Sage Computing Continuum

SAGE
sagecontinuum.org

A Software-Defined Sensor Network
Cyberinfrastructure for AI@Edge Computing

A Disturbance in the Continuum

AI@Edge and Sage

Pete Beckman: Co-Director Northwestern University / Argonne Inst. for Science and Engineering
Collaborators: Ilkay Altintas, Charlie Catlett, Scott Collis, Nicola Ferrier, Kate Keahey, Eugene Kelly, Jim Olds, Mike Papka, Dan Reed, Raj Sankaran, Sean Shahkarami, Joe Swantek, Valerie Taylor, Doug Toomey, Frank Vernon, Rommel Zulueta, and many more….

Instrument

Data

Analysis

The Computing Continuum

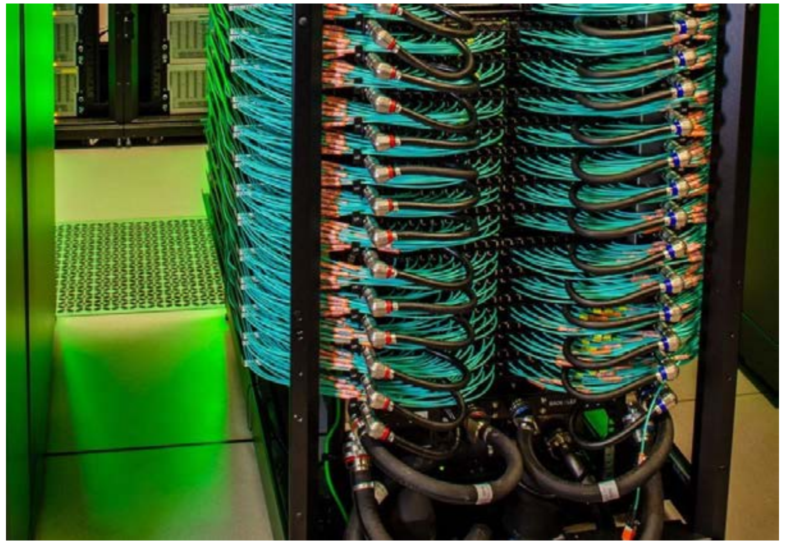
Accelerate scientific discovery: close the loop

Instrument

HPC/Cloud

Edge Computing

Experimental control, New AI



IoT **Facilities**

Analyse full resolution data, find highest value data for the science

HPC Analysis

Problem: Bespoke tools and processes make integration difficult

SAGE sagecontinuum.org

Sage Computing Continuum

2013

SAGE

Cyberinfrastructure for

AI at the Edge

sagecontinuum.org

AI@Edge Summer 2022

(Student Outing: June 2022)

Leadership Team


Pete Beckman
(NU: Director)


Nicola Ferrier
(UC: Deputy Dir.)


Scott Collis
(NU: Instruments, Atmos)


Valerie Taylor
UC: Edu, Broader Impacts)


Eugene Kelly
(CSU; Ecosys, NEON)


Mike Papka
NIU: Edu, Broader Impacts


Raj Sankaran
NU: Node Arch


Ilkay Altintas
(SDSC: Data)


Charlie Catlett
(UIllinois: Urban)


Jim Olds
(GMU; Life Sci, Risk)


Dan Reed
(Utah; Architecture)


Kathy Bailey
Proj Mgmt


Helen Taaffe
Proj Mgmt


Joe Swantek
NU: Software


Irene Qualters
(LANL; Advisory Committee Chair)

MSRI-1: 1935984
Start: October 1, 2019

Sage Computing Continuum

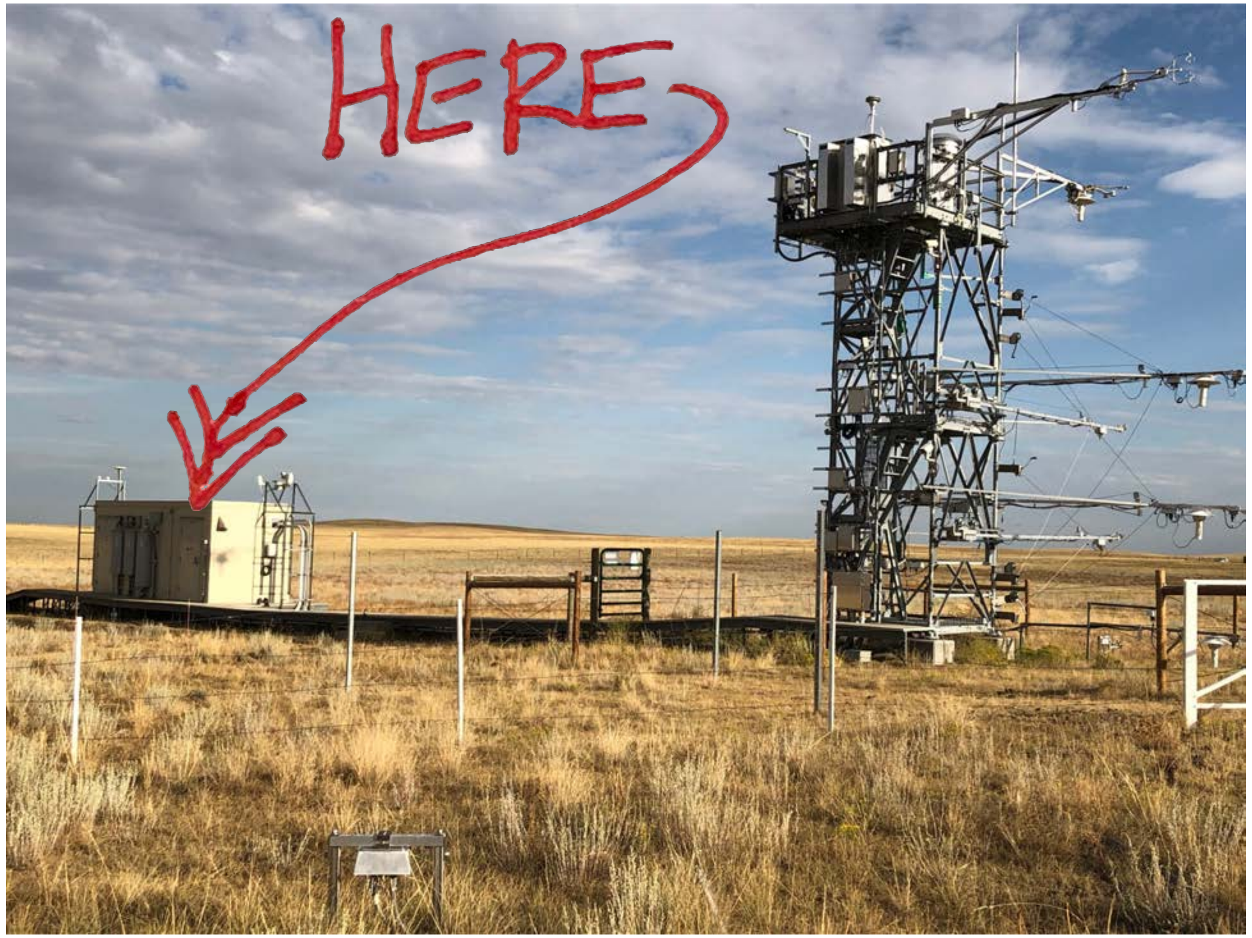
Sage Goals

- New kind of National AI Cyberinfrastructure
    - High-quality, resilient, well-documented software
    - Leverage best Open Source frameworks
        - PyTorch, OpenCV, TensorFlow, Kubernetes, Docker, etc.
- Build community of AI@Edge scientists
    - New AI-based measurements
        - Software-defined sensors
    - New AI algorithms for edge
- Deploy testbed into production facilities
- Provide new capabilities for live data and triggered responses
- Teach and train students, explore new ideas

SAGE
sagecontinuum.org

Building on
NSF Array of Things
(2016-2018)

Put AI@Edge



(Sensors sample at 40hz, aggregate to 30min)

Analyse full resolution data,
find highest value data for
the science

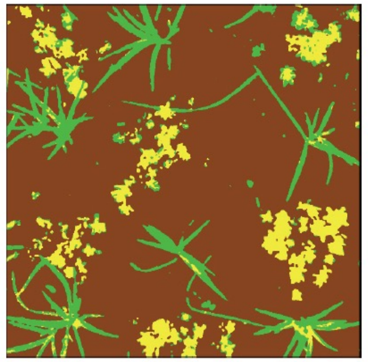
What is a "Software Defined Sensor"?

AI-Based Measurement & Anomaly Detection, & Control

Your software container running here

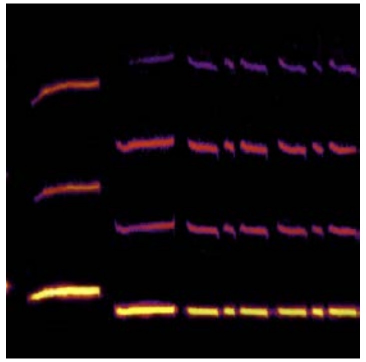
Analysis produces live results

Plant Species

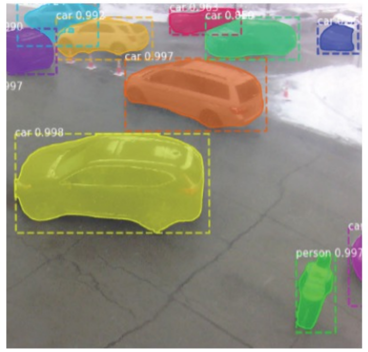


Pedestrian Flow

Birdsong

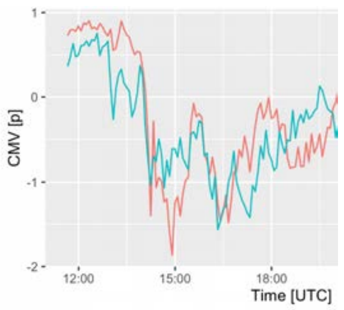


Traffic Flow



Wildlife

Cloud Motion Vectors



Wildfires: detecting smoke

Flooding / surface water

SAGE
sagecontinuum.org

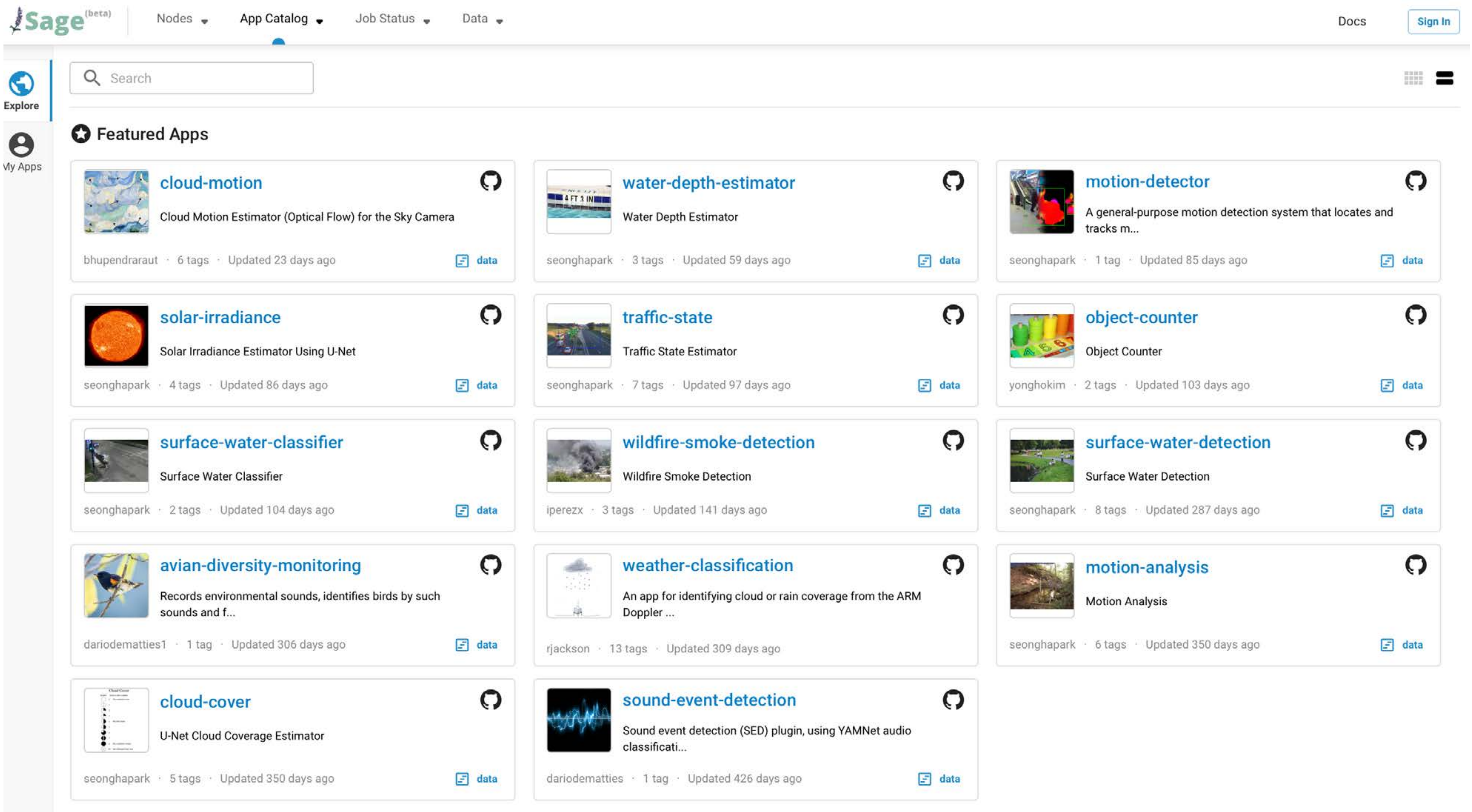
Software Defined Sensors – AI@Edge

AI@Edge	Methodology	Sites
Motion Detector (V)	Background subtraction and Dense Optical Flow [C]	**U**
Motion Analysis (V)	Random Decision Forest [C]	**G**, **N**
Smoke and Fire Detection (I)	SmokeyNet DNN [F]	**N**, **R**
Solar Irradiance (I)	pvlib [D]	**U**, **G**, **O**
Cloud Cover (I)	U-Net [P]	**G**, **H**, **N**, **O**, **T**, **U**
Cloud Motion (V)	Phase Correlation [C]	**G**, **H**, **N**, **O**, **T**, **U**
Weather Classification (I)	Gradient Boosting Tree [X]	**M**
Traffic State (V)	YOLO V7 and Sort [P]	**U**
Object Counter (I)	YOLO V7 [P]	**G**, **H**, **N**, **O**, **T**, **U**
Water Depth Estimation (I)	U-Net [P]	**N**
Surface Water Detection (V)	DeepLab v2 and ResNet 101 [P]	**G**
Surface Water Classification (I)	ResNet50 [P]	**U**
Avian Diversity (A)	BirdNET DNN ResNet [F]	**G**, **H**, **N**, **O**, **T**, **U**
Sound Event (A)	VGG based YAMNet DNN [F]	**G**, **H**, **N**, **O**, **T**, **U**

Framework: [D]=Pandas, [K]=Keras, [C]=OpenCV, [P]=PyTorch, [F]=TensorFlow, [X]=XGBoost
Sites: **M**=ARM, **G**=GLIFWC, **H**=HPWREN, **N**=NEONS, **O**=OHAZ, **R**=Rural, **T**=TNC, **U**=Urban
Data:(I)=Image, (A)=Audio, (V)=Video

A National AI@Edge Resource for the Community

The Edge Code Repository



Sage Computing Continuum

SAGE
sagecontinuum.org

Sage Computing Continuum

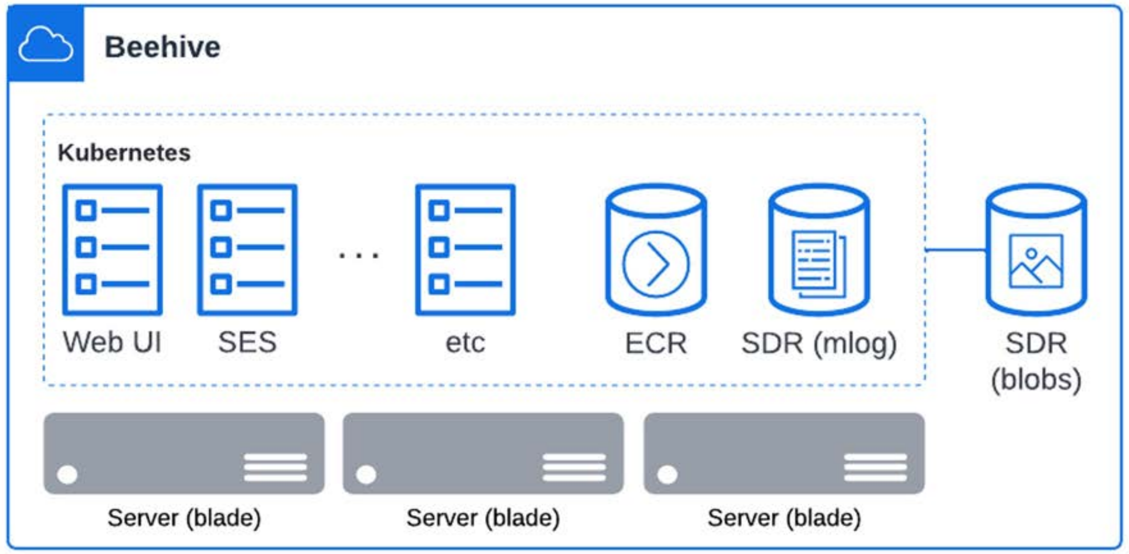
Sage Software Architecture

Sage Nodes

OpenCV docker PYTORCH TensorFlow Lite K3S ROS

- Built upon standard AI Stack
- **Containers** on **Kubernetes**
- **Multi-tenancy**
- **“Goal-based” Scheduler**
- Local control for actuation
- Extreme cybersecurity
- Publish data to Beehive

Cloud Infrastructure



AI@Edge ”Plugin” from Edge Code Repository (ECR) (the “App Store”)

Beehive manages

- Sage Edge Scheduler (SES)
- Sage Data Repository (log entries)
- Sage Data Repository (binary files)
- User Interface components

SAGE sagecontinuum.org

“Digital Twin” <> “HPC Simulation”
“Edge Computing” <> “Physically Moved Computation”

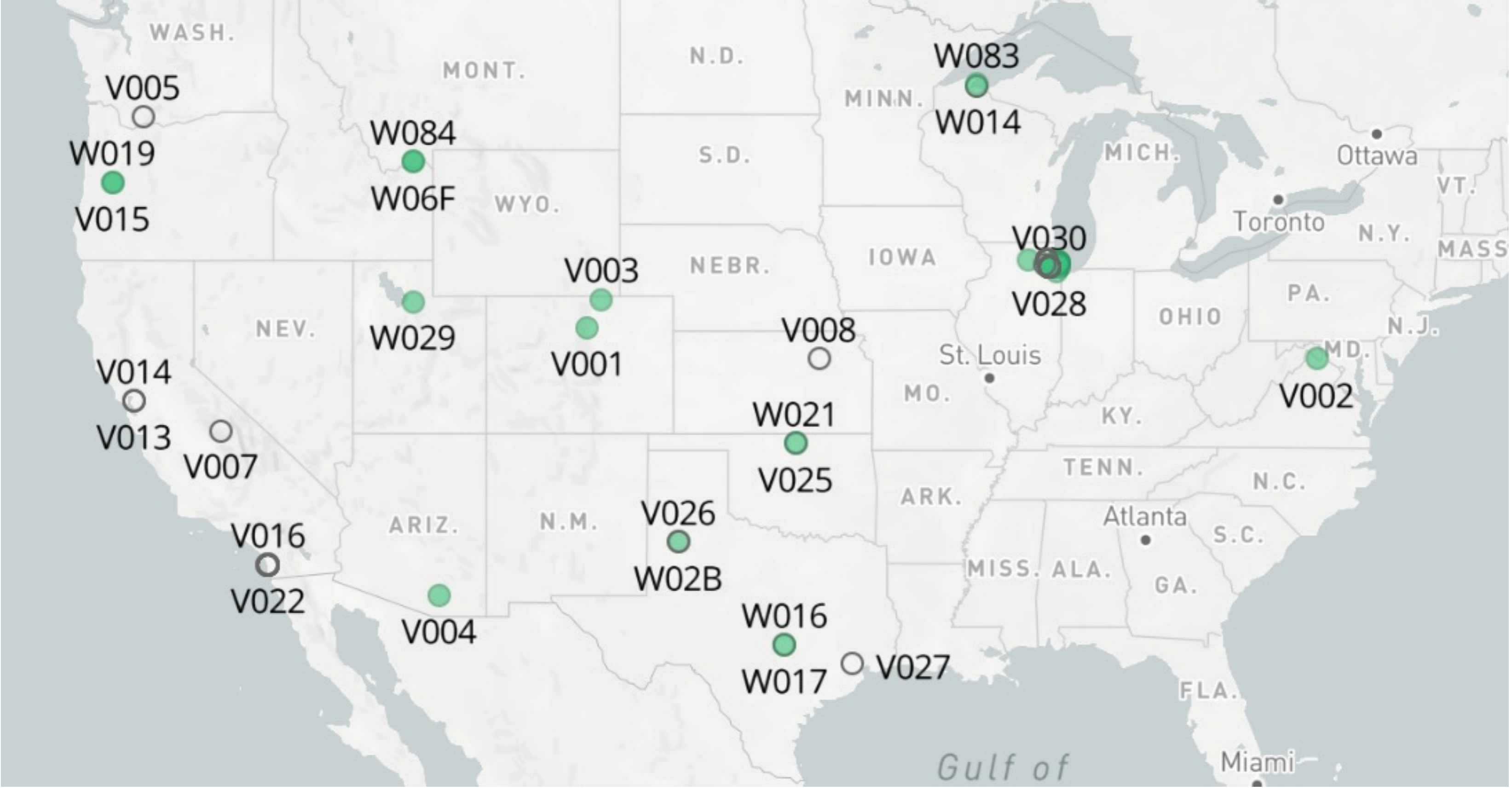
Is Being Edgy Really Different?

- *Extreme cybersecurity*
  - No physical or network security ==
    - no open ports, phone home, apps cannot leak data, full encryption
- *Autonomy*: Operates disconnected for hours, days, weeks
  - Local decisions: computing tasks, experimental control, data preserved
- *Multi-tenancy* managed by ***Science Goal (rule)*** based (not queued job)
  - Multi-objective autonomy
- *Secure Edge Apps*: managed builds -> provenance and policy/cybersecurity

Sage Computing Continuum

SAGE
sagecontinuum.org

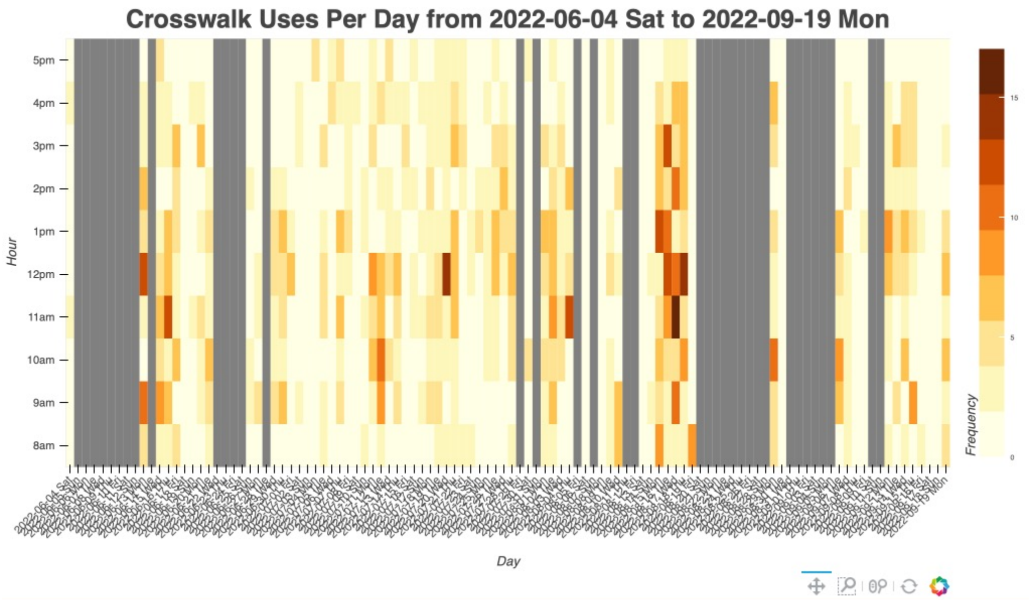
Sage Computing Continuum

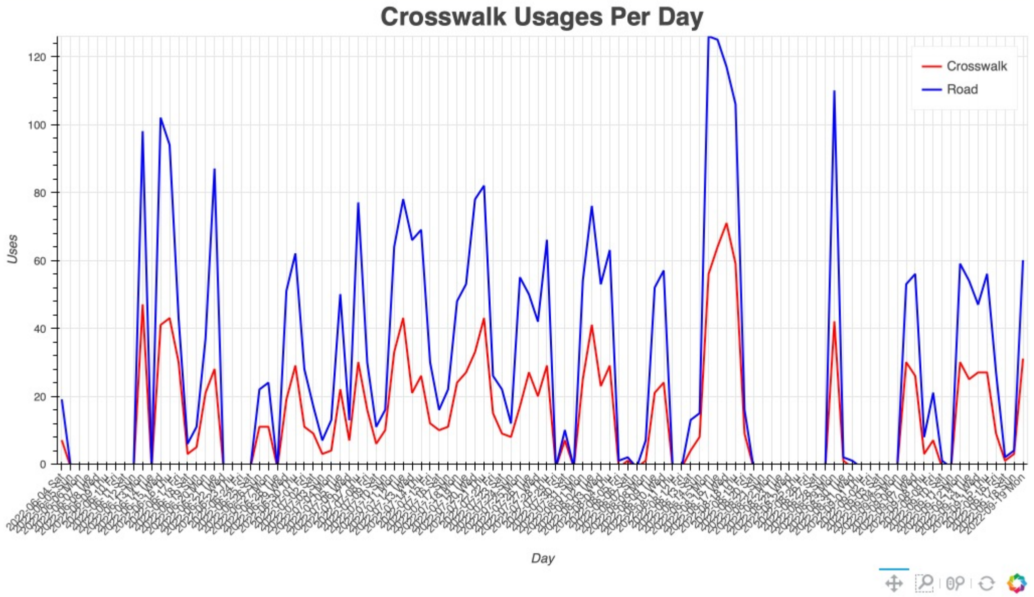


Undergraduate Research: Pedestrian Detection and Paths

http://snick.cs.niu.edu

Pedestrian data processed to understand patterns and transformed for top-down view then bundled to highlight patterns





NIU experimental node with wired network connection

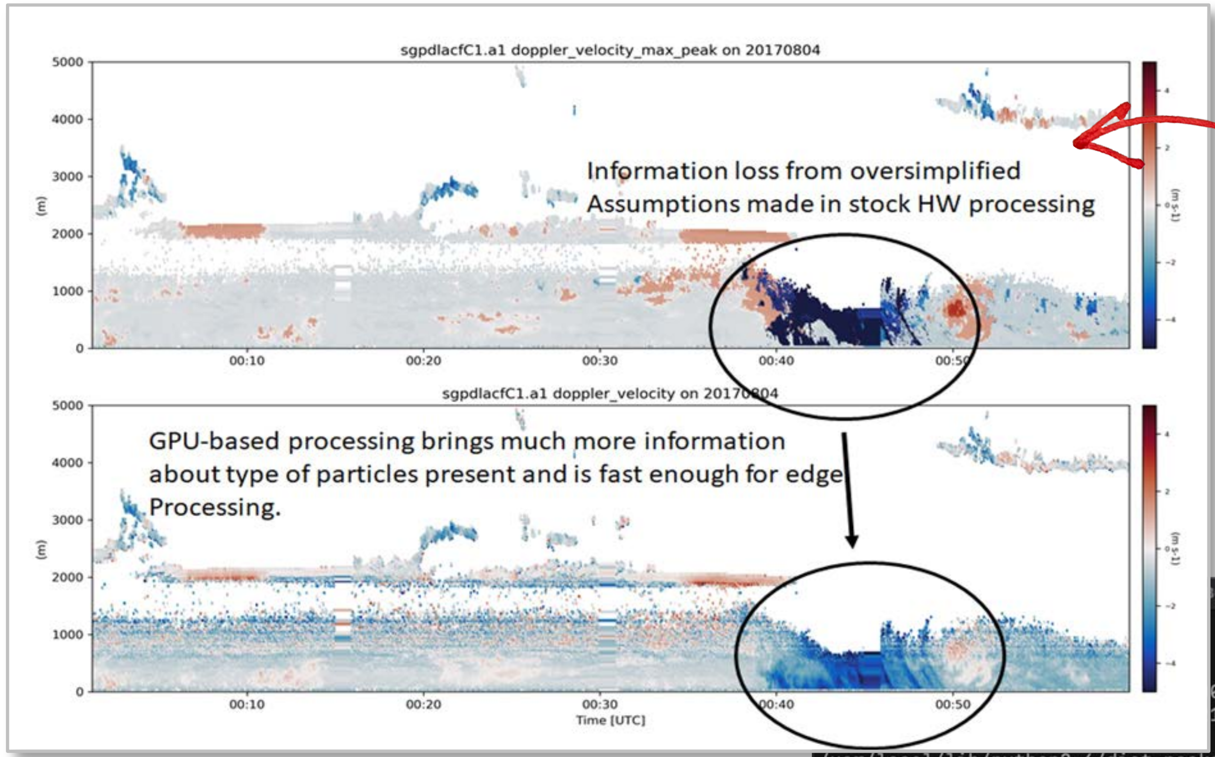
- Experiment with sampling rate and resolution

YOLO based model for identifying people and to check for use of crosswalk

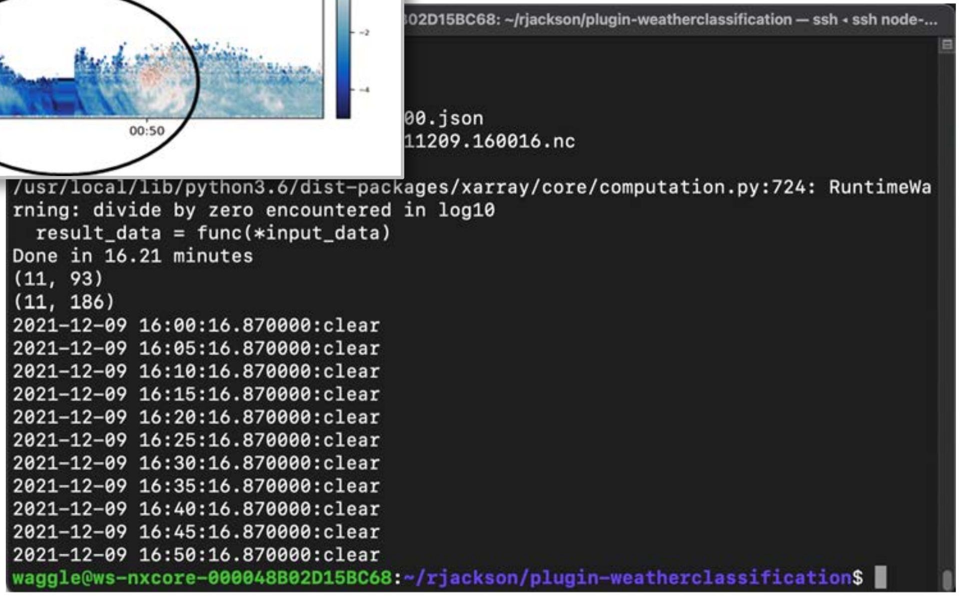
Justin Derus, **Wesley Kwiecinski**, Pratool Bharti, Michael Papka: Northern Illinois University

Sage Computing Continuum

ML-guided Doppler LIDAR Processing



Analyse full resolution data, find highest value data for the science

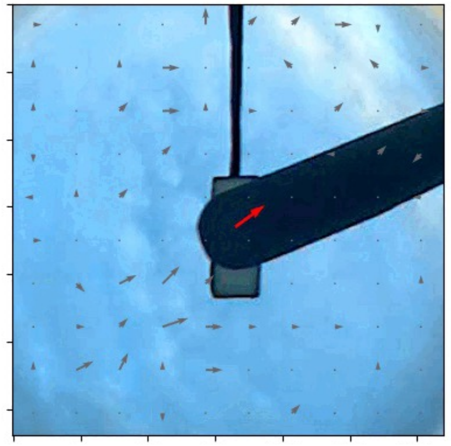


- Waggle node installed at the ARM SGP site
- Algorithm developed using XGBoost.
- Automated ML doppler LIDAR spectra classifier
- Next step: autonomous selection and operation of LIDAR scanning modes

SAGE sagecontinuum.org

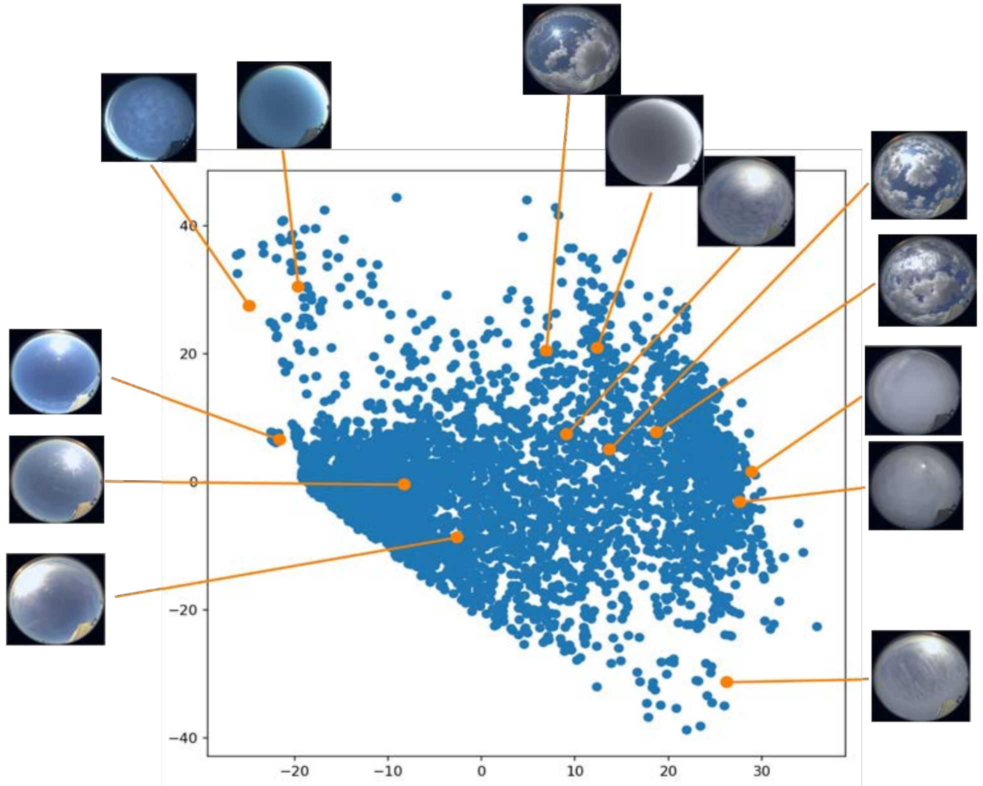
Research Credit: Bobby Jackson: rjackson@anl.gov & Scott Collis: scollis@anl.gov

Edge computing unleashed on understanding climate

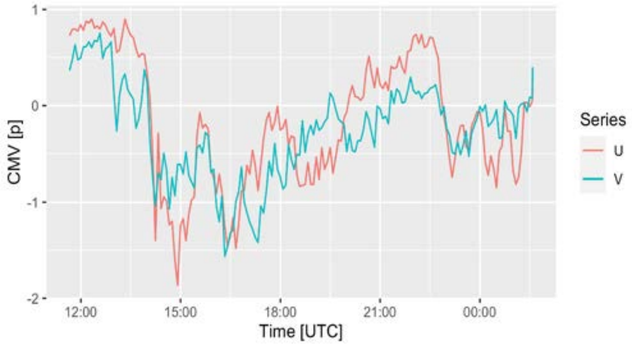


Analyzing cloud patterns

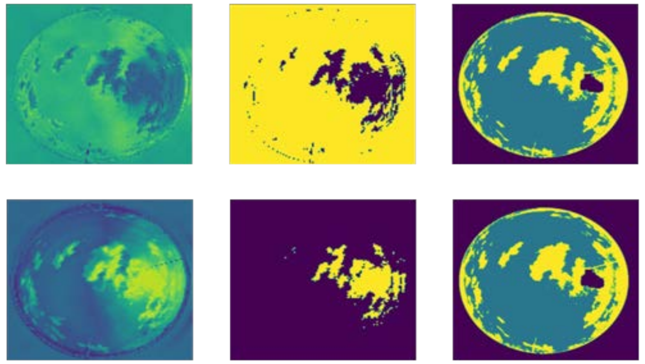
Self-learning Cloud Patterns



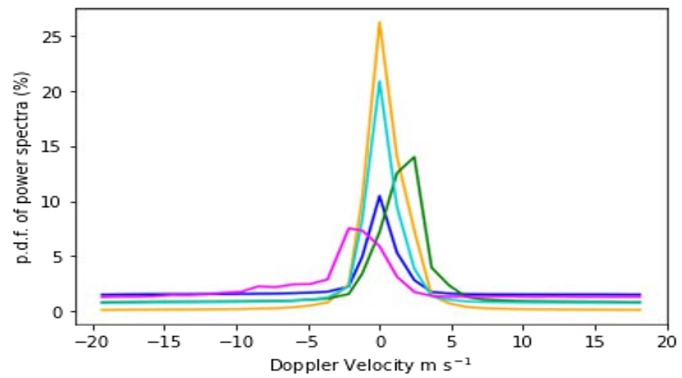
Cloud Motion Vector

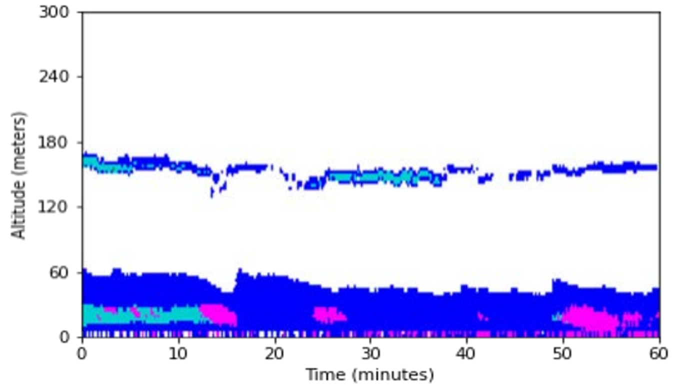


Cloud Cover Estimation

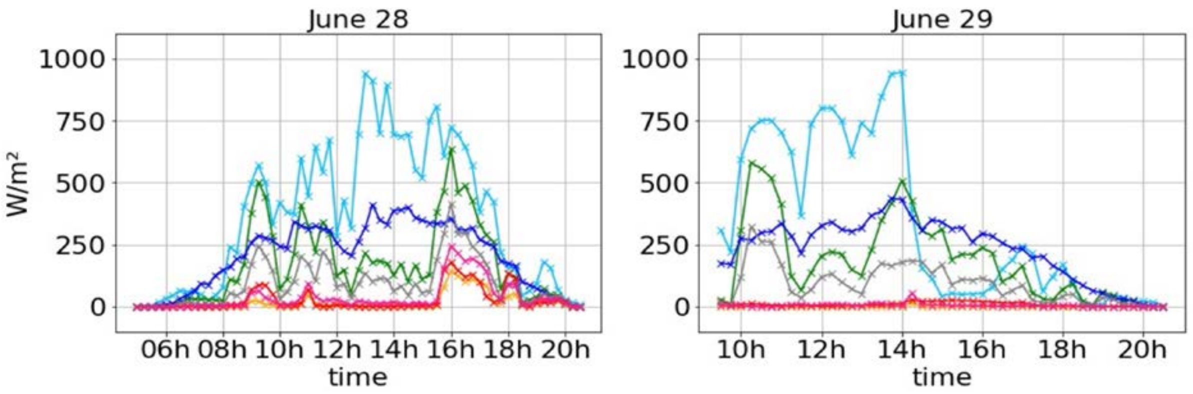


Automated LIDAR particle identification



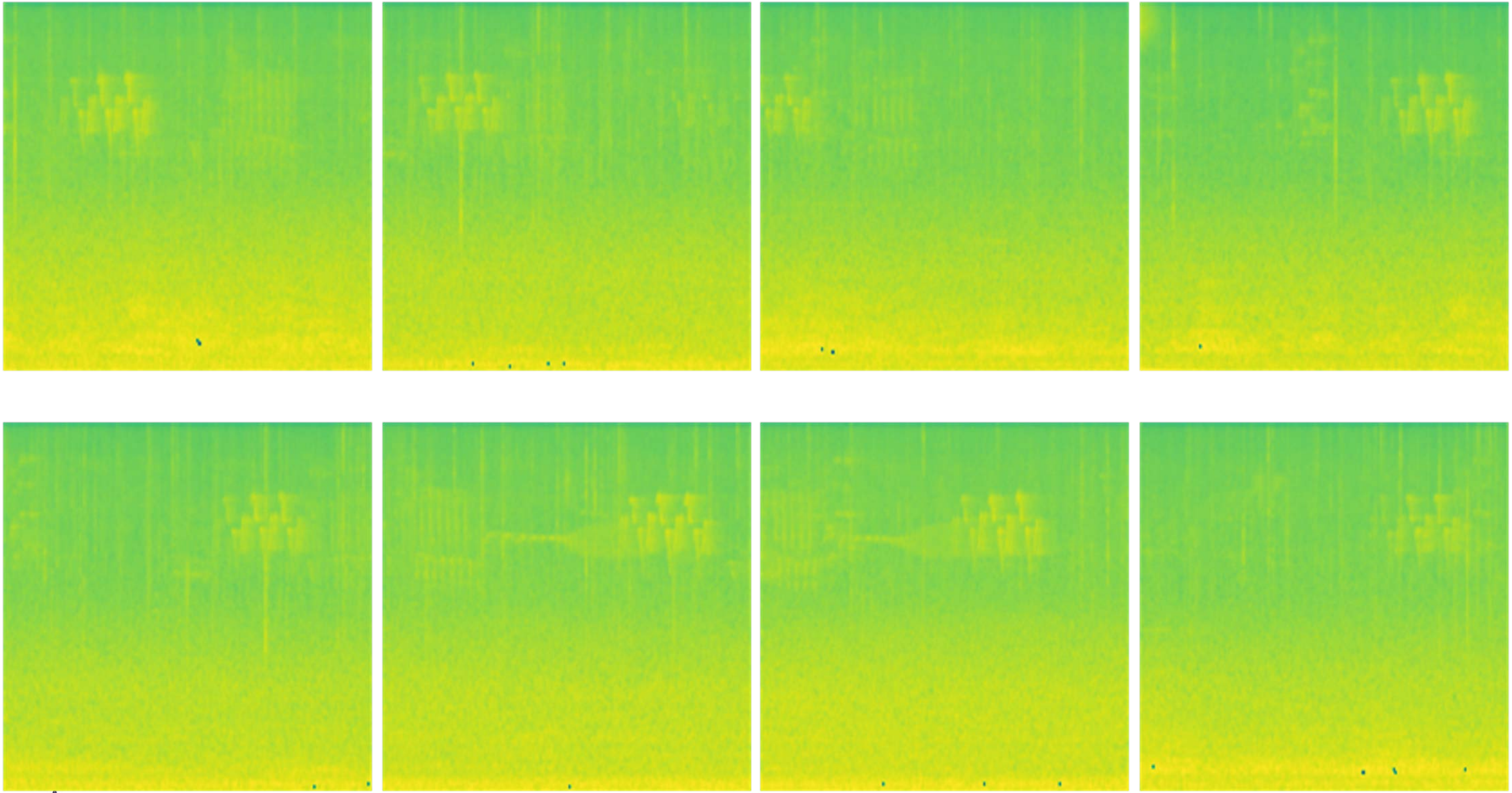


Solar energy estimation

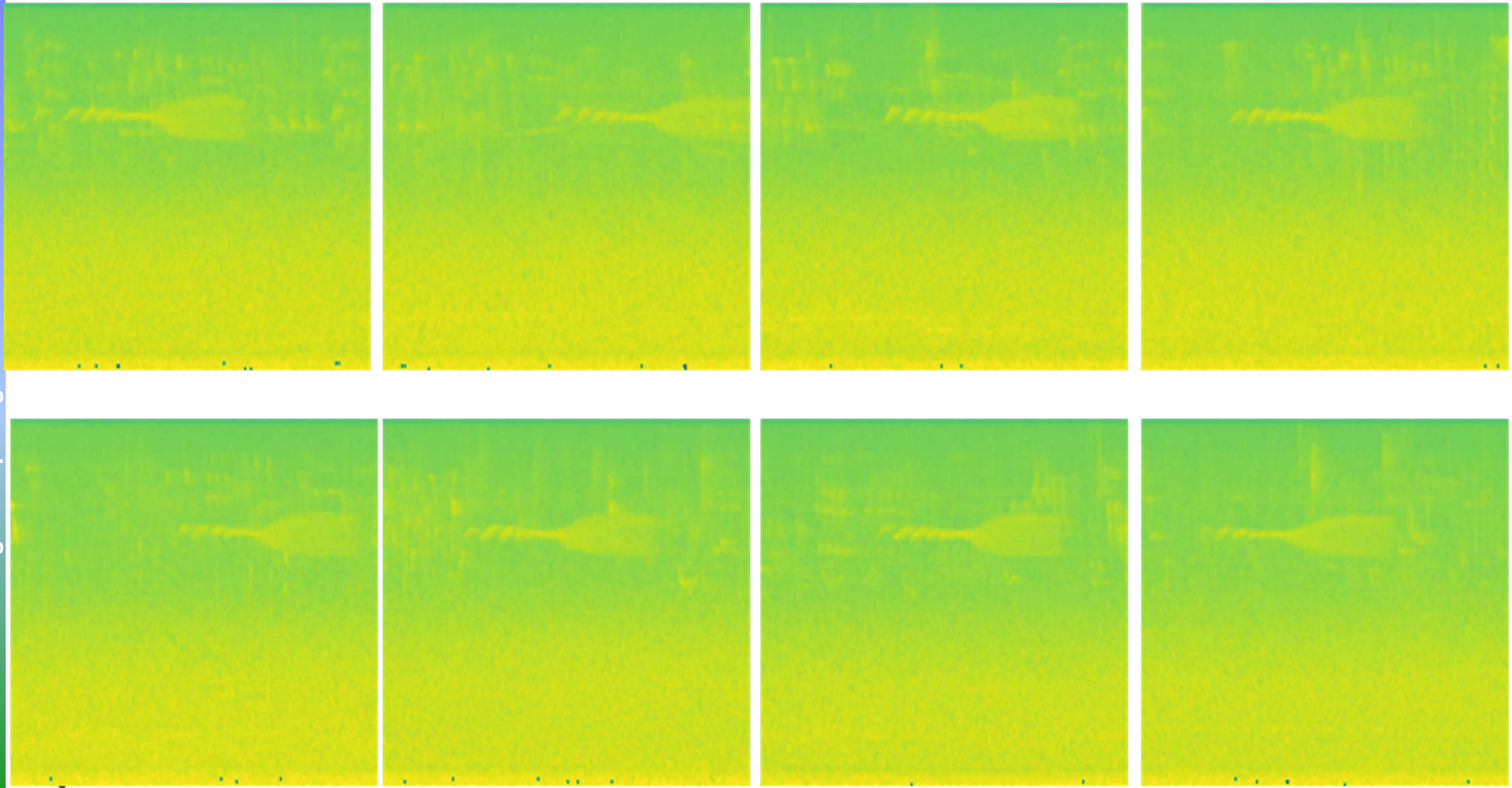


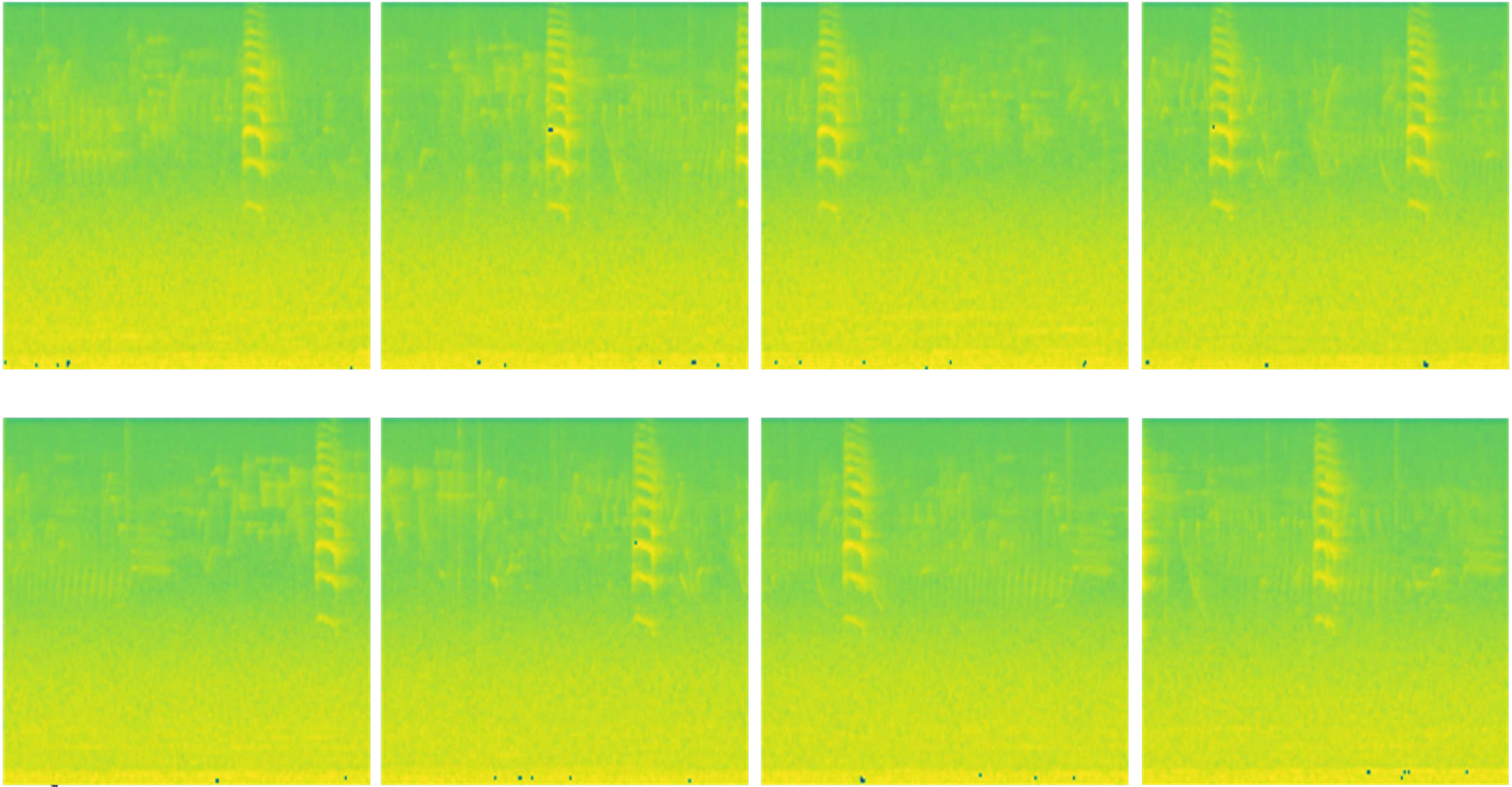
Sage Computing Continuum

Sage Computing Continuum



SAGE
sagecontinuum.org

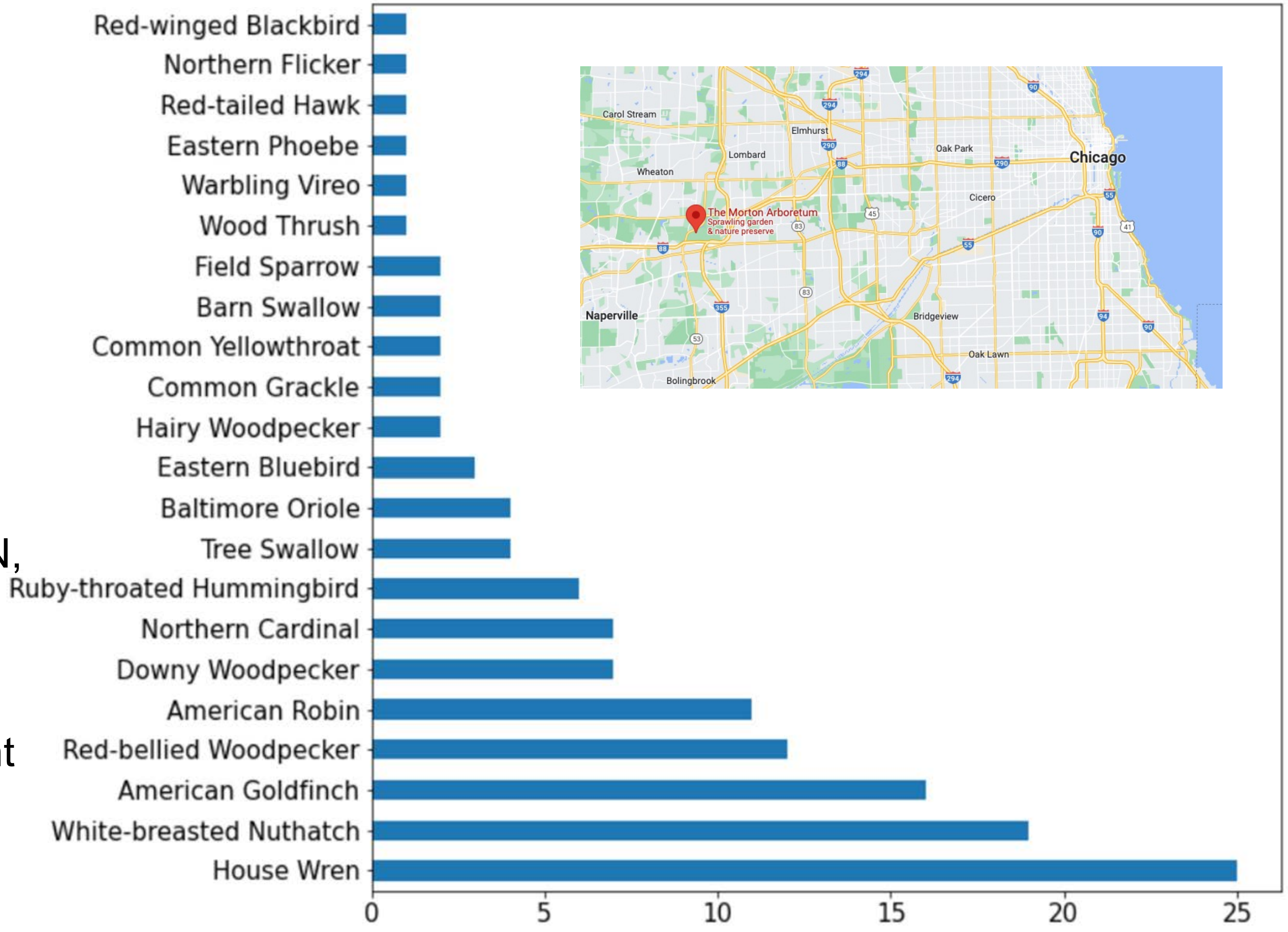




Avian diversity monitoring

- Bird diversity changes as a metric to track the current environmental conditions
- We automate Avian Diversity Monitoring by using a DNN, called BirdNET [1], capable of identifying 984 North American and European bird species by sound. Weekly cumulative detections of non-migratory species occurrence was highly correlated with human point count observations
- It will be possible to get exposure to many organisms occupying diverse areas without needing to detect them during demanding and expensive human fieldwork

[1] Stefan Kahl, Connor M. Wood, Maximilian Eibl and Holger Klinck. BirdNET: A deep learning solution for avian diversity monitoring. Ecological Informatics Volume 61, March 2021.



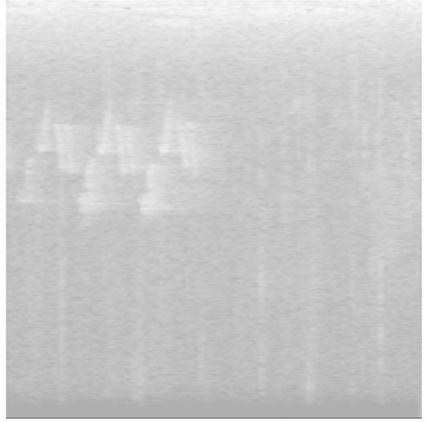
Morton Arboretum Avian Detection, June 28, 2021 (24 hour)

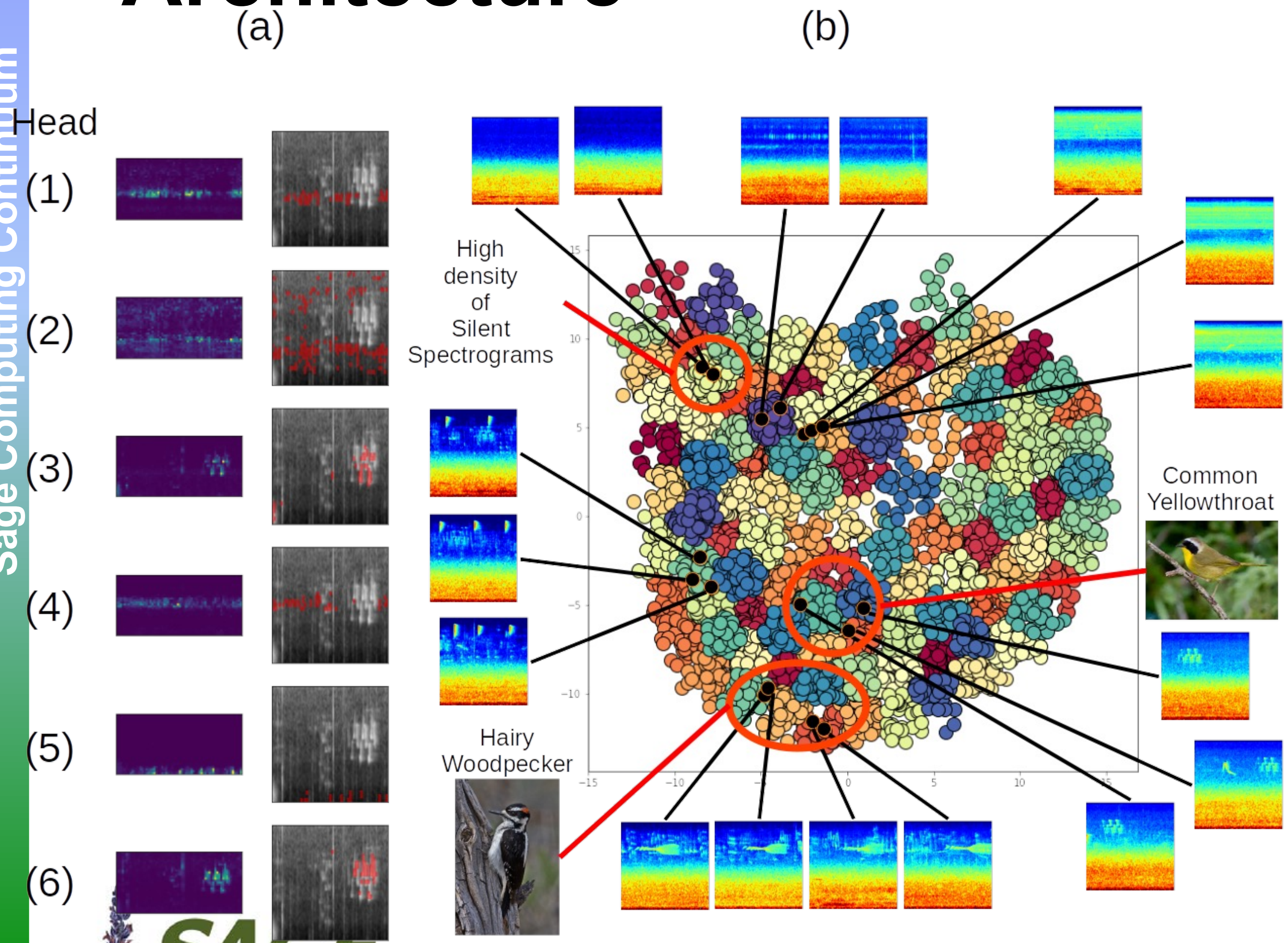
Sage Computing Continuum

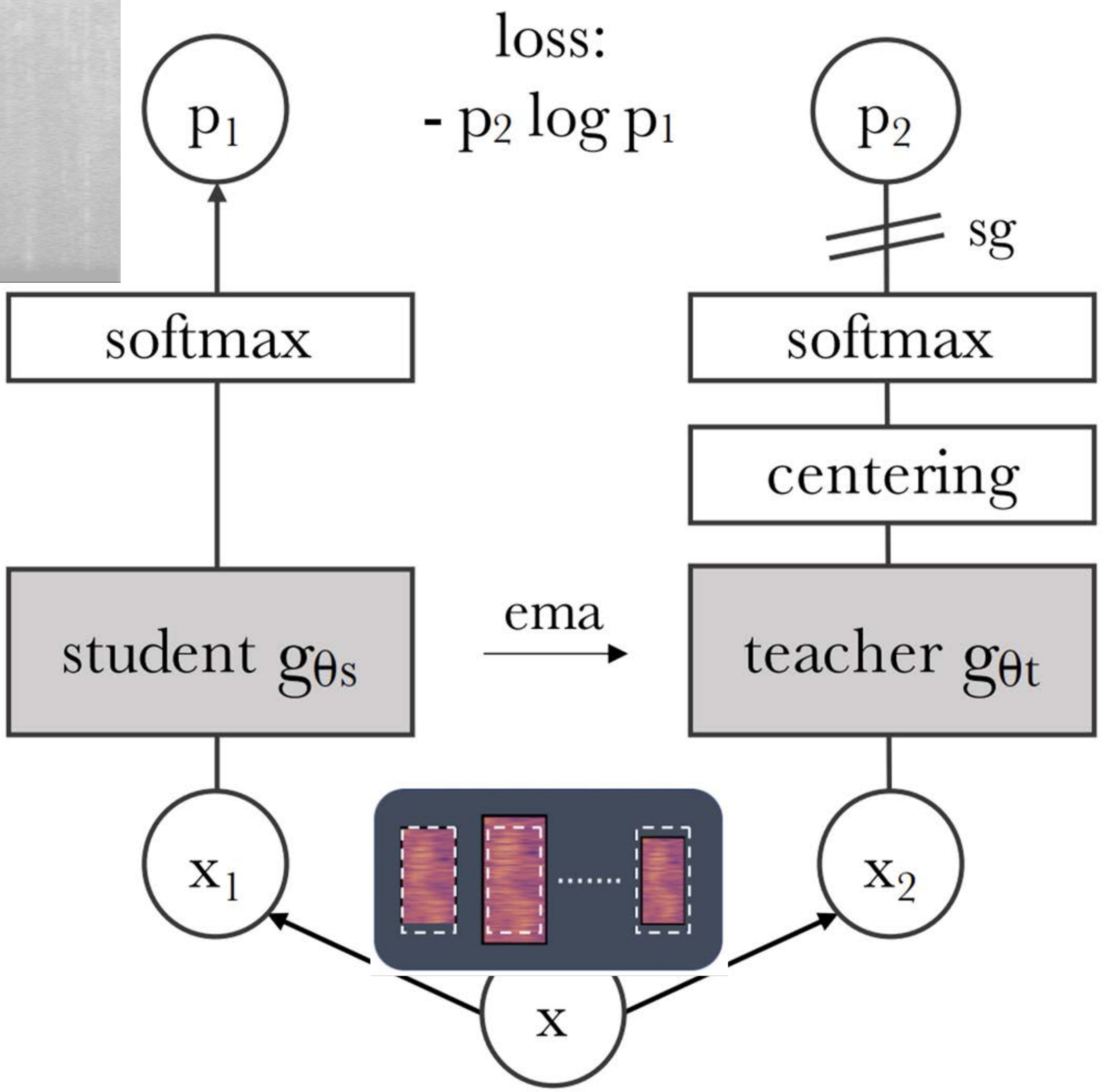
sagecontinuum.org

Research Credit: Dario Dematties, Bhupendra Raut, Nicola Ferrier

Self-supervised Avian Diversity Monitoring Joint-Embedding Architecture







Caron, M., Touvron, H., Misra, I., Jégou, H., Mairal, J., Bojanowski, P. and Joulin, A., 2021. Emerging properties in self-supervised vision transformers. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 9650-9660).

Sage Computing Continuum

SAGE
sagecontinuum.org

Sage Computing Continuum

Detecting measuring stick for estimating water and/or snow d

Importance and motivation

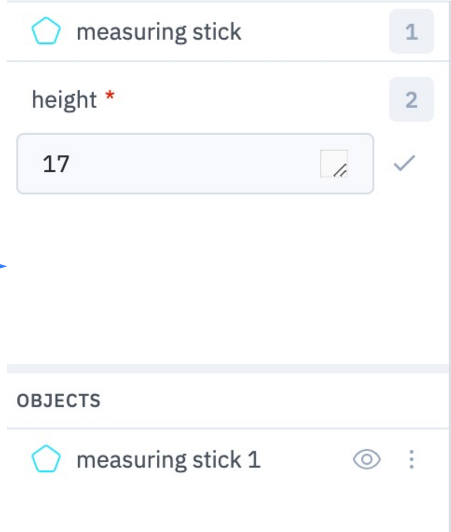
- Monitoring water depth changes in realtime
- Estimating water and/or depth using images to:
  - Analyze current local weather conditions
  - forecast and analyze potential flooding in local streams, rivers, or other water reservoirs when there are heavy rainstorms, rapid snow melting events, or hurricanes[1, 2]

1) N. H. Jafari, X. Li, Q. Chen, C.-Y. Le, L. P. Betzer, and Y. Liang, “Real-time water level monitoring using live cameras and computer vision techniques,”Computers & Geosciences, vol. 147, p. 104642, 2021.
2) L. Sabbatini, L. Palma, A. Belli, F. Sini, and P. Pierleoni, “A computer vision system for staff gauge in river flood monitoring,” Inventions, vol. 6, no. 4, p. 79, 2021.

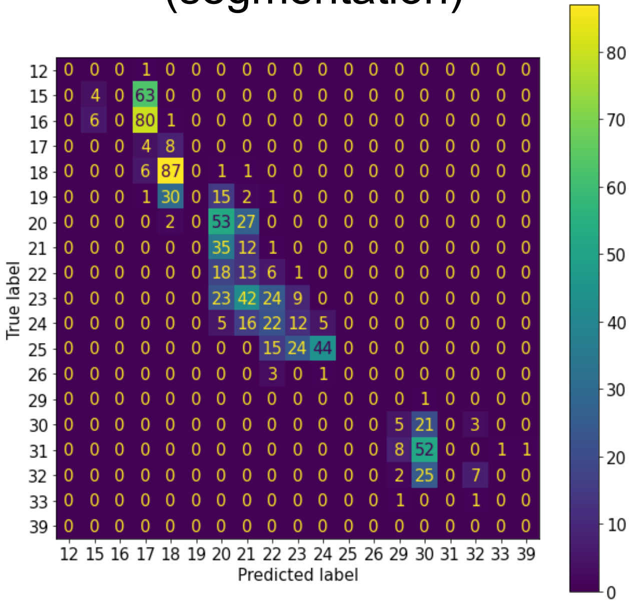
Detecting measuring stick for estimating water and/or snow depth

- Create ground truth images using (human) LabelBox
- Trained a Deep Neural Network models

  → U-shaped Network (U-Net)
- Convert the lowest pixel value to the height of the water level

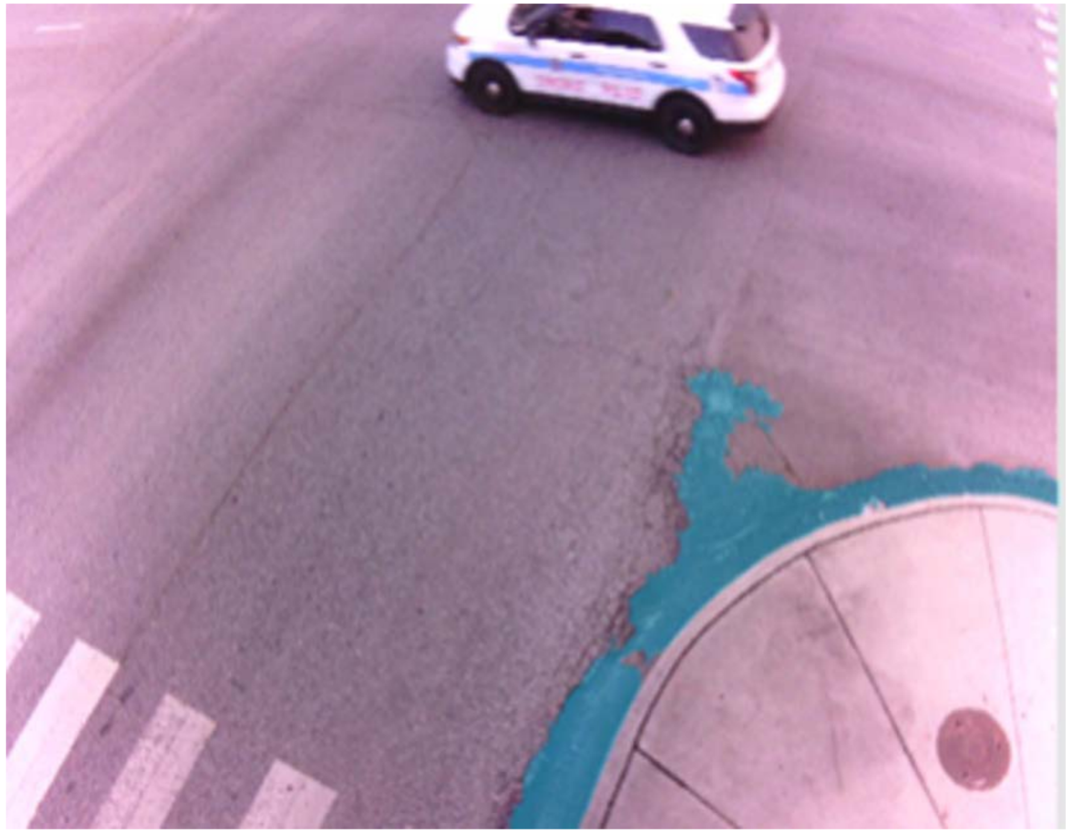


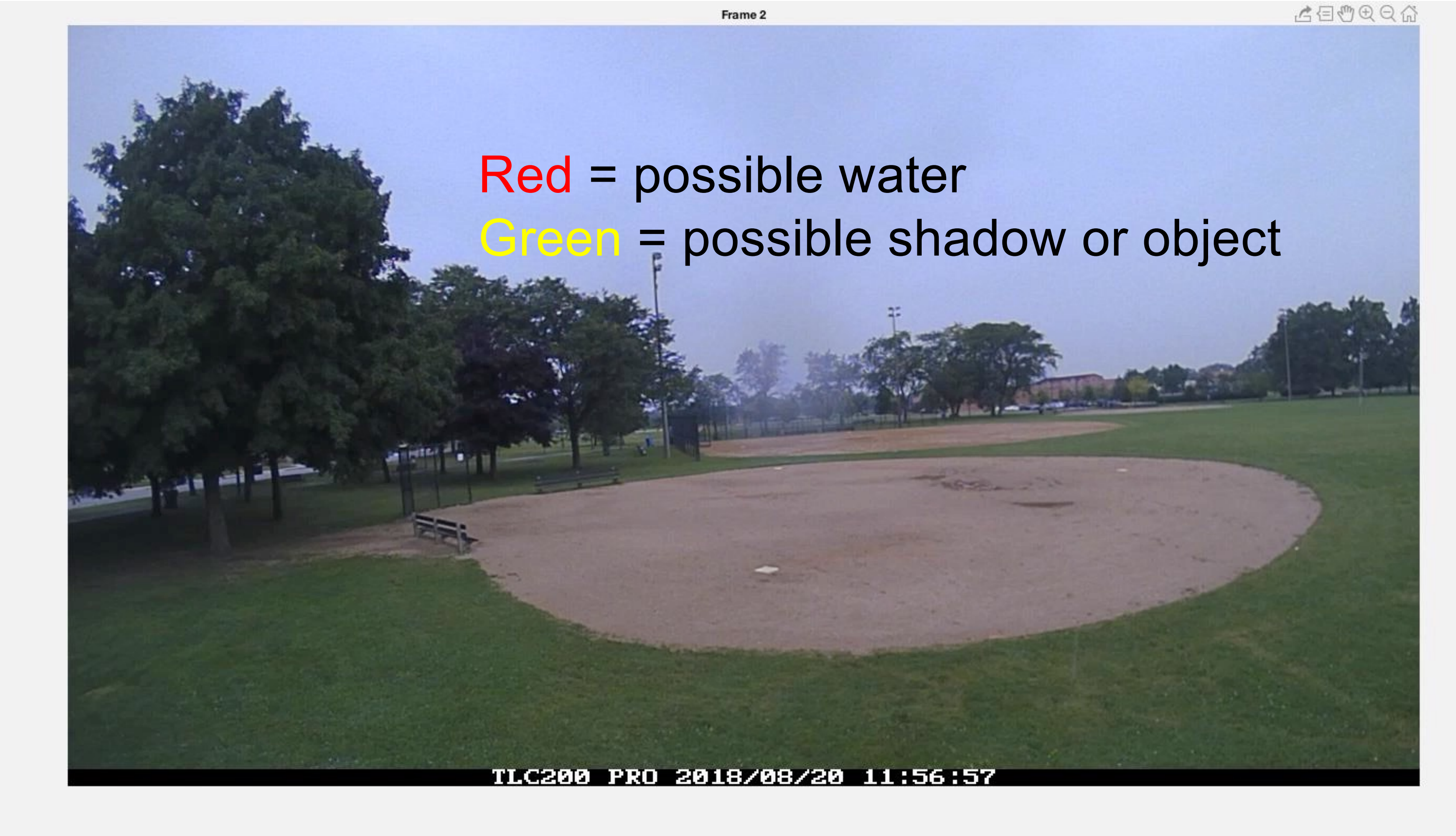
U-Net
(segmentation)

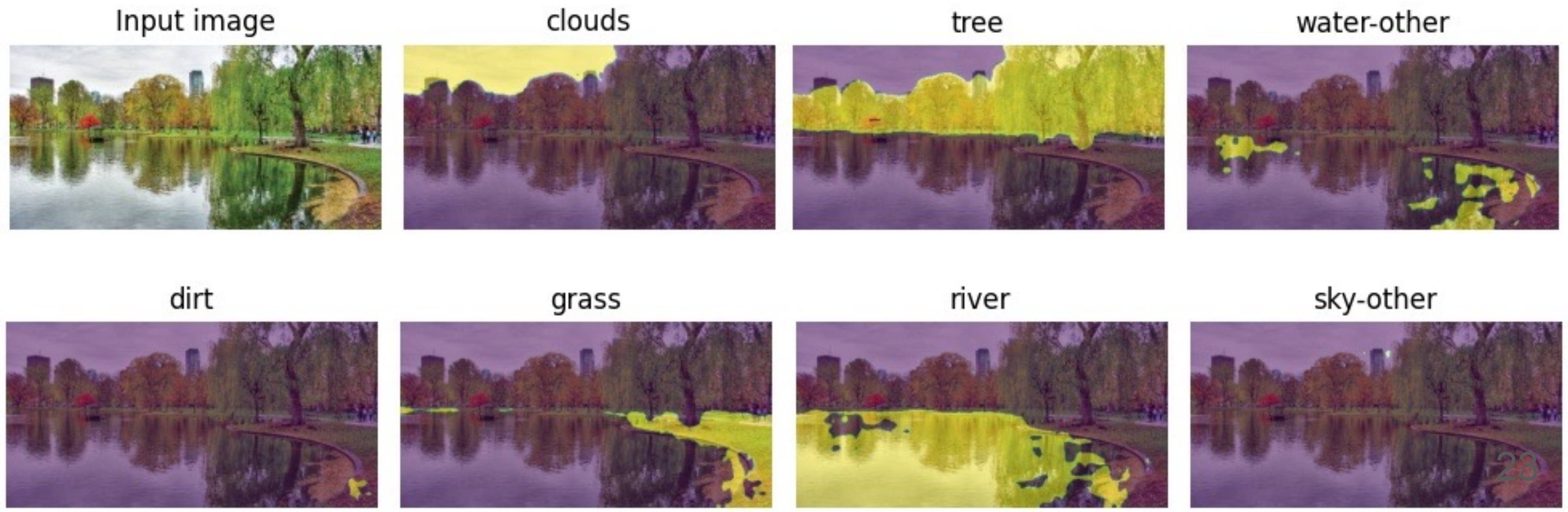


SAGE
sagecontinuum.org

Surface Water Detection







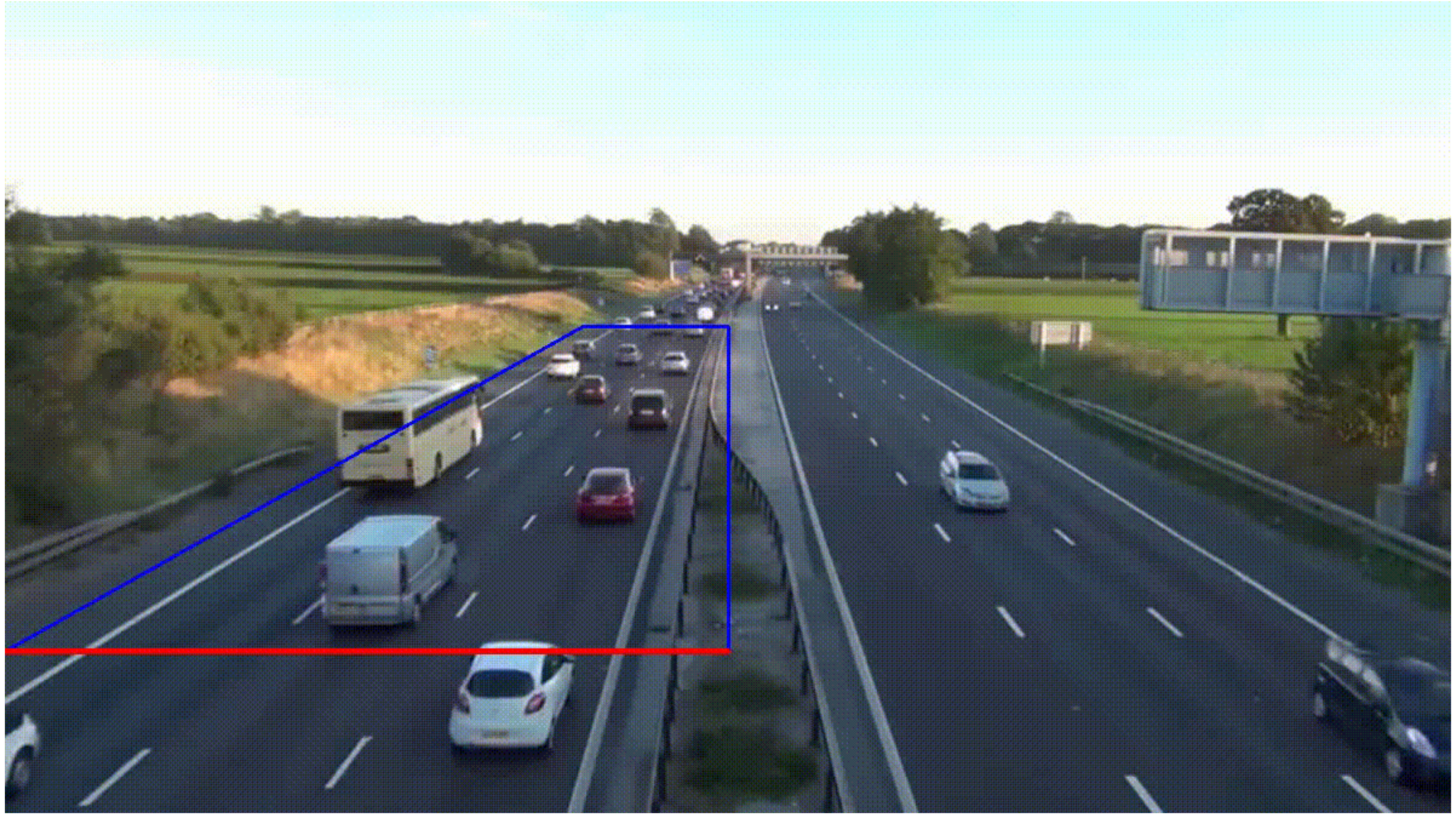
Linked with HPC, can be used to build hydrology models and predictive capabilities

SAGE
sagecontinuum.org

Nicola Ferrier, Uchicago

Measuring traffic state using vehicle tracking

In dense urban areas and places like airports, understanding and analyzing traffic state can be a starting point for brining traffic regulations that enable effective and efficient movement of people, goods, and services.

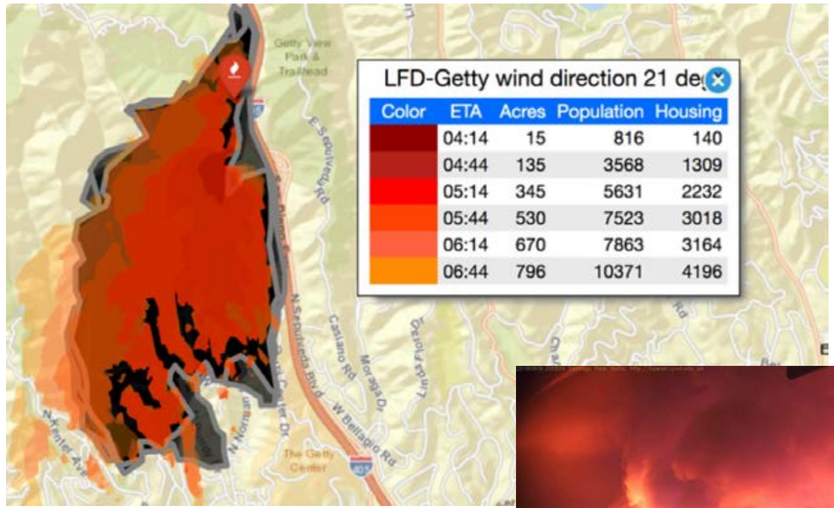


Ilkay Altintas, UCSD, Co-PI for SAGE

Wildfire Detection and Prediction

WIFIRE

Exploring wildfire detection at the edge linked to HPC simulations

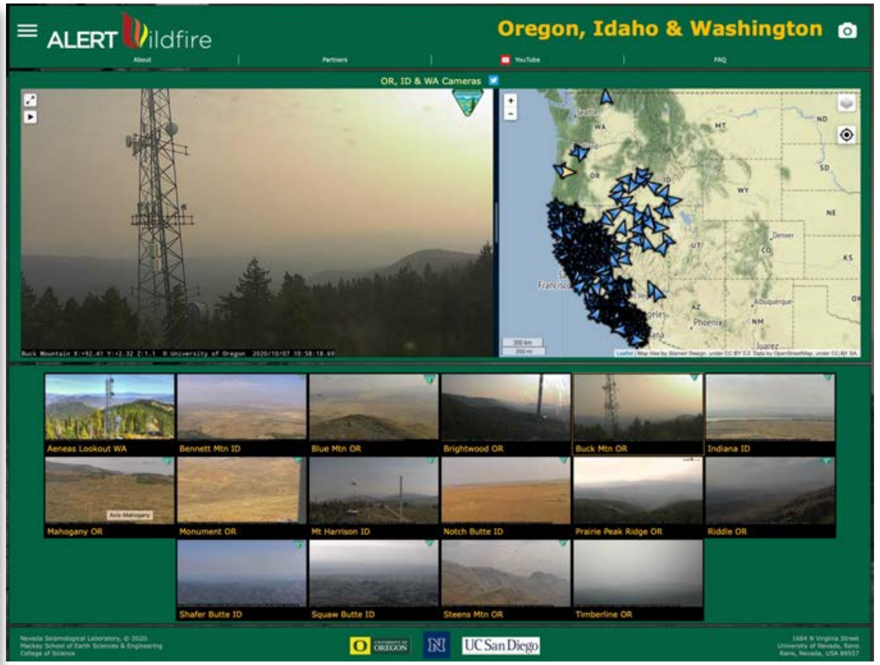


ALERTWildfire: A unique wildfire detection and monitoring system

UNIVERSITY OF OREGON

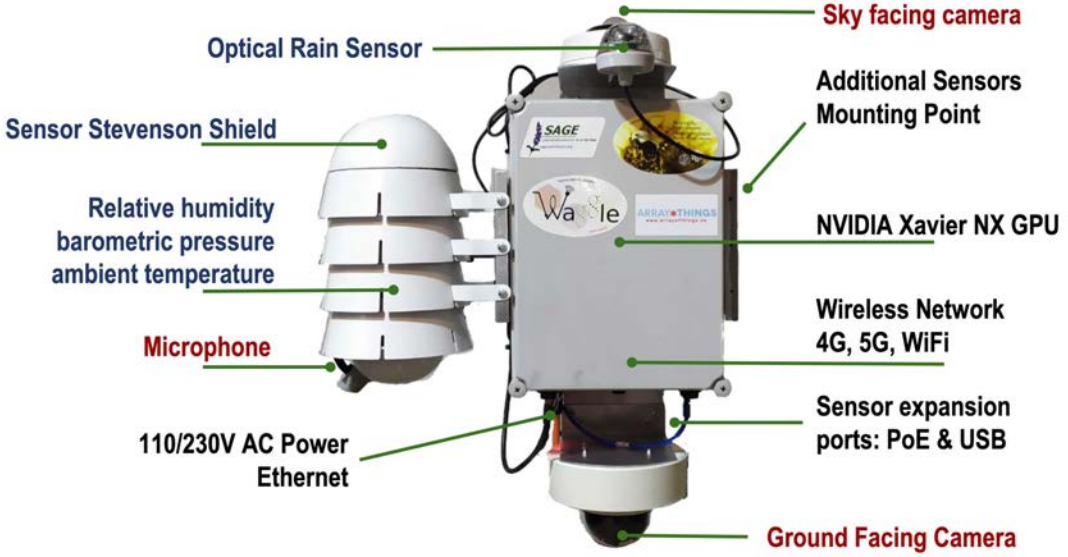
OHAZ

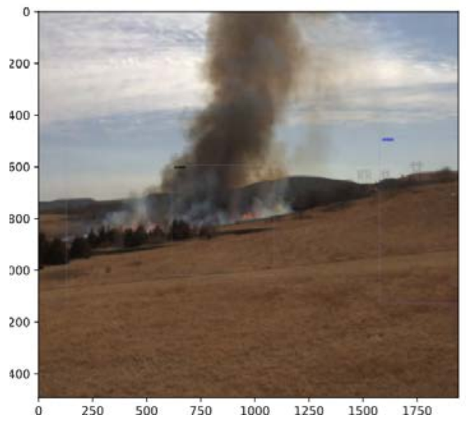
Collaboration: Doug Toomey, UOregon

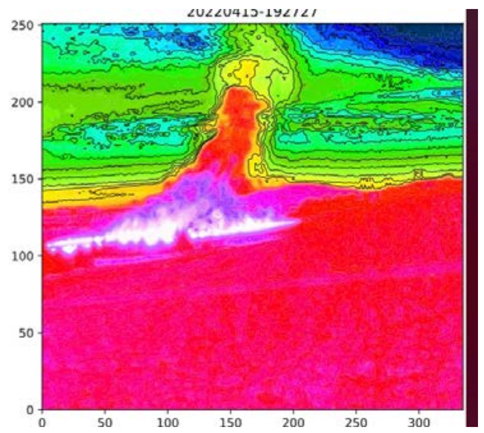


NEON Mobile Deployment Platform (MPD) with Sage Konza Prairie for controlled burn: April 2022.

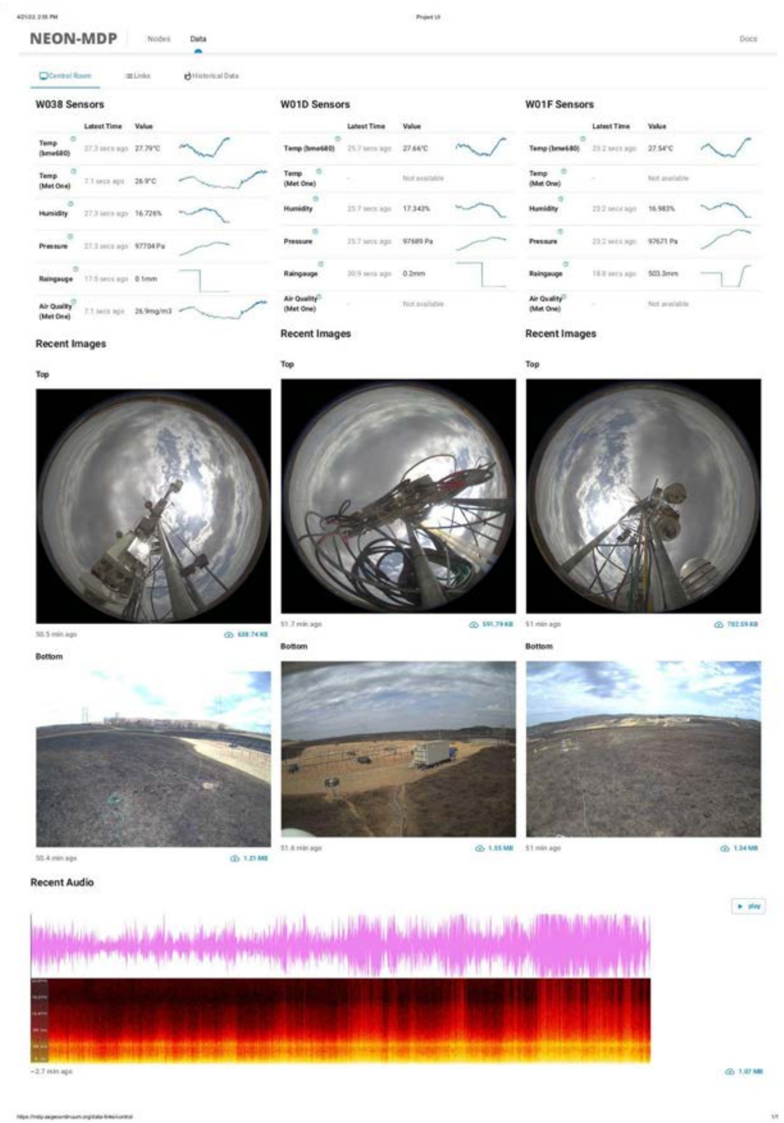
Sage Co-PI: Eugene Kelly, Colorado State
eugene.kelly@colostate.edu

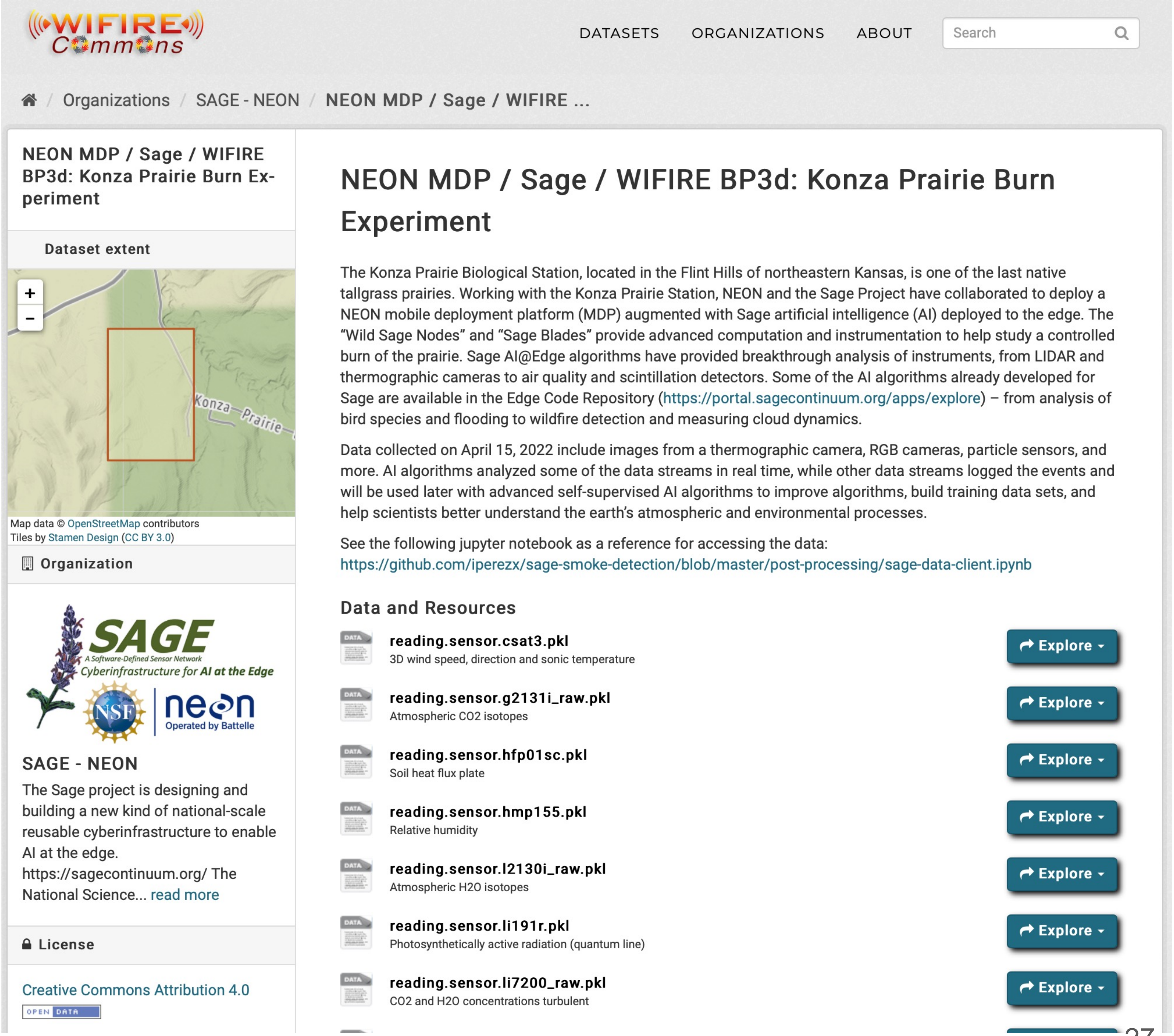






Data from the experiment available to the community!





WIFIRE Commons

DATASETS ORGANIZATIONS ABOUT Search

Organizations / SAGE - NEON / NEON MDP / Sage / WIFIRE ...

NEON MDP / Sage / WIFIRE BP3d: Konza Prairie Burn Experiment

Dataset extent

Map data © OpenStreetMap contributors
Tiles by Stamen Design (CC BY 3.0)

Organization

SAGE - NEON

The Sage project is designing and building a new kind of national-scale reusable cyberinfrastructure to enable AI at the edge. https://sagecontinuum.org/ The National Science... read more

License

Creative Commons Attribution 4.0

NEON MDP / Sage / WIFIRE BP3d: Konza Prairie Burn Experiment

The Konza Prairie Biological Station, located in the Flint Hills of northeastern Kansas, is one of the last native tallgrass prairies. Working with the Konza Prairie Station, NEON and the Sage Project have collaborated to deploy a NEON mobile deployment platform (MDP) augmented with Sage artificial intelligence (AI) deployed to the edge. The “Wild Sage Nodes” and “Sage Blades” provide advanced computation and instrumentation to help study a controlled burn of the prairie. Sage AI@Edge algorithms have provided breakthrough analysis of instruments, from LIDAR and thermographic cameras to air quality and scintillation detectors. Some of the AI algorithms already developed for Sage are available in the Edge Code Repository (https://portal.sagecontinuum.org/apps/explore) – from analysis of bird species and flooding to wildfire detection and measuring cloud dynamics.

Data collected on April 15, 2022 include images from a thermographic camera, RGB cameras, particle sensors, and more. AI algorithms analyzed some of the data streams in real time, while other data streams logged the events and will be used later with advanced self-supervised AI algorithms to improve algorithms, build training data sets, and help scientists better understand the earth’s atmospheric and environmental processes.

See the following jupyter notebook as a reference for accessing the data:
https://github.com/iperezx/sage-smoke-detection/blob/master/post-processing/sage-data-client.ipynb

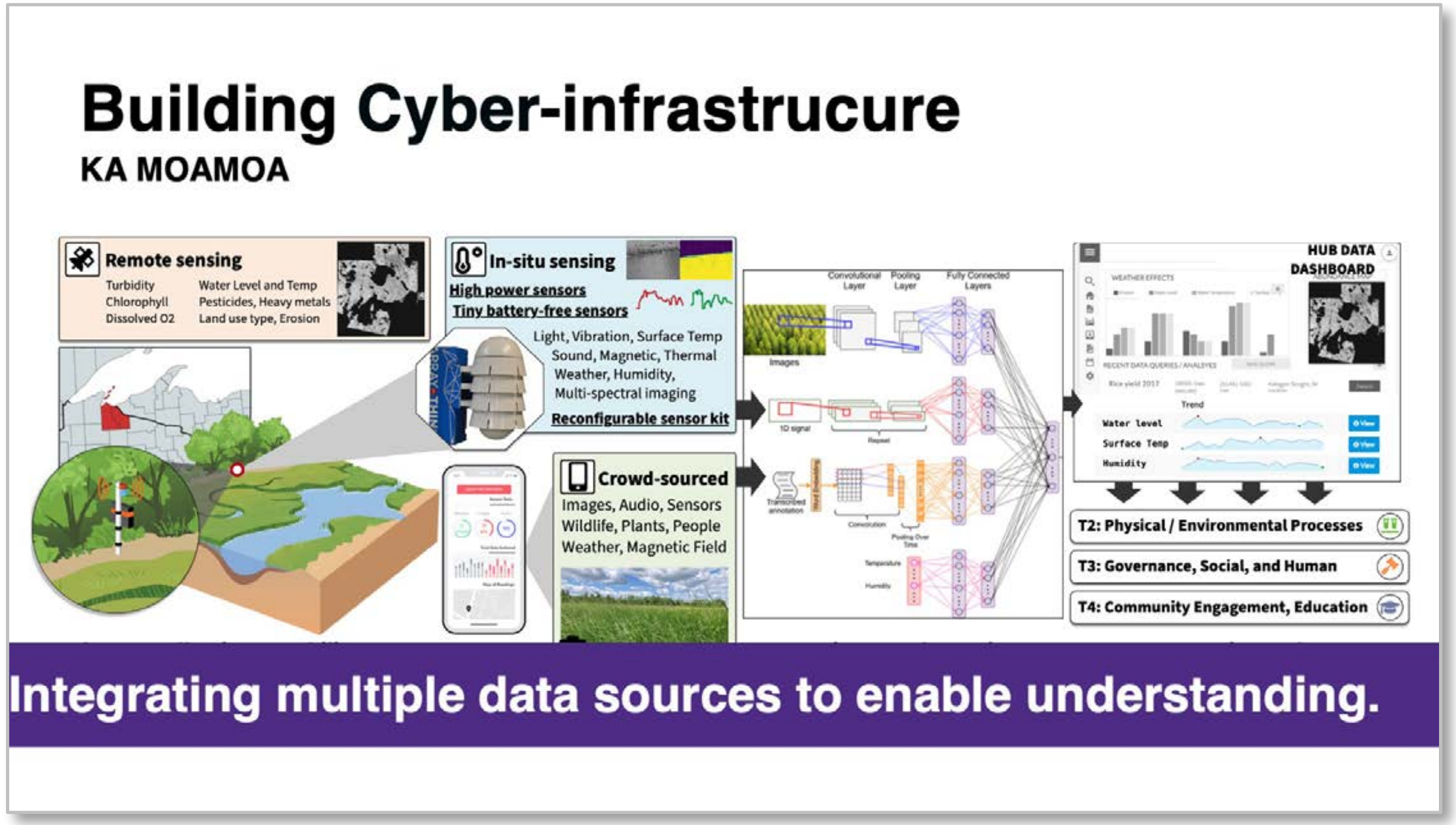
Data and Resources

- **reading.sensor.csat3.pkl** — 3D wind speed, direction and sonic temperature — Explore
- **reading.sensor.g2131i_raw.pkl** — Atmospheric CO2 isotopes — Explore
- **reading.sensor.hfp01sc.pkl** — Soil heat flux plate — Explore
- **reading.sensor.hmp155.pkl** — Relative humidity — Explore
- **reading.sensor.l2130i_raw.pkl** — Atmospheric H2O isotopes — Explore
- **reading.sensor.li191r.pkl** — Photosynthetically active radiation (quantum line) — Explore
- **reading.sensor.li7200_raw.pkl** — CO2 and H2O concentrations turbulent — Explore

Ismael Perez and Ilkay Altintas, UCSD

NSF Coastlines and People

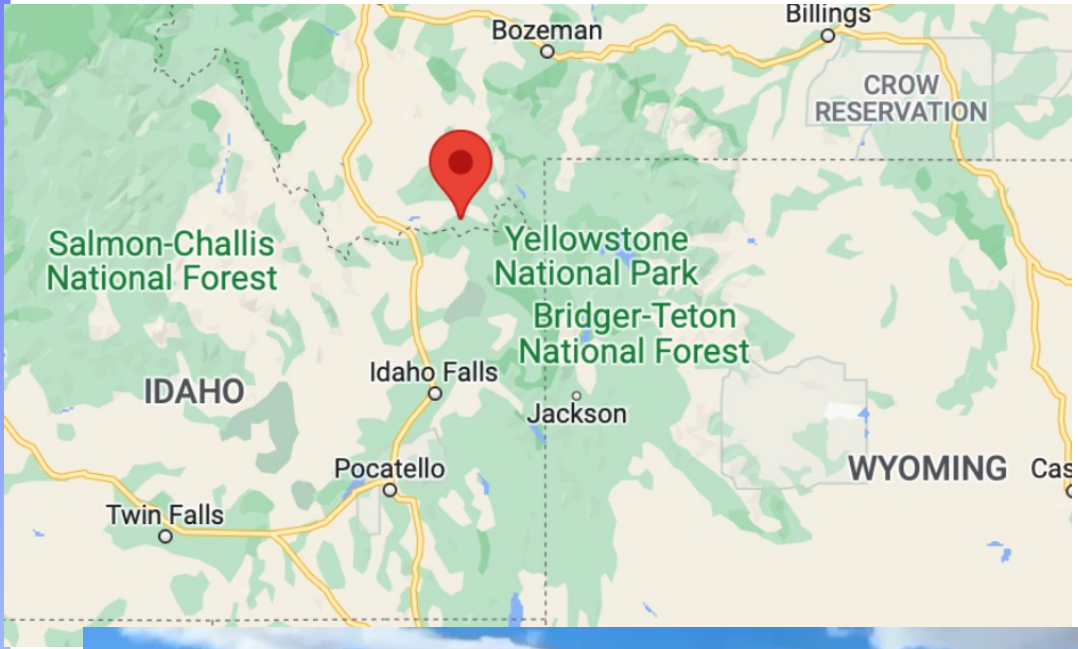
Strengthening Resilience of Manoomin, the Sentinel Species of the Great Lakes, with Data-Science Supported Seventh Generation Stewardship
PI: Josiah Hester



Jonathan Gilbert, Biological Services Director, Great Lakes Indian Fish & Wildlife Commission (GLIFWC)

Sage Computing Continuum

Wild Sage Node Deployment: University of Utah's Taft-Nicholson Center in Montana



Motivated by the success of the Konza burn, we are planning deployment of 5 mobile Sage towers. The first deployment will be at a remote site in Montana .

Two phase deployment:

Initial Deployment on campus with line power and Starlink Internet (university network as back up).

Final deployment on a hilltop powered by solar and wind, and Starlink.

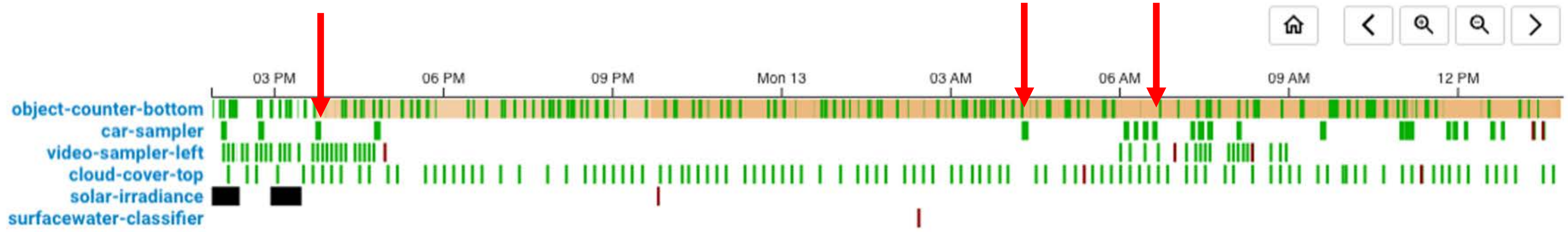
Sensors: *Sage node with cameras, microphone, TPH, precipitation, dust and thermal camera.*

EDGE SCHEDULER AND JOBS

- ***Science Goals (rules)*** are being developed and improved. Scientists can activate edge applications to capture critical data
- More advanced science rules will come available to scientists
- (https://docs.waggle-edge.ai/docs/tutorials/schedule-jobs#creating-job-description-with-advanced-science-rules-for-supporting-realistic-science-mission)

Science Rules

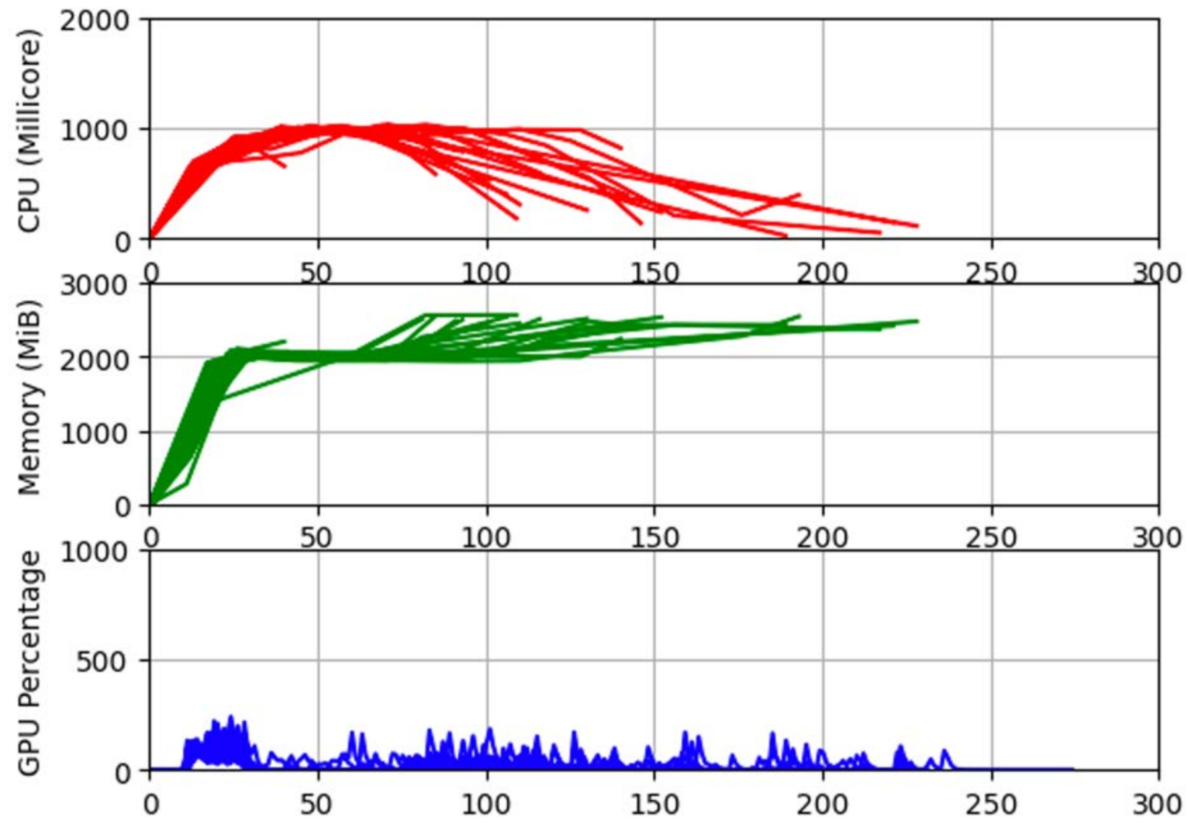
```
schedule(cloud-cover-top): cronjob('cloud-cover-top', '*/10 * * * *')
schedule(solar-irradiance): True
schedule(object-counter-bottom): cronjob('object-counter-bottom', '*/5 * * * *')
schedule(car-sampler): any(v('env.count.car', since='-5m', last=True) >= 1)
publish(env.event.car.detected):  any(v('env.count.car', since='-5m', last=True) >= 1)
set(env.event.car.detected, value=true): any(v('env.count.car', since='-5m', last=True) >= 1)
```



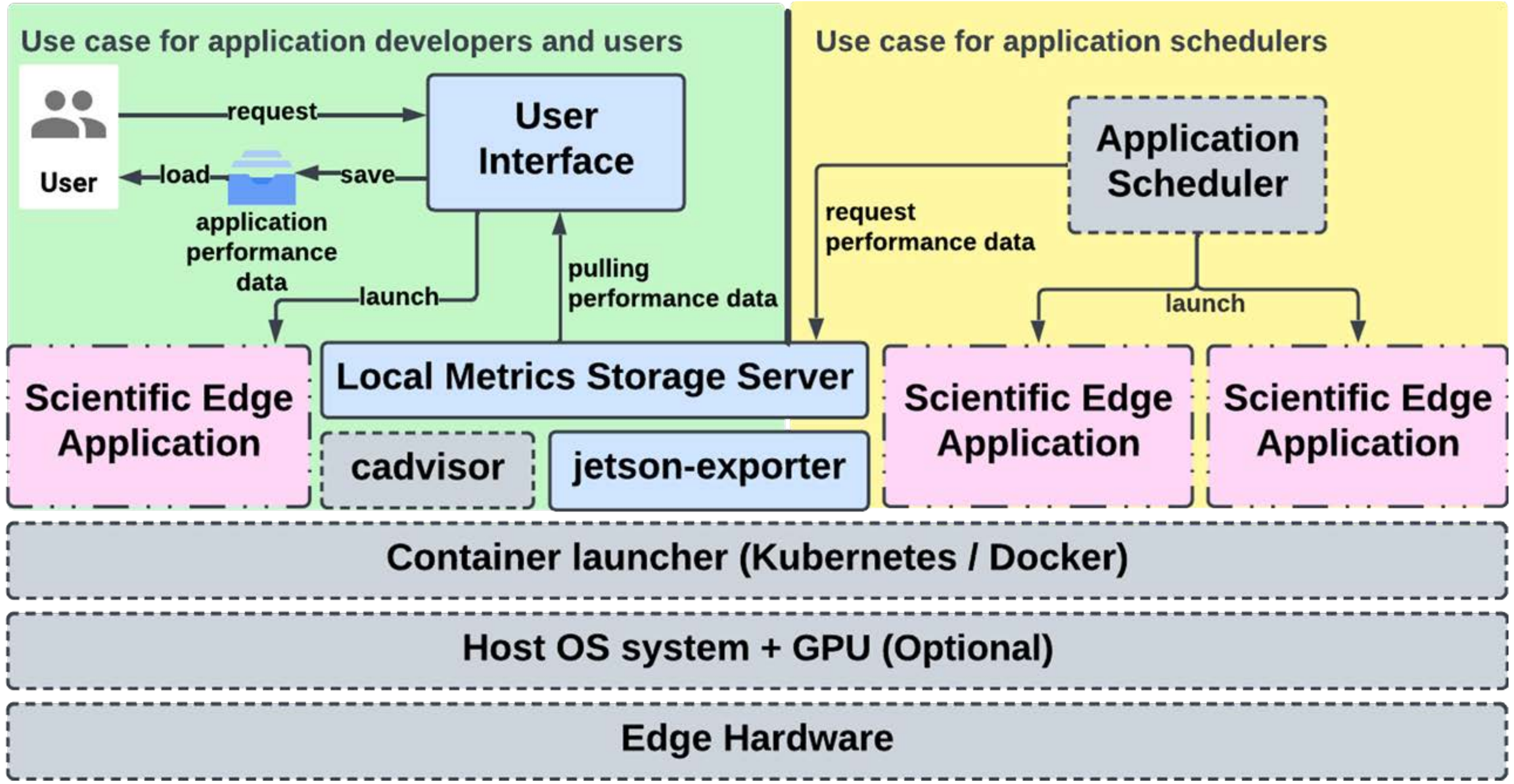
the car-sampler application runs whenever the object counter application detects a car (*obviously the car-sampler ran mostly on the rush-hours*)

PERFORMANCE PROFILING OF EDGE APPLICATIONS

- Edge applications need to be performance profiled to inform the scheduler the amount of resource needed to run the application
- Waggle's pluginctl tool allows performance profiling, https://github.com/waggle-sensor/edge-scheduler/blob/main/docs/pluginctl/tutorial_profiling.md



an illustration of overlaid resource profile of 24 runs of the object counter application



Waggle application performance profiling framework

Sage Computing Continuum

sagecontinuum.org

Exciting, Hard, Challenging, CS Problems:

From Instrument to the HPC/Cloud

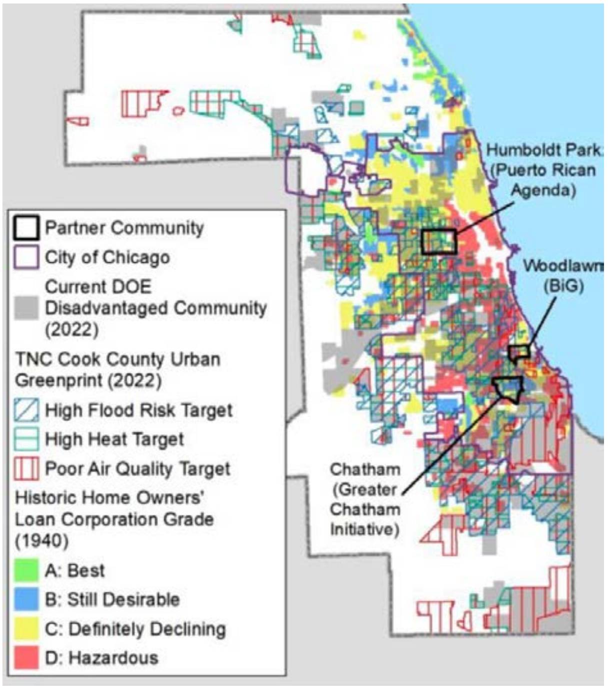
POWDER

Instrument → **HPC/Cloud**

USRP B210 SDR

Utah Field Museum

- Programming model for the Digital Continuum
- Lightweight AI training / model adaptation at instrument edge
- Self-supervised learning with multiple instruments
- Container technology for Cloud/HPC and the edge
- Cooperative sharing: multi-tenancy
- Control loops for actuation, steering
- Movement (drones, robots)
- Digital twin / MODEX for setting local edge goals
- Microelectronics for low-power AI@edge
- Advanced networking: wireless and satellite

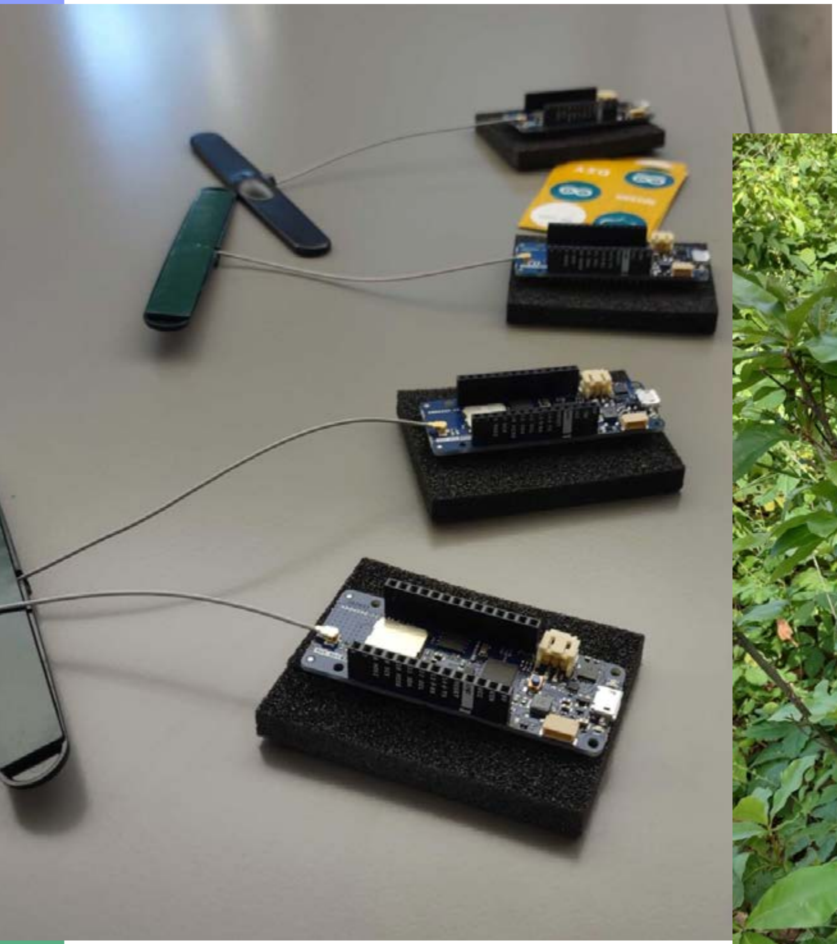


CROCUS
Community Research on Climate & Urban Science

SAGE
sagecontinuum.org

Sage Computing Continuum

”More room at the bottom” : LoRaWAN & 5G

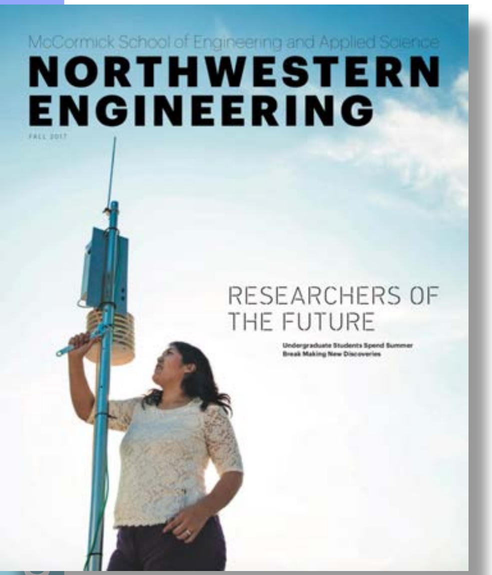


SAGE
sagecontinuum.org

Larry Hartman: Uoregon
Eugene Kelly: Colorado State

SAGE Education Kits

- **Outreach and Engagement**
  - **Low-cost Sage node (~$300)**
    - Hardware
    - NVIDIA Jetson Nano
    - Camera
    - Microphone
    - Environmental sensor (temperature, humidity, barometric pressure, and VOC gas)
    - Setup instructions
    - Jupyter notebook demonstrations for each sensor
    - Jupyter notebook project that combines microphone and sensor
  - **Workshops/Camps**
  - **Undergraduate and Graduate Research Efforts**



NVIDIA Jetson Nano based kit for Students

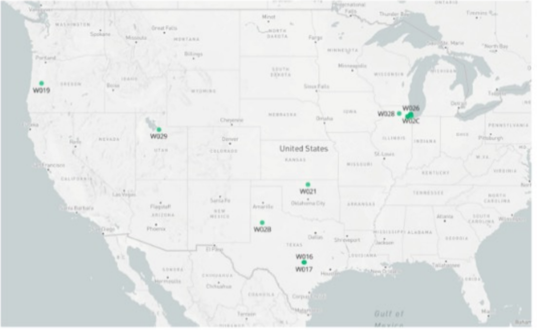
SageEDU nodes: github.com/ddiLab/SageEdu

Sample Project using SageEDU node: Identify if air conditioner is running

- Microphone to detect noise
- Sensor to read temperature

Aldo Malaquias Cabrera (UChicago), **Andrew Garcia (UChicago)**, **Eric Sledge (Claflin)**, **Joseph Katula (NIU)**, Rajesh Sankaran, Valerie Taylor, Michael Papka

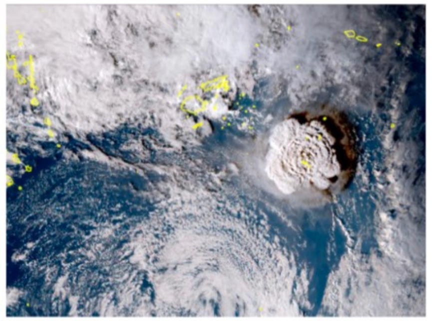
Anomaly Detection and Multi-messenger Science



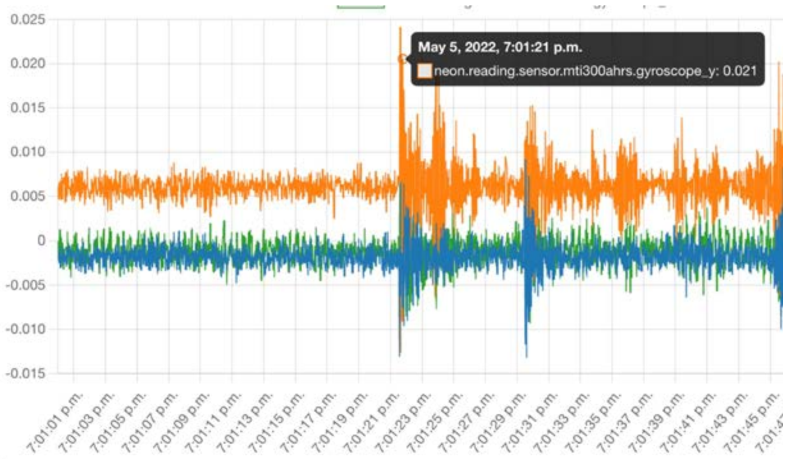
2022 Hunga Tonga Eruption

		UTC		
		First Peak	Second Peak	Third Peak
GC Distance	Node Location	1/16		1/17
10357KM	Billings (W021)	6:58		19:08
10039KM	Austin (W016)	7:14	23:33	19:23
9851KM	Lubbock (W02B)	7:21	23:41	19:31
8964KM	Eugene (W019)	8:16	23:56	20:26

Chicago area nodes registered little to no changes in atmoshperic pressure readings.



(Sensors sample at 40hz, aggregate to 30min)

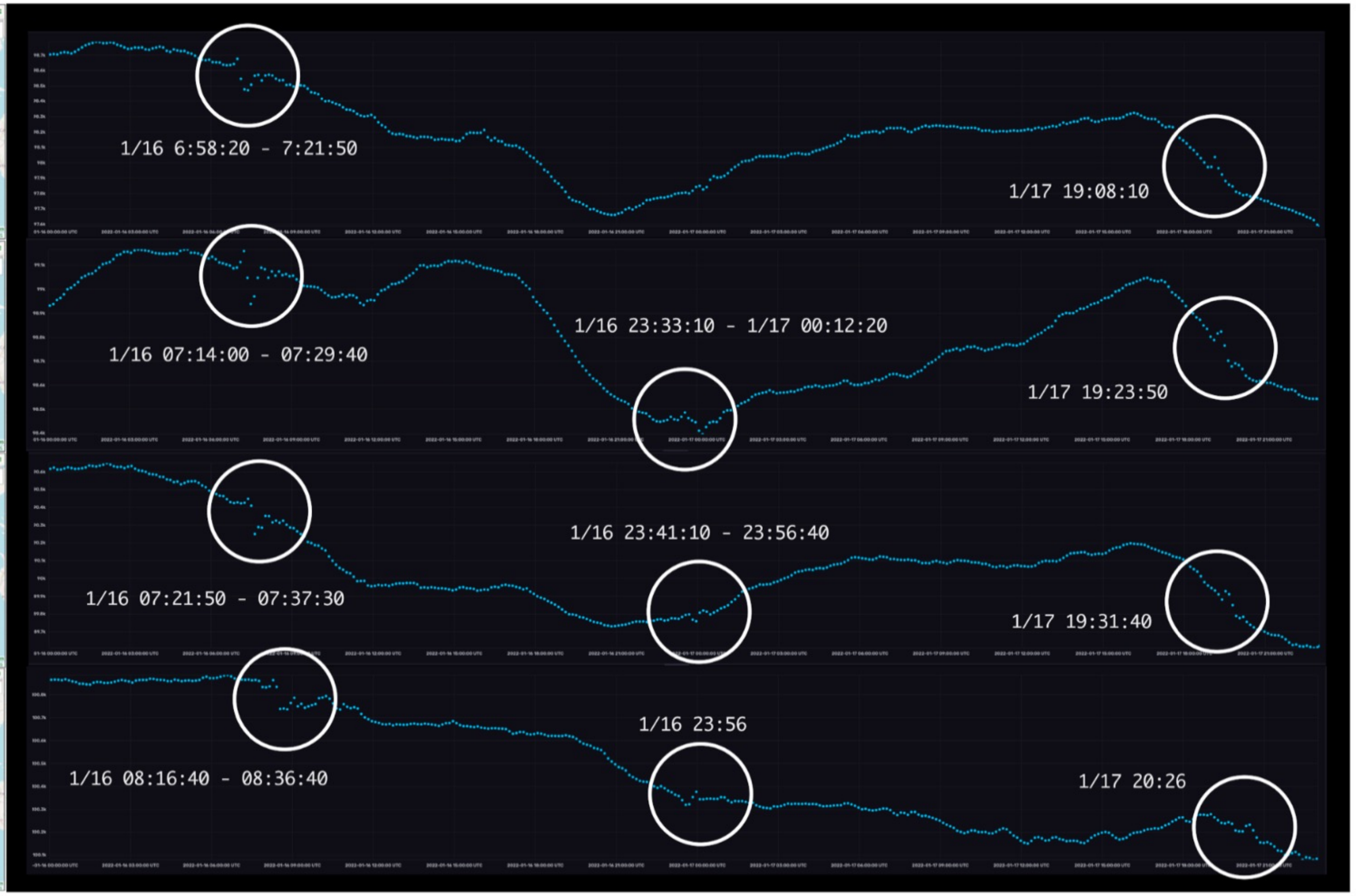


Billings, OK
10357KM

Austin, TX
10039KM

Lubbock, TX
9851KM

Eugene, OR
8964KM

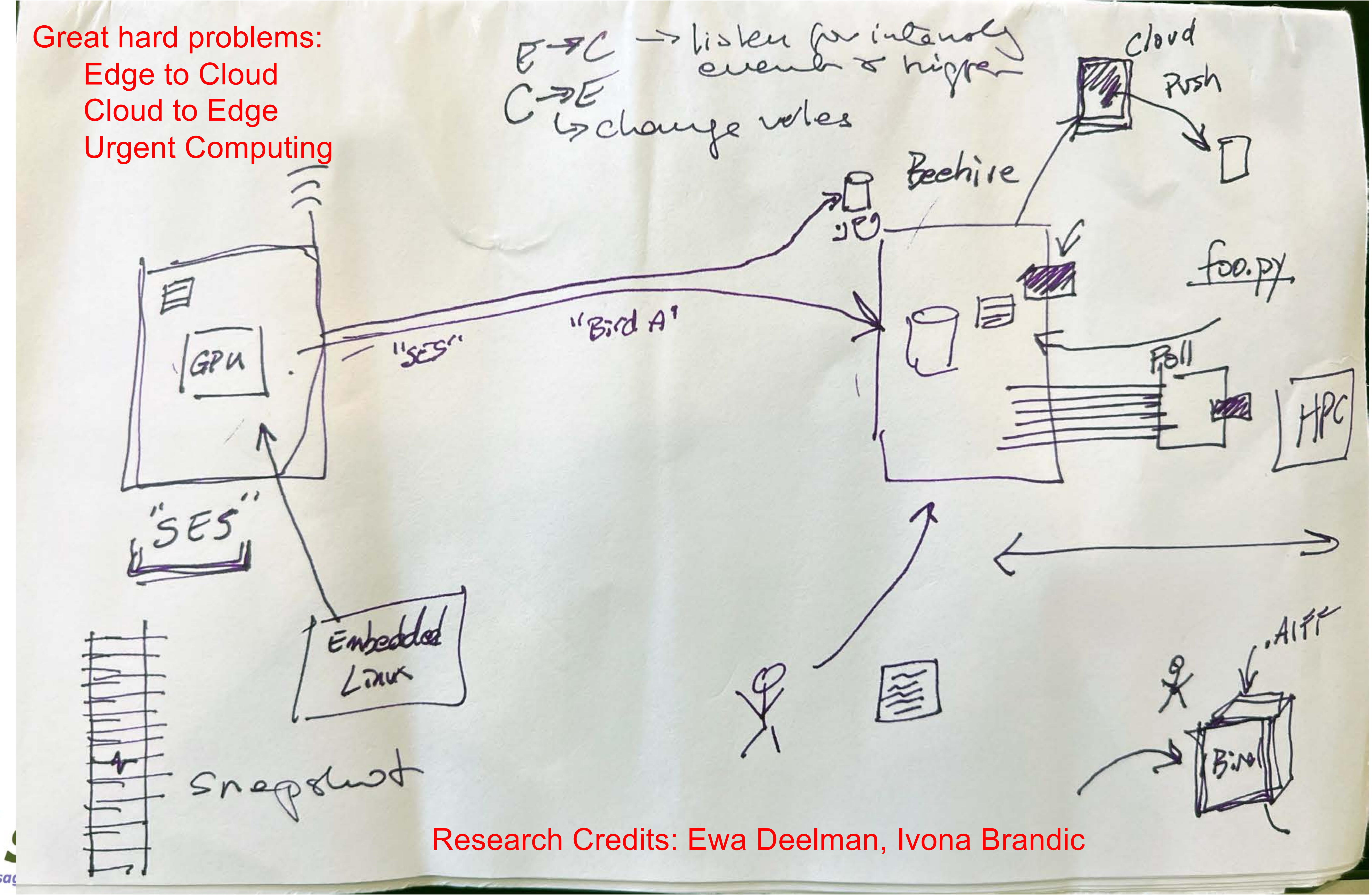


SAGE
sagecontinuum.org

Preliminary data graphs from Sage Nodes not yet curated for peer review. http://sagecontinuum.org/
Pressure data measured by BOSCH BME680 Sensors. Individual sensors were not calibrated post install. (1/27/2022)

Great hard problems:

- Edge to Cloud
- Cloud to Edge
- Urgent Computing



Research Credits: Ewa Deelman, Ivona Brandic

AI@Edge at Argonne APS

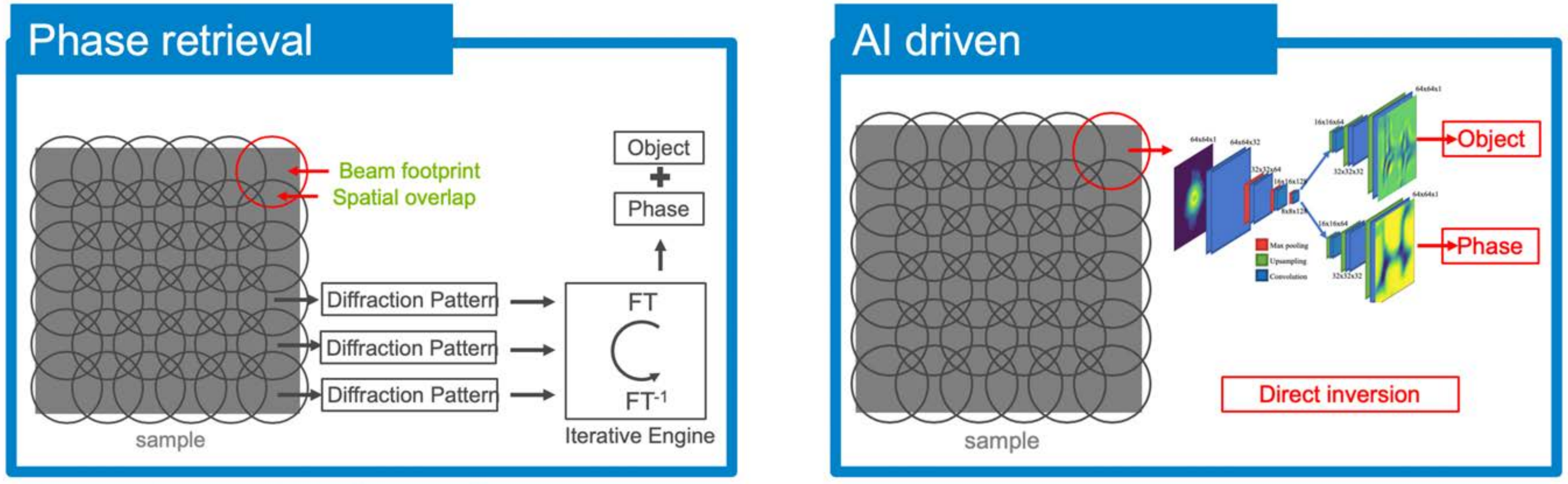
AI@EDGE ENABLES REAL-TIME PTYCHOGRAPHY

Train AI @ ALCF, deploy AI @ beamline

- Real-time imaging: >100X faster than phase retrieval
  - Demonstrated live inference at 100 Hz on 512x512 detector im
- Lower-dose imaging : 25X less data than phase retrieval
- Future work: other techniques, closed-loop experimental steering

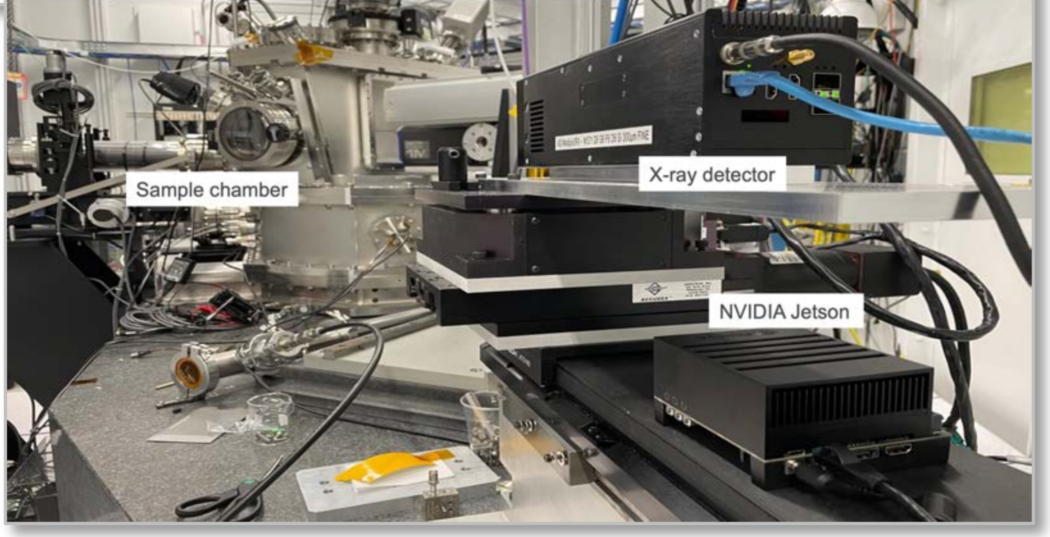
REINVENTING COHERENT IMAGING DATA INVERSION

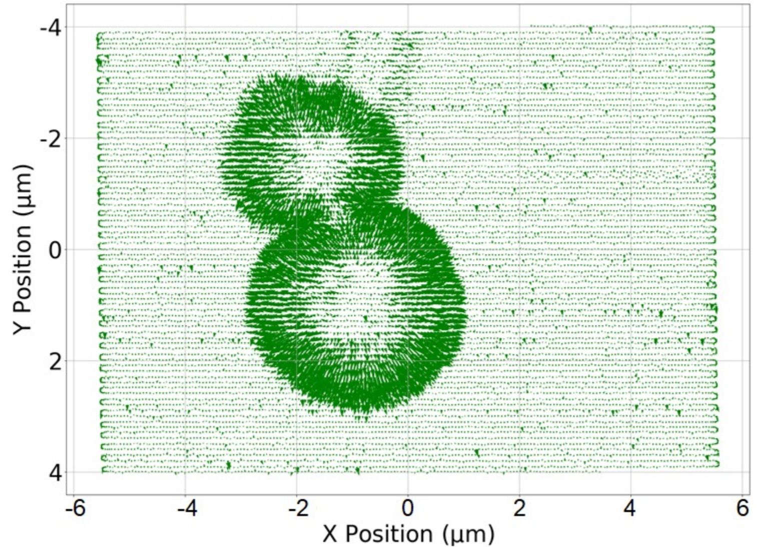
PtychoNN: Accelerating analysis



Requires >PFLOPs of on-demand to keep up with experiments → **PtychoNN is >100X faster Requires 25X less data**

AI@EDGE FOR PTYCHOGRAPHY





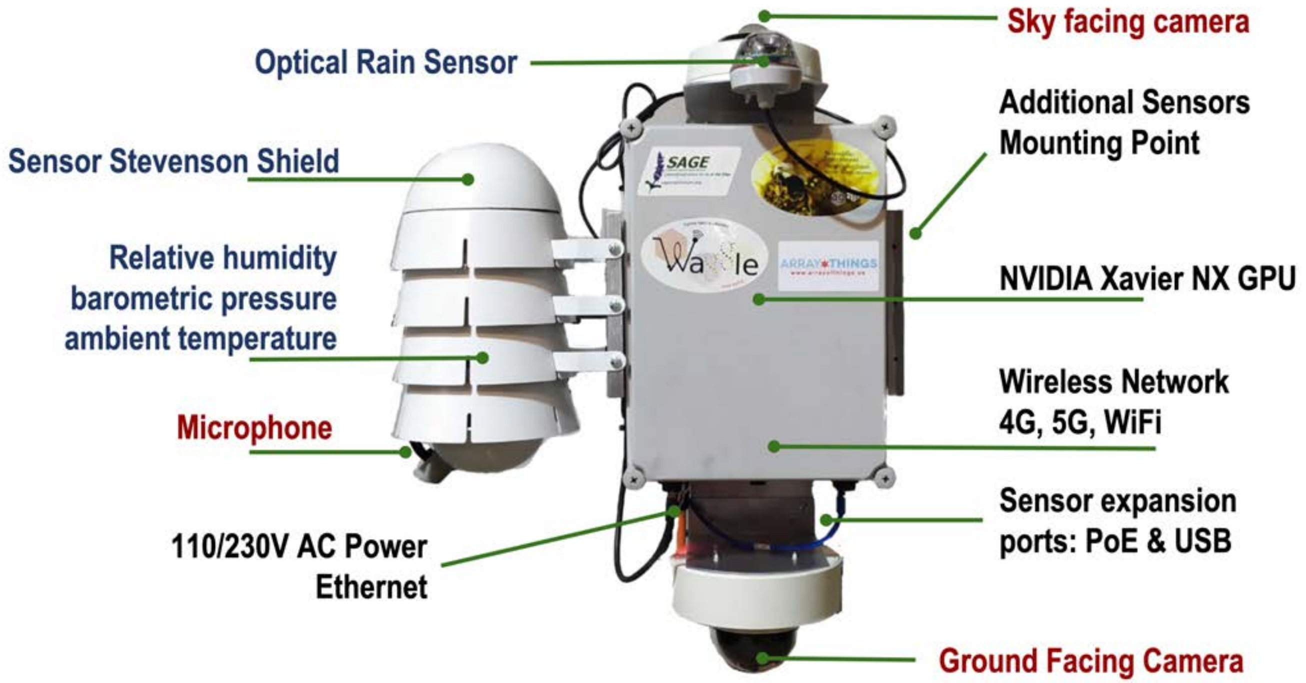
Mathew Cherukara
<mcherukara@anl.gov>

SAGE
sagecontinuum.org

Sage Computing Continuum

Let’s Solve some Hard Problems Together!

- Fun with Edge2Cloud, Cloud2Edge: new software workflow tricks?
- Explore Urgent Computing with Edge-Cloud
- New AI@Edge Applications
- Integrate small IoT sensors
- More instruments!
- Actuation!
- National AI@Edge Cyberinfrastructure



SAGE
sagecontinuum.org

AI@Edge science problems for students. Get involved!

- Measuring river depth against graduated marker
- Auto-steering of PTZ cameras based on local AI
- Measuring snow depth against graduated marker
- Measuring vegetative states, growth rates
- Self-supervised learning: IR, LiDAR, audio, and RGB
- Vehicle types and flow speeds
- Quantify flower blooming (color, count)
- Outlying conditions from previous sensor data
- Calculating biodiversity based on audio
- Measuring surface water coverage
- Measuring lightning via RF (software defined radios)
- Measuring visibility across a field
- Measuring rime ice thickness
- Measuring ice coverage on a large body of water
- Measuring water flow speed
- Classifying wildlife behaviors
- Improved wildfire detection algorithms
- Wildlife tracking in open fields (speed, direction, count)
- Ultrasonic bat detection
- Measuring pedestrian movement dynamics
- Measuring land changes (riverbeds, plant coverage)
- Measuring water turbidity, debris movement, floating waste
- Measuring vehicle dynamics: identification of sliding, crashes, mishaps
- Measuring bike usage, bike lane dynamics
- Identifying urban "near misses"
- Measuring bird flocks and dynamics

Sage Computing Continuum

Questions?

Join us!

- **Participate in next Hackathon**
- **Deploy nodes, write AI@Edge code**
- **Add SAGE to a proposal**
- **Develop AI algorithms @ Edge!**

Getting started with Sage! - https://docs.sagecontinuum.org

Sage AI@Edge Apps - https://portal.sagecontinuum.org/apps/explore

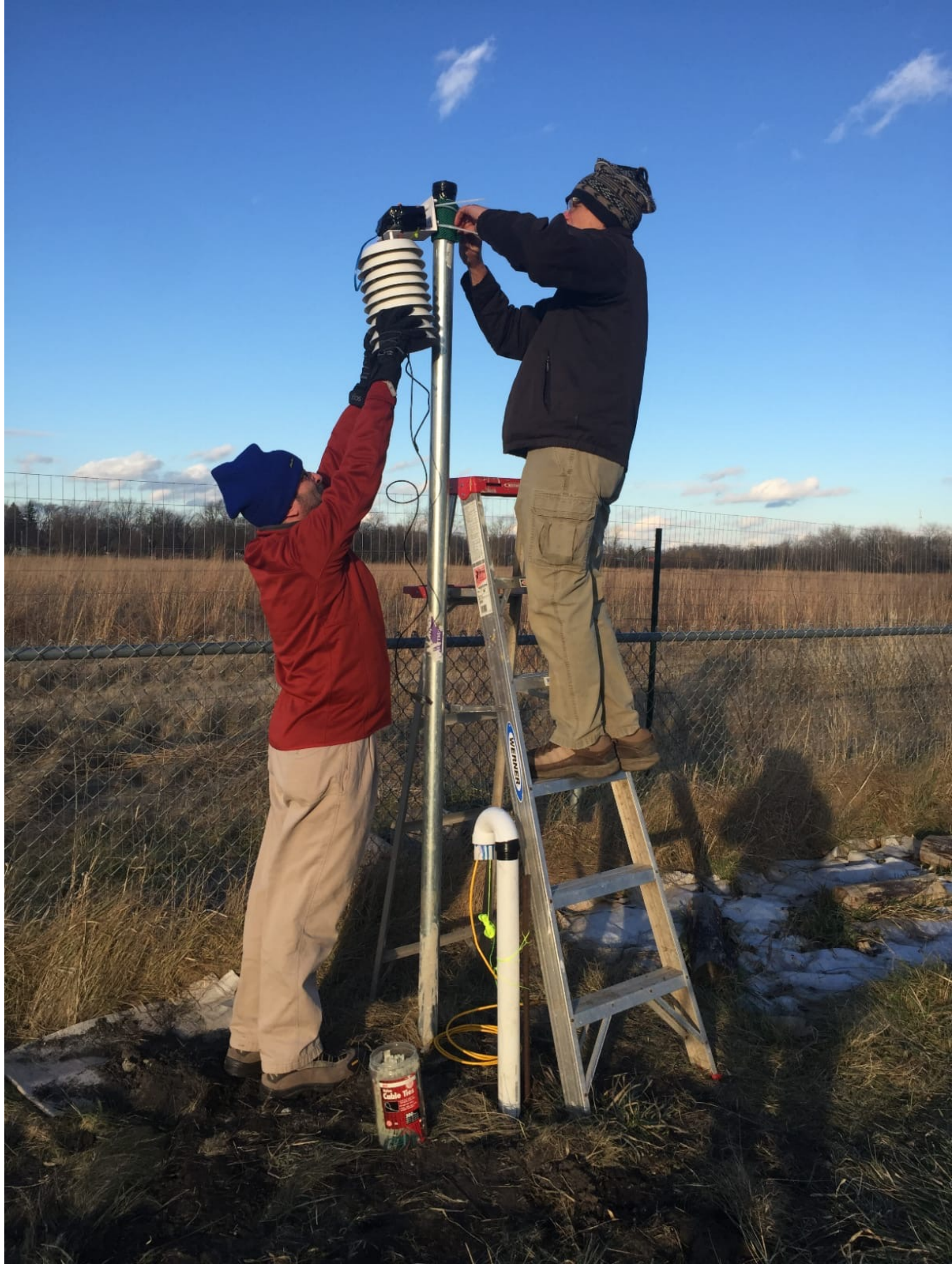
Sage Data - https://portal.sagecontinuum.org/data

Sage Konza MDP Campaign - https://mdp.sagecontinuum.org

Overall Sage system status - https://admin.sagecontinuum.org/status

Waggle Github - https://github.com/waggle-sensor

Sage Continuum Github - https://github.com/waggle-sensor



Professors Aaron Packman and William Miller, Northwestern University
Gensburg-Markham Prairie, The Nature Conservancy
Photo Credits: Liliana Hernandez-Gonzalez, Northwestern University
Dec 2015

SAGE
sagecontinuum.org