# **SCINet Newsletter: January 2024**

Research Spotlight | News | Training | Support | Connect

# **RESEARCH SPOTLIGHT**

# **SCINet Resources for Photogrammetry**

By: Craig D. Woodruff, Computational Biologist, NWRC, Boise, Idaho

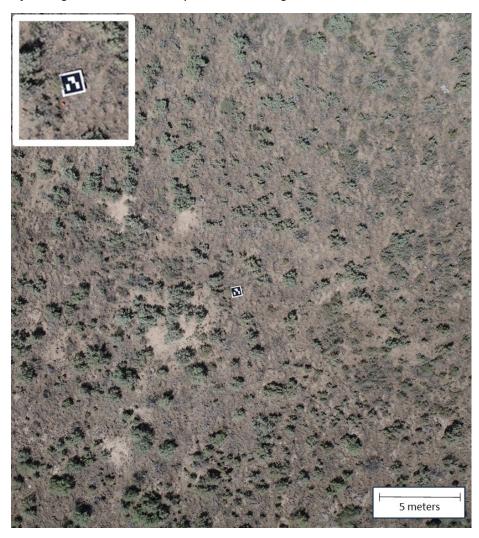


Figure 1: A coded target for automatic ground control point identification is shown in a Wyoming Big Sagebrush-dominated study site in the Reynolds Creek Experimental Watershed. A zoomed-in nested image of the coded target is shown in the upper left corner.

In this Research Spotlight, I would like to highlight SCINet resources for the Unoccupie d Aerial Systems (UAS) research program at the ARS Northwest Watershed Research Center (NWRC) in Boise, Idaho. Scientists at the NWRC use UAS to research long-term vegetation dynamics in sagebrush-dominated rangelands. Multispectral, hyperspectral, and natural color imagery are collected to study the effects of fire, grazing, and invasive weed encroachment. A large amount of UAS data has been collected since 2017, and a streamlined processing workflow that is reproducible on the SCINet high-performance computing (HPC) clusters, Atlas and Ceres, is necessary to consistently process these data. HPC image processing tutorials, supporting code, and reference materials are all available on SCINet's Geospatial Workbook website. These resources have greatly increased the accessibility of the HPC clusters for new and experienced individuals needing to process UAS data. The tutorials not only provide the overview of the necessary steps with examples, code snippets, and animations but they also provide a wealth of knowledge on HPC programming. In the past couple years, the preferred, OCIO-approved UAS processing software has changed to the open-source OpenDroneMap (ODM), and these tutorials have paved the way for implementing this software on SCINet resources.

Processing UAS-collected data can be tricky, and it requires a human component. When we collect the data, we lay out targets referred to as ground control points (GCPs) which help to increase the accuracy of the "stitched together" final products. The area of interest may have a large elevational gradient making the GCPs all the more important to derive accurate orthomosaics, digital surface models, and digital elevation models. Before we process the imagery, someone is required to geotag the GCPs in the images with their respective global positioning system (GPS) coordinate values. This step virtually anchors the images to the ground surface and allows the ODM software to combine the hundreds of images into a single product. (Please see the article in the "Training" section below for a summary of a recent SCINet UAS workshop about this process.) Once this is done, any SCINet user can implement the tutorial on processing imagery on Atlas. But the big time sink and potential source of human error is the GCP tagging step, and if it were automated, the processing pipeline would be completely reproducible. One of the newest advances in UAS data collection is coded targets that allow the GCP geotagging step to be automated.

Coded targets, as seen in Figure 1, are printed QR code-style GCPs. When setting these targets out before collecting UAS data, you simply need to record which coded target is at which GCP location. This information is then fed into the processing pipeline, and the GCPs are automatically detected. Aleksandra Badaczewska-Dawid, with the SCINet Virtual Research Support Core (VRSC), and Pat Clark, a Range Scientist at the NWRC, have worked together to develop the tutorial, code, and real-world examples of implementing coded target detection on SCINet's HPC clusters. This suite of supporting documentation on coded targets fills the final gap in the processing workflow by automating the manual, time-intensive portion of image processing.

### **SCINet and AI COE Fellows**

Please welcome our newest SCINet and ARS Artificial Intelligence Center of Excellence (AI-COE) fellows!



**Dr. Rishi Bhandari** recently joined Dr. David Kang's Insect Microbiome Lab at the USDA-ARS Biological Control of Insect Microbiome Laboratory in Columbia, Missouri. He earned a B.S. in Agriculture from Tribhuvan University in Nepal and an M.Sc. and Ph.D. in Plant Pathology from Auburn University in Alabama. His doctoral research focused on utilizing metagenomics and machine learning to study phyllosphere microbial communities in tomato and pepper, specifically bacterial leaf spot disease and the impact of climate change. His work also focuses on comparative genomics, assessing non-pathogenic *Xanthomonas*' role in virulence traits and disease development.

Dr. Bhandari's research in the Insect Microbiome Lab focuses on exploring the microbiome's role in insect pest control, emphasizing environmental influences, within-community interactions, and their effects on host immunity, metabolism, reproduction, and survival. Using a multidisciplinary approach, he will combine machine learning and statistical modeling techniques to investigate how the microbial community assembly negatively impacts the pest performance of *Drosophila*.



Dr. Annie Hatmaker earned her Ph.D. in Biological Sciences from Vanderbilt University. She was awarded a National Institutes of Health National Research Service Award predoctoral fellowship to complete her dissertation research, which focused on the evolution of virulence within the fungal pathogen *Aspergillus flavus* and related species. Prior to her Ph.D., Dr. Hatmaker completed her B.S. at the University of Tennessee, Knoxville, in Ecology and Evolutionary Biology. She then earned her M.Sc. in Genomics and Bioinformatics from the University of Tennessee Institute of Agriculture's Department of Entomology and Plant Pathology. After her M.Sc., she worked as an ORISE ASTRO Fellow and Post-Master's

Research Associate at Oak Ridge National Laboratory in the Microbial Ecology and Physiology group.

As a SCINet fellow, she will study members of the fungal genus *Fusarium*, with an emphasis on species responsible for Fusarium head blight, branching into machine learning to examine regulatory elements of fungal genomes using large genomic and transcriptomic datasets. She is mentored by Dr. Milton Drott at the Cereal Disease Laboratory in St. Paul, Minnesota.



Dr. Joyce Marumo earned her Ph.D. in Zoology from the University of Aberdeen in Scotland, United Kingdom, and holds a B.Sc. (*Cum laude*) and a M.Sc. degree in Agriculture-Animal Production from the University of Limpopo in South Africa. Her doctoral research focused on the environmental and social factors influencing milk production in Scottish dairy cows. In 2021, she joined Cornell University, New York, working in the Dairy Cattle Nutrition lab as a Postdoctoral Associate alongside Professor Mike Van Amburgh. At Cornell, her responsibilities included updating and modifying the Cornell Net Carbohydrate and Protein System (CNCPS). This work aimed to quantify, manage, and potentially mitigate

greenhouse gas emissions from dairy cattle. She has authored and co-authored papers in the areas of dairy cattle behavior and physiology, and nutrition.

Dr. Marumo has experience in teaching, research, and managerial roles. She holds the title of Associate Fellow in Teaching in Higher Education accredited by the Higher Education Academy in the UK. Additionally, she is a Postdoctoral member of the American Dairy Science Association, a Professional Natural Scientist registered with the South African Council for Natural Scientific Professions (SACNASP), a member of the National Postdoctoral Association (NPA), and a member of the International Organization of Women in Science for the Developing World (OWSD).

Currently, Dr. Marumo has joined SCINet/AI-COE as a Postdoctoral Fellow and she is grateful to collaborate with ARS molecular biologist Dr. Paula Chen. Dr. Marumo's research focuses on developing a cross-kingdom genome editing toolkit using AI. She plans to develop her skills and knowledge in using machine learning applications and AI methods with gene editing tools. She also looks forward to collaborating with other ARS scientists.



Dr. Ronnie Serfa Juan is an ASEAN (Association of Southeast Asian Nations) Electronics engineer with 15 years in academia in the Philippines. Holding a Ph.D. in Electronic Engineering from Cheongju University in South Korea, his expertise spans AI, machine learning, thermography, and image processing. During his Ph.D., he led research on a "Smart Automotive Advanced Driver Assistance System," a cutting-edge driver assistance technology project, and his contributions extend to computer systems, networking, structural engineering monitoring, and photovoltaic applications.

Currently, Dr. Serfa Juan is an SCINet/AI-COE fellow under the mentorship of Paul S. Armstrong, and his work will concentrate on insect detection and classification using image processing and AI implementation. Moreover, he aims to harness the power of the Internet of Things (IoT) to provide real-time solutions for detecting and addressing insect infestations in grain or other storage facilities. This innovative approach, combining the strengths of IoT and AI, aims to revolutionize precision pest management in agricultural settings.

#### **NEWS**



We are pleased to welcome Dr. Lavida Rogers, Ph.D., our new Training Coordinator at SCINet. Lavida holds a B.S. in Biology from the University of the Virgin Islands and a Ph.D. in Microbiology and Molecular Genetics from Michigan State University. She is a trained computational biologist and an experienced educator. Before joining the SCINet Office, Dr. Rogers taught upper-level biology courses and provided guidance and mentorship to undergraduate students in both research and academics at the University of the Virgin

Islands. Additionally, she served as a program coordinator for a summer training initiative tailored for K-12 teachers in the Virgin Islands and for an NIH-funded training grant for undergraduates.

In her role as Training Coordinator, Dr. Rogers will develop and facilitate SCINet training activities. With proficiency in programming languages such as R, Python, Bash, and Mathematica as well as teaching and curriculum development experience, she is well-equipped to enhance our training programs.

### **Apply to Serve on the Scientific Advisory Committee!**

The SCINet Scientific Advisory Committee (SAC) is seeking representatives from the Northeast Area (2 positions), Pacific West Area (1 position), and Statistics (1530 series) (1 position) to serve for 3-year terms beginning March 2024.

The SAC helps ensure that SCINet meets the scientific computing needs of ARS researchers. SAC members represent a broad range of scientific research at ARS and membership includes two researchers from each of the five ARS geographic areas, a statistician, and an "at-large" member. The SAC meets monthly to provide input on planning, education, and communication activities of the SCINet Initiative. For example, SAC members help organize instructional workshops, maintain a body of scientific points-of-contact across ARS research units, and help identify unmet scientific computing needs in ARS. There is also opportunity for input on policy initiatives. SAC progress and issues are elevated to the Executive Committee for review and approval.

Please consider sharing your time and talents by serving on the SAC. These are 3-year terms which begin Spring 2024 and will end Spring 2027. <u>Please send your nominations</u> (including self-nominations) to Kathleen Yeater (Kathleen.yeater@usda.gov) by COB on <u>Friday February 16, 2024.</u> Please provide e-mail contact information of nominees. Self-nominations are encouraged; if you self-nominate, please provide a paragraph summarizing your desire to serve on the committee.

### **Internships Update**

Many thanks to the 61 ARS scientists who applied to serve as Al-COE/SCINet graduate student internship mentors in 2024!

We plan to support 30 interns in 2024, and our university partners are working right now to recruit student participants. After our university partners select the students who will participate in this year's program, we will match each student with an internship opportunity based on the mutual interests of the student and their prospective mentor.

We expect to announce intern/mentor pairings in early March. Again, thank you for your support of the AI-COE/SCINet graduate student internships program!

#### **TRAINING**

### **Training Opportunities**



Getting Started: With the expansive list of free training available online, finding the right training to meet your learning needs can be daunting. Take the first steps in getting started with the SCINet Introductory Learning Pathway. Learn about SCINet, how to sign up for an account, and what is possible when supported by SCINet infrastructure. Then dive in with hands-on tutorials available across multiple searchable platforms to find the information you need for just in time learning.

**Practicum AI:** Developed and presented by the University of Florida and customized for USDA-ARS with funding from ARS's Artificial Intelligence (AI) Center of Excellence, *Practicum AI* is a hands-on, applied artificial intelligence curriculum intended for learners with limited coding and math background. Using hands-on exercises and graphically based, conceptual content, the program starts from introductory content and builds your AI knowledge, enabling you to design and conduct AI work. Our initial offering of these courses wrapped up on January 10, but if there is sufficient interest, we will offer these courses again in the future. To join the waitlist for future offerings of these courses, please <u>fill out the registration form</u>. There is no cost for ARS researchers to participate in this program.

For more information about the courses, please see the <u>SCINet website</u>.

Coursera.org Courses: The SCINet Office and the AI-COE are excited to provide training opportunities through Coursera. Coursera licenses are available to ARS scientists and support staff for training focused on scientific computing, data science, artificial intelligence, and related topics. Successful completion of courses and specializations result in widely recognized certificates and credentials. Please visit the SCINet Coursera Training Page to request a license. Licenses will be assigned on a rolling basis and are active for three months. Users may be able to extend their licenses upon request.

The Carpentries Instructor Training: SCINet is collaborating with The Carpentries to offer The Carpentries' Instructor Training Course for ARS scientists. In this course, you will learn about evidence-based practices for effective and inclusive teaching, with a particular focus on teaching computational skills. There is no fee charged to course participants, but seats are limited. If you are interested in becoming a Carpentries certified instructor, please complete this form.

Please help us improve our training offerings! What scientific computing training do you need? The SCINet Office's goal is to provide training opportunities and resources that meet the needs of ARS researchers, so we would be grateful if you could complete a short form and let us know how we can best help you learn the computing skills you need. Your feedback will help us decide where we should focus our efforts over the next year and beyond.

Training opportunities are continually being updated on the <u>SCINet Upcoming Events</u> <u>webpage</u>. For more information on any of the above trainings, registration questions, or suggestions, please <u>email SCINet-training@usda.gov</u>.

### **Workshop Reports**

#### **Unoccupied Aerial System (UAS) Workshop:**

Andrea Albright (SCINet Postdoctoral Fellow, ARS Southeast Watershed Research Laboratory) and Aleksandra Badaczewska-Dawid (SCINet Virtual Research Support Core (VRSC)), developed and led a workshop focused on processing UAS image data using the open-source software OpenDroneMap (ODM) on SCINet's supercomputer, Ceres.

The workshop was held November 7-9, 2023. There were four sessions that described the entire image processing workflow. Session one provided a basic overview of image processing with some details about UAS imagery and introduced the Julia programming language for use on Ceres. Session two demonstrated the process of collecting Ground Control Points (GCPs) and then creating the corresponding GCP file needed for processing. Session three demonstrated submitting a SLURM script using ODM in a container. And finally, session four discussed the results of ODM image processing and potential improvements to the workflow.

Links to the slides and recordings of the workshop can be found at <u>the UAS workshop</u> <u>event page</u>.

#### **SUPPORT**

## Getting Started with SCINet Is As Easy as 1,2,3

If you do not already have a SCINet account, we hope you will consider joining the 2,300+ researchers who do. Follow the steps below to get your SCINet account.



- 1. Request a SCINet account to get started.
- 2. Read the <u>SCINet FAQs</u> covering general info, accounts/login, software, storage, data transfer, support/policy/O&M, parallel computing, and technical issues.
- 3. Register for a <u>SCINet Forum</u> account to connect to other users, ask questions, and learn how SCINet can enable your research.
- P.S. Don't forget to complete your annual security training! This is required to maintain your account.

For technical assistance with your SCINet account, please email <a href="mailto:scinet\_vrsc@usda.gov">scinet\_vrsc@usda.gov</a>.

### Support email addresses

All requests for help with user accounts, login problems, resource requests, or support for the Ceres HPC cluster should be sent to the SCINet Virtual Research Support Core (VRSC) at <a href="mailto:scinet\_vrsc@usda.gov">scinet\_vrsc@usda.gov</a>. Help requests specific to the Atlas HPC cluster should be sent to <a href="mailto:help-usda@hpc.msstate.edu">help-usda@hpc.msstate.edu</a>.

Many emails are currently being sent to other SCINet email boxes. For the most expedient response to your support requests, be sure to send them to <a href="mailto:scinet\_vrsc@usda.gov">scinet\_vrsc@usda.gov</a> or to <a href="mailto:help-usda@hpc.msstate.edu">help-usda@hpc.msstate.edu</a> for Atlas-specific requests.

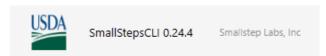
### SCINet User Tip: Streamlined SSH Access to SCINet

Logging on to SCINet's supercomputers via SSH is now easier than ever thanks to a new software package available in the CEC Software Center!

**SmallStepsCLI** is software that supports LincPass authentication for SSH, and you can now install it directly from the Software Center. After installing SmallStepsCLI and completing some initial configuration steps, you will be able to use your LincPass card or YubiKey to log on to SCINet's supercompters via SSH.

Please also note that Software Center installation of SmallStepsCLI is currently only available for Windows users. If you are a Mac user, please see <u>additional instructions on the SCINet website</u>.

Name Publisher



#### Installing SmallStepsCLI for Windows

- 1. Open the Software Center application and type "smallstep" into the search box in the upper-right corner.
- 2. The search will return **SmallStepsCLI 0.24.4** (or similar). Click the **SmallStepsCLI** application.
- On the following screen, an Install button should appear. Click Install.
  If you only see "Uninstall" and "Repair" as available buttons, SmallSteps has already been installed on your computer.
- 4. The installation will take 5-10 minutes to complete. Once completed, the status of the application will show as "Installed".
- 5. Open a PowerShell window and run the following commands, replacing "user.name" with your own SCINet username:
  - step ca bootstrap --ca-url https://step-ca.scinet.usda.gov --fingerprint adb703fd18f176937743b20228d52af7a705d63a0471cd67428660be5fd006bf
  - step ssh config --set Provisioner=keycloak --set User=user.name

Please see the <u>SCINet website</u> for more information and troubleshooting help.

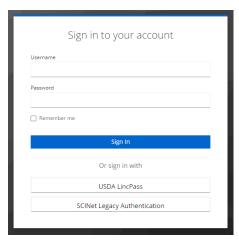
### **Usage Instructions:**

In PowerShell, if you are connecting to Ceres, run ssh user.name@ceres.scinet.usda.gov

Or, if you are connecting to Atlas, run ssh user.name@atlas-login.hpc.msstate.edu

(Change "user.name" to your own SCINet username.)

Your default web browser should open automatically and take you to the SCINet authentication page.



Choose **USDA LincPass** as your sign-in option. If you do not have a LincPass card, you will need to use a YubiKey; please see additional instructions on the SCINet website.

- You will then go through a typical eAuth-based login.
- You will select your applicable option (either USDA or Non-USDA), select login with your PIV/CAC, and enter your PIN.

Go back to your shell and you should see "CA: https://step-ca.scinet.usda.gov" followed by your regular login.

After these steps, command-line *ssh* works normally. The only difference is that it will not prompt you again for authentication for the rest of the day (16 hours).

Do you have tips to share? Email them to <u>SCINet-Office@usda.gov</u> to be included in future newsletters.

#### **SCINet Corner**

SCINet Corner is a VRSC-moderated virtual space for people to share knowledge, discuss best practices, learn about new opportunities, and explore resources to support progress on their projects.

The next SCINet Corner will be held on February 15, 2024, 1-2pm ET. You can register for this and future SCINet Corners <u>here</u>.

Have a question that just can't wait? Want to see what other users are doing? Reach out to the ever-expanding SCINet Forum community for ideas, support, or just someone to bounce ideas off of at <a href="https://forum.scinet.usda.gov/">https://forum.scinet.usda.gov/</a>.

#### CONNECT

### The SCINet Community

Every newsletter highlights SCINet community members as a way to connect the ARS scientific computing community. To see all the SCINet community updates and review past newsletters, visit the Newsletter Archive.

#### Contribute

Do you use SCINet for your research? We would love to share your story! Email <u>SCINet-Office@usda.gov</u> to contribute content, ask questions, or provide feedback on the SCINet newsletter or website.

### **SCINet Leadership Team**

Brian Stucky, Acting Chief Science Information Officer Rob Butler, SCINet Program Manager Jeremy Edwards, Science Advisory Committee (SAC) Chair Jeff Silverstein, Acting Associate Administrator

Note: This newsletter is edited to comply with ARS editorial standards.

**SCINet Website** 

Stay Connected with the USDA Agricultural Research Service 5601 Sunnyside Avenue, Beltsville, MD 20705









