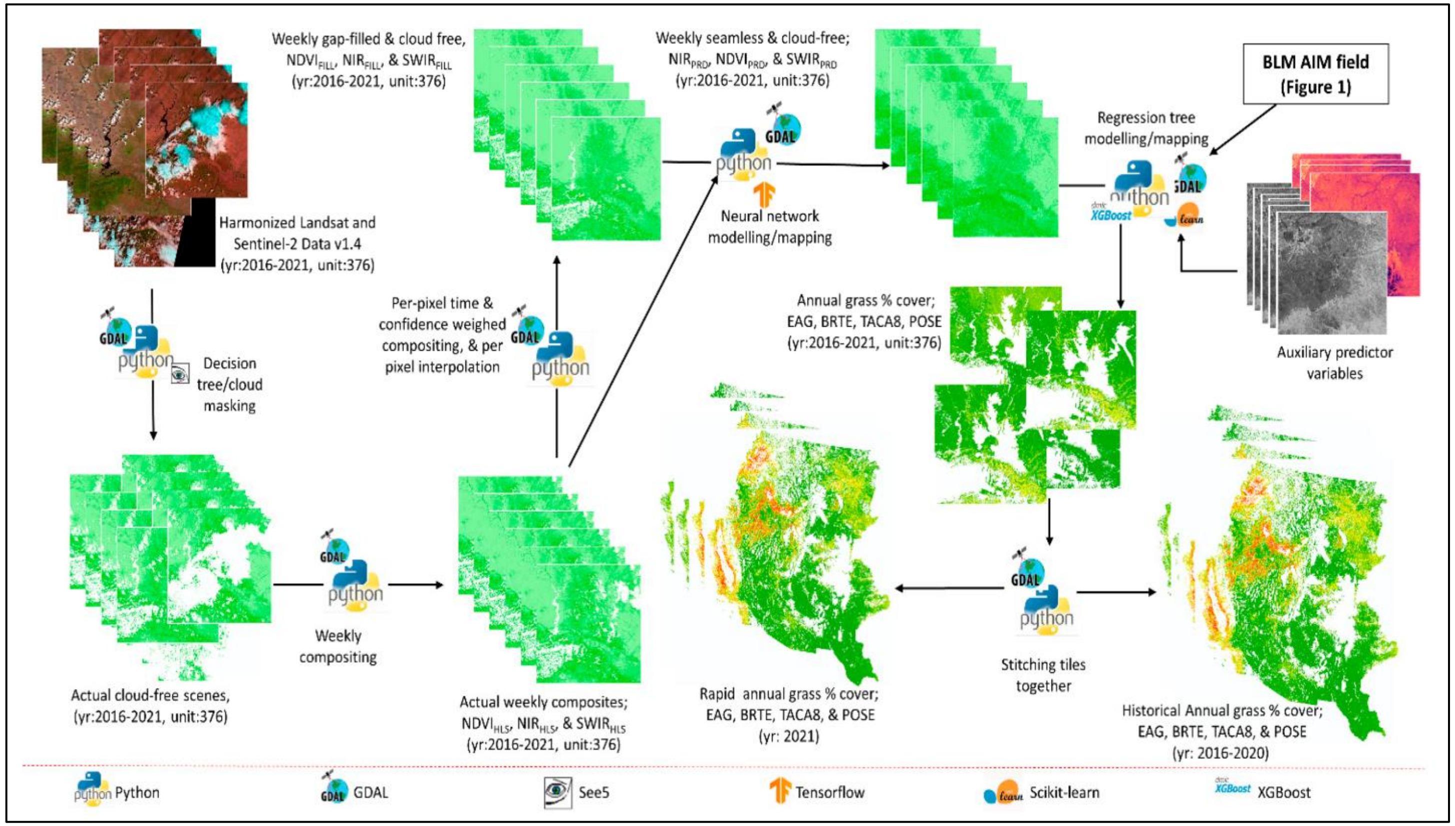


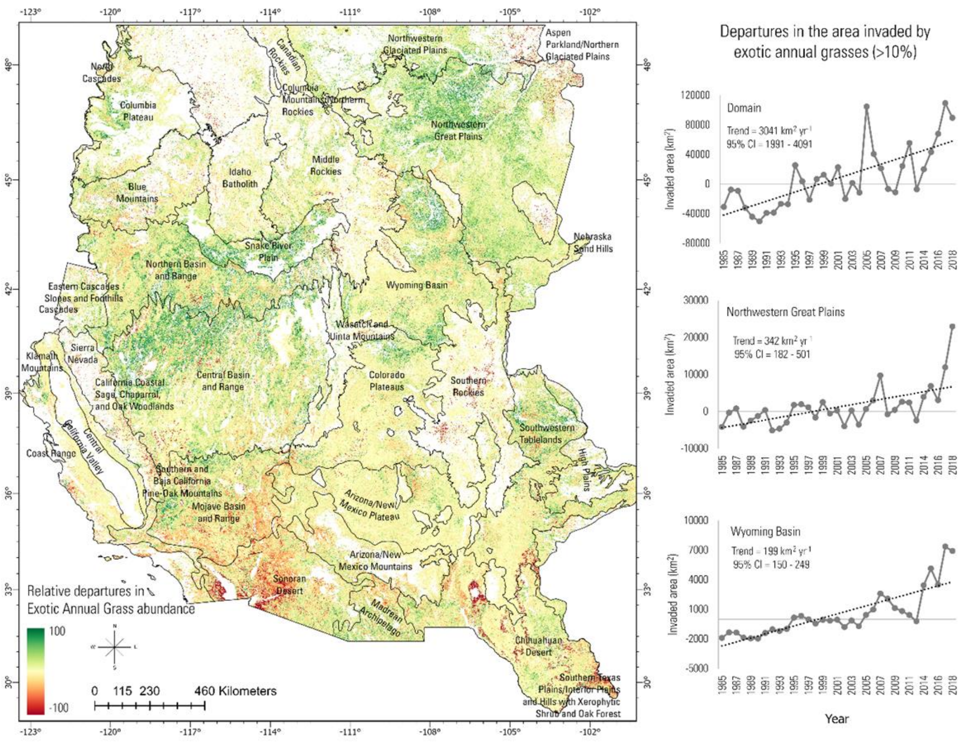
Start of sustained growth of areas dominated by invasive annual grasses

- **Inputs**: Remotely sensed observations (i.e., Harmonized Landsat and Sentinel-2, invasive annual grass cover) and interpretations of major plant life-cycle events (e.g., start of sustained growth).
- **Models**: Multi-task learning w/ deep neural networks (DNNs) & Xgboost.
- **Compute:** USGS High-Performance Computers (i.e., Denali, Tallgrass).
- **Outputs**: Seamless composites of spectral data → land surface phenology.

Invasive Annual Grasses

Rangeland Exotic Plant Monitoring System: Rapid & historical estimates



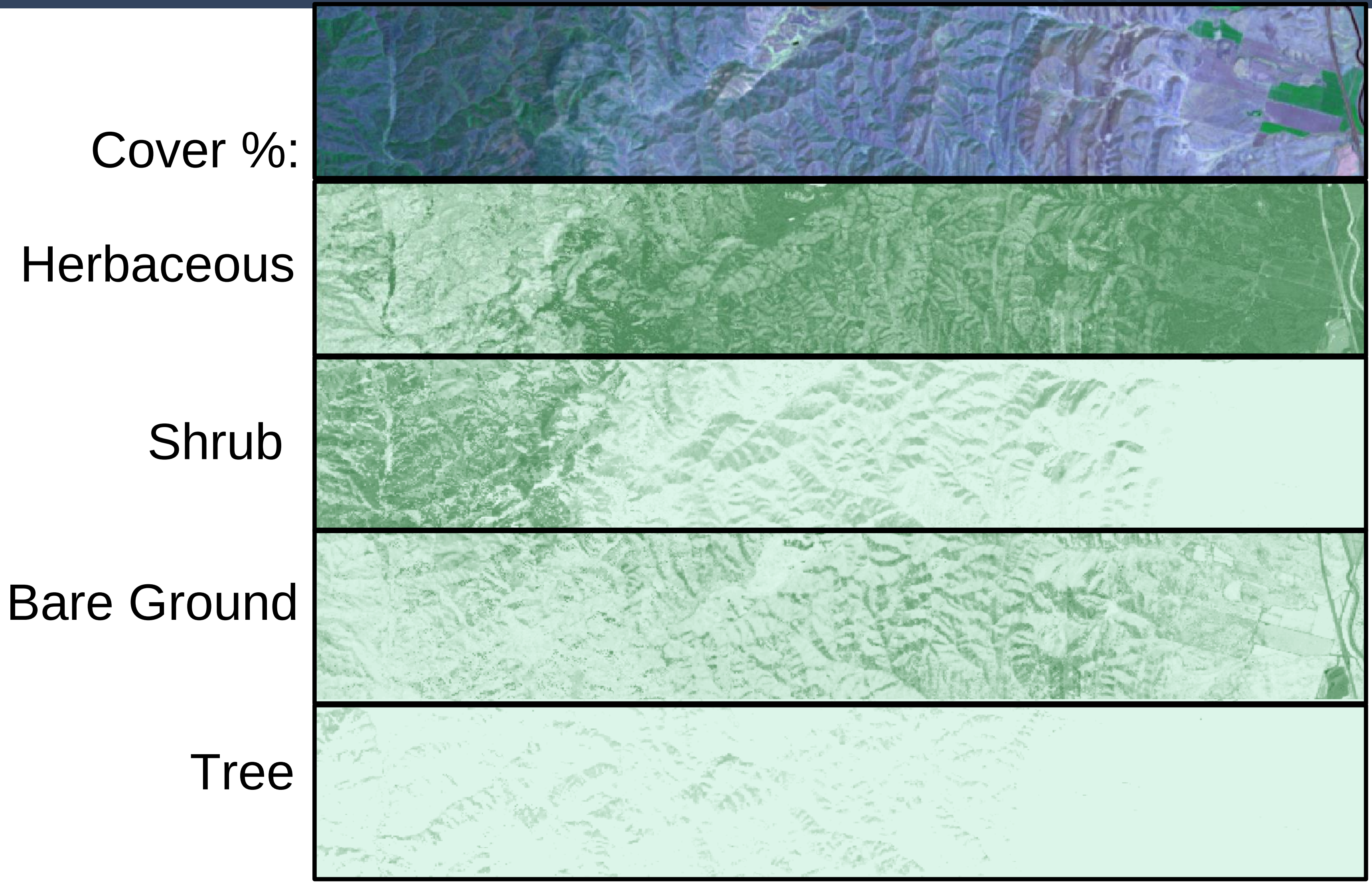


Departures and trends in invasive annual grass (%) cover

- **Inputs:** Field and high-resolution observations of invasive annual grass (%) cover and remotely-sensed and derived data (e.g., Harmonized Landsat and Sentinel-2, Rangeland Analysis Platform + Rangeland Condition Monitoring Assessment and Projection).
- **Compute:** USGS High-performance computers (HPCs) + Amazon Web Services (AWS) + Google Earth Engine (GEE).
- **Models:** Ensembles of XGBoost models + DNNs + LSTMs + RNNs.

RCMAP Fractional Components

- Neural Network model used to predict the cover of several rangeland components simultaneously
- Test accuracy demonstrated a ~9% increase in accuracy moving from regression trees to neural networks

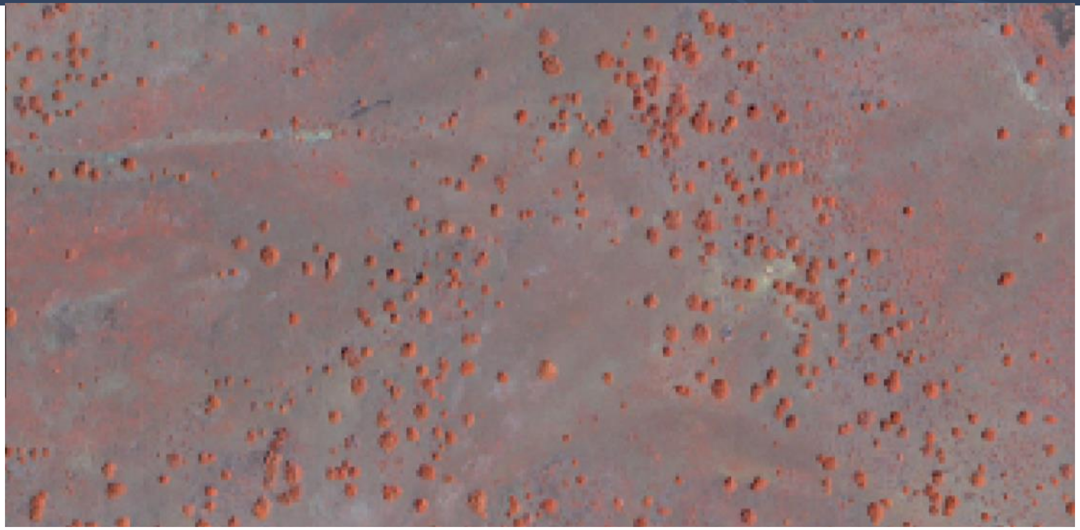


\+ Sagebrush, Annual Herbaceous, and Litter

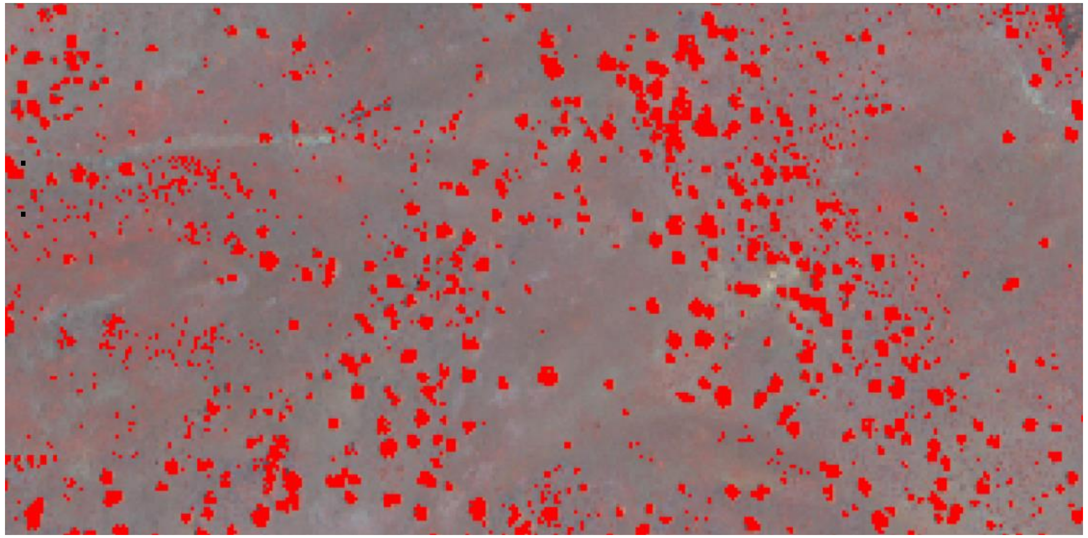
*RCMAP (Rangeland Condition Monitoring and Projection)

High-Resolution UNET Tree Cover

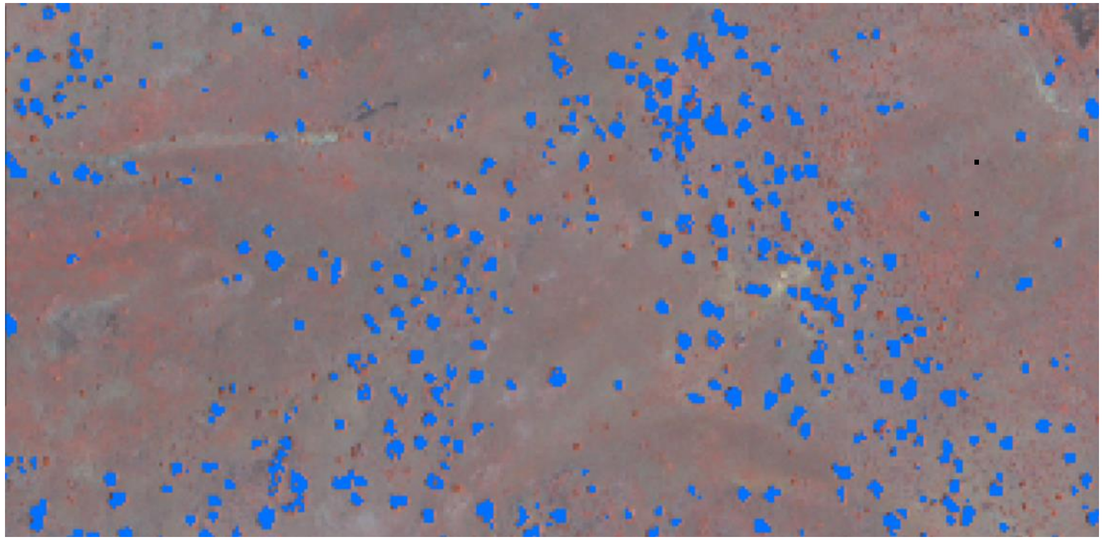
Application of Convolutional Neural Networks (UNET) to WorldView 2-m imagery (right) compared to traditional classification using unsupervised classification.



Unsupervised Classification

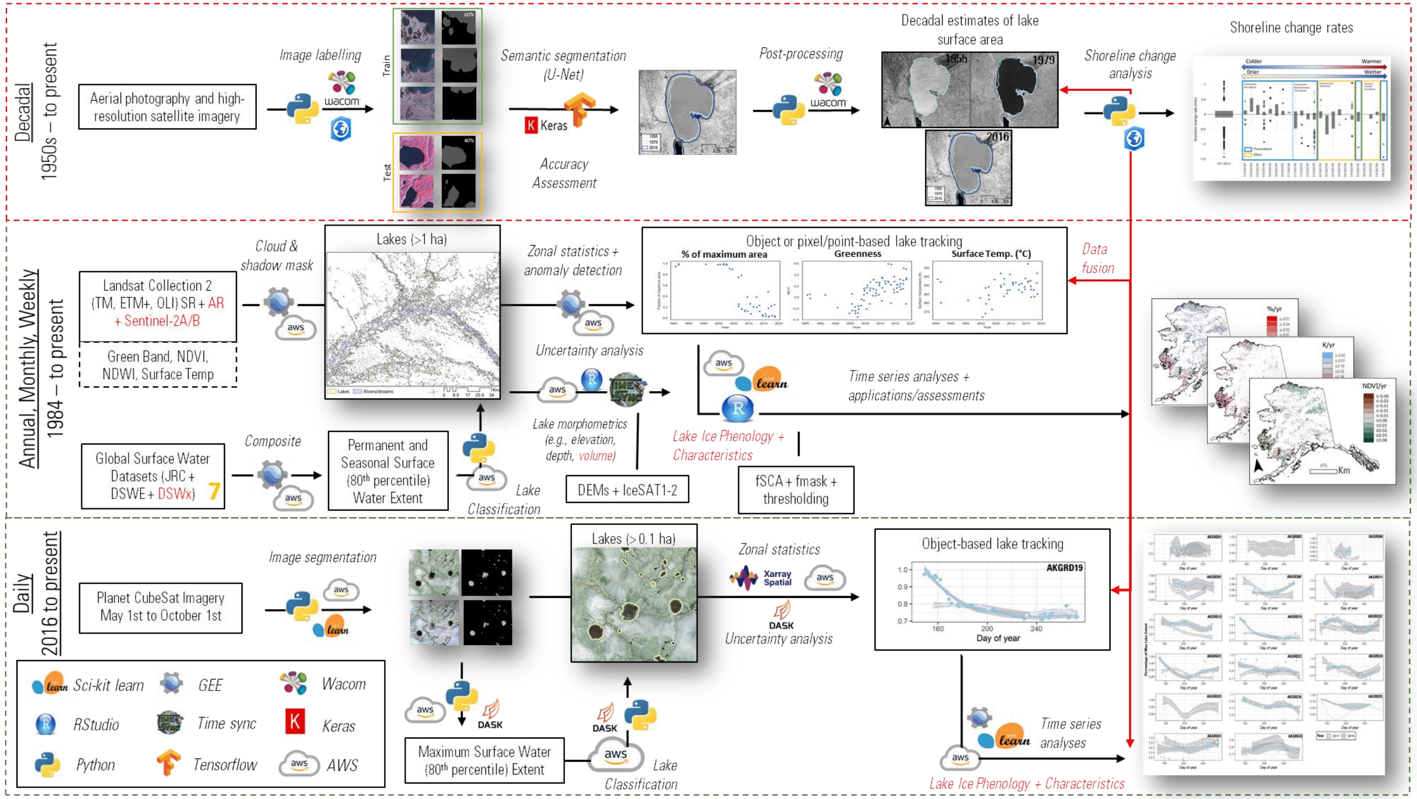


UNET

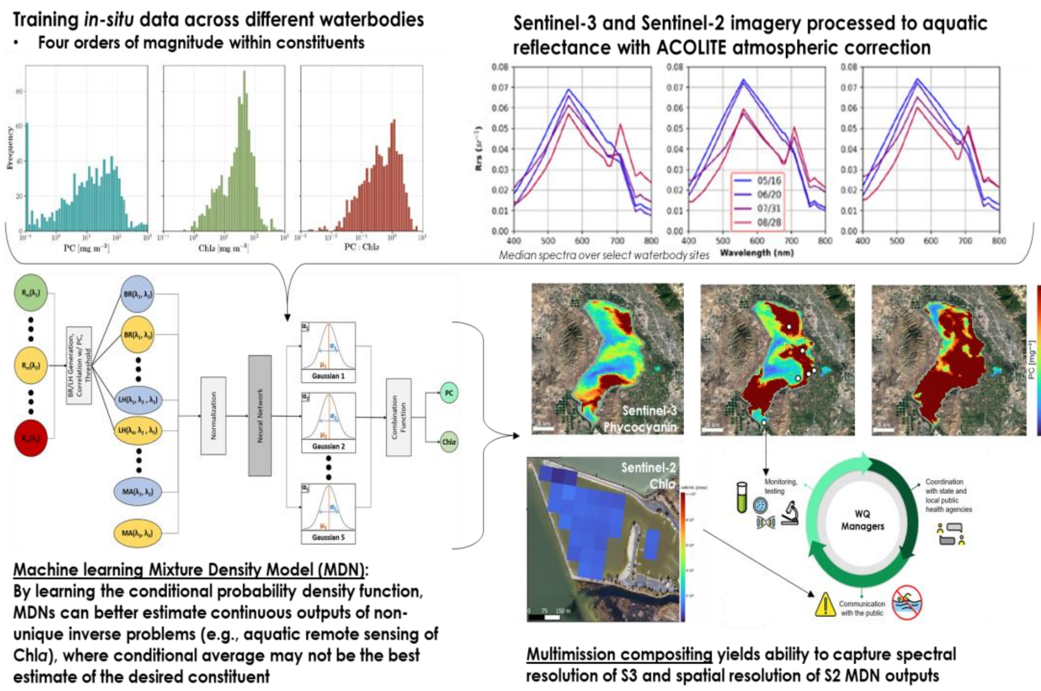


Surface Water Conditions

Cloud and DL-enabled lake monitoring



Lake monitoring system for tracking surface conditions (i.e., area, color, volume, surface temperature)



Sentinel-2 and 3 chlorophyll-a (Chl*a*) & phycocyanin (Pc) retrievals via MDNs trained on field obs.

- **Inputs**: Field and high-resolution observations of surface water conditions and remotely-sensed data (e.g., Sentinel 2-3, Landsat, Planet, Maxar, IceSAT-2).
- **Compute**: AWS + GEE + USGS HPCs
- **Models**: Autoencoders + U-Nets & Mixture Density Networks (MDNs)

Scalable AI/ML Workflows w/ EO Data

Challenges & Opportunities:

- Filling gaps between software libraries commonly used in geoscience data analysis (e.g., Xarray, Dask) and libraries commonly used for deep learning (e.g., TensorFlow, PyTorch).

- Lower barriers to entry + improve modelling tasks
  - Building catalogs/libraries of deep learning models, EO training/testing data, and reproducible workflows (e.g., Jupyter Notebooks)

- Questions: Neal Pastick (njpastick@usgs.gov)