Outlook

**Meeting assets for GeoComp & ML 2025 course are ready!**

**From** Zoom <no-reply@zoom.us>
**Date** Thu 10/2/2025 9:59 AM
**To** Amatulli, Giuseppe <giuseppe.amatulli@yale.edu>

zoom

# Meeting assets for GeoComp & ML 2025 course are ready!

## Meeting summary

### Quick recap

The meeting focused on technical discussions about transitioning fire weather calculations in Alaska and implementing various data processing tools like GDAL, GRASS, and PK Tools for handling TIFF and NetCDF files. Participants addressed performance issues with virtual machines and discussed the advantages of using Ubuntu over Windows for Bash and Linux-based development. The group explored various tools and techniques for handling and analyzing spatial data, including rasterization, masking operations, and quality control layers in satellite images, with Giuseppe providing demonstrations and explanations of different functionalities.

### Next steps

- Giuseppe to post two options for next week's classes on Discord: either a full week of GRASS or keeping the current plan .
- All students to vote on Discord for their preference between the two class options for next week.
- All students to prepare their datasets for the upcoming catch-up session to apply knowledge from classes and ask questions.
- Quinn to try re-scripting something from R using GDAL and PK tools.
- Kula to close at least two terminal processes to free up RAM on her virtual machine.
- Nafeesa to start applying basic statistics and resolution changes on her downloaded TIFF file.

### Summary

#### Alaska Fire Weather Script Transition

Quinn discussed transitioning scripts for fire weather calculations in Alaska from Terra to GDAL, noting that while GDAL would handle data integration better, some calculations might need to be rewritten. Giuseppe explained that zonal statistics and GIS functions could be more straightforward

in GRASS, which he would be teaching the following week. The conversation ended with 14 participants, and no major issues were reported.

**Optimizing CPU and Memory Usage**

Kula encountered performance issues while running GDAL commands on a virtual machine and sought help from Giuseppe. They identified that Kula was requesting 5 CPU processes but was only allocating 1 CPU, causing slow performance and high memory usage. Giuseppe advised Kula to close unnecessary terminal processes, set the virtual machine to use 3 CPUs, and run only one terminal at a time for better performance. He also recommended using tools like htop to monitor system resources and avoid saturating RAM.

**Python and Bash Data Processing**

The meeting focused on technical discussions about using Python and Bash in data processing, particularly for handling TIFF and NetCDF files. Giuseppe explained the importance of creating Python environments for managing different library versions and demonstrated how to install and use CDO for climate data processing. Nafeesa raised questions about using Python scripts in Jupyter and running Bash commands on Windows, which led to a discussion about the advantages of using Ubuntu over Windows for Bash and Linux-based development due to better integration and support for high-performance computing.

**PK Tools Development and Functionality**

Giuseppe discussed the development and functionality of PK Tools, highlighting its similarities and differences with GDELT, particularly in raster preprocessing and filtering capabilities. He explained the use of PK Tools for tasks like aggregation, filtering, and reclassification, emphasizing its efficiency and versatility. Katarina inquired about hexagonal analysis and nearest neighbor operations, to which Giuseppe clarified that while PK Tools does not support these directly, they can be addressed using other software or libraries.

**Digital Calculation and Raster Analysis**

The meeting focused on various aspects of digital calculation and raster analysis, including predefined functions like median, mean, and deviation, as well as zonal statistics for vector and raster data. Giuseppe explained the use of PK tools for these calculations and mentioned that atmospheric correction and unmixing analysis are more specialized topics typically handled by remote sensing-specific software. The group discussed practical exercises involving masking operations, including thresholding and creating masks for data analysis, and Giuseppe demonstrated how to perform these operations directly in binary format using GDAL commands.

**Satellite Image Quality Control Techniques**

Giuseppe explained the concept of quality control layers in satellite images, which provide information about pixel quality and can be used for masking operations. He demonstrated how to threshold and mask images using different values, and emphasized the importance of understanding the difference between get and set commands in GIS software. Sofia asked about using a vector file to mask the borders of a country, to which Giuseppe replied that a vector file needs to be converted to raster format for masking purposes, and suggested using tools like GDAL rasterize or PK raster for this conversion.

**QGIS Rasterization and Masking Techniques**

Giuseppe explained how to perform rasterization and masking operations in QGIS, emphasizing the importance of understanding masking concepts for various applications. He demonstrated a composite operation using PK composite, showing how to calculate means, standard deviations, and other statistics while applying masks to handle overlapping and no-data areas. Giuseppe highlighted the flexibility and efficiency of PK composite for creating global composites and thresholding, noting its advantages over other tools like GeoCalc.

**Spatial Data Analysis Tools Overview**

Giuseppe discussed various tools and techniques for handling and analyzing spatial data, focusing on PK tools and GDAL. He explained how to perform operations such as compositing, filtering, reclassifying, and calculating statistics on raster data. Giuseppe emphasized the importance of understanding the tools' capabilities and limitations, particularly for large datasets. He also touched on using scripting languages like Bash and Python for automation and workflow organization. Giuseppe encouraged participants to experiment with the tools and explore advanced functionalities.

AI can make mistakes. Review for accuracy.

Please rate the accuracy of this summary.

Edit summary Share

---

Thank you,

Zoom Support Team
https://support.zoom.us

Zoom.com
55 Almaden Blvd
San Jose, CA 95113

+1.888.799.9666