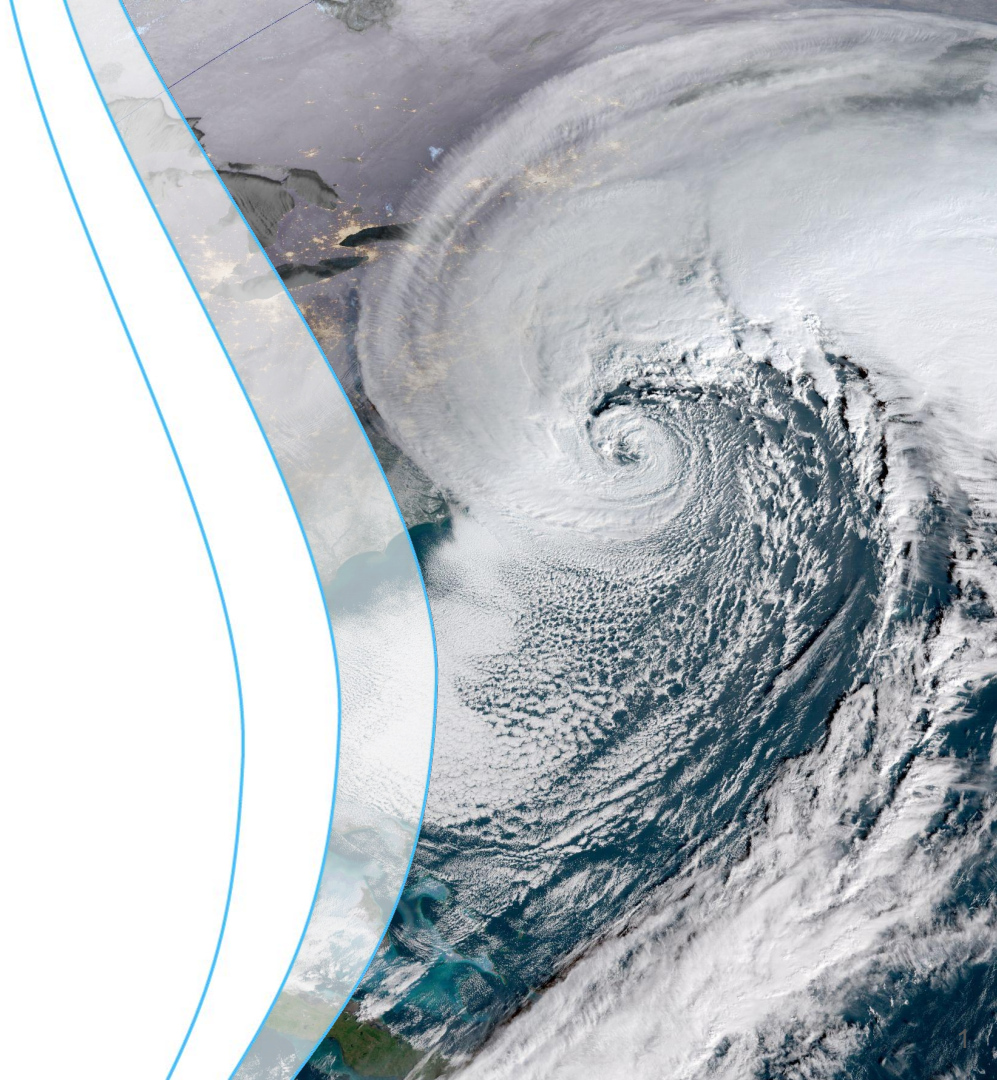

MARCH 5, 2026

AI Assistants in ACOM Software Development

KATE THAYER-CALDER
CGD/PPC at NSF NCAR

Material notes duis aute irure in reprehenderit velit esse dolor ex commodo magna lorem sit ipsum dolor nonummy aliquip cilum ut labore et dolore. Mattis elementum gravida eget consequat imperdiet sed. Mi erat hac urna facilisis facilisis.

A few quick stories....

- Kyle and AI - the good!
- Kyle and AI - not so good
- David and Victor (CheMPAS-A) - multiple Claude AI agents
- Kate and Claude - early days

Kyle and AI Agents: Positive Notes

- ➔ Kyle has been using Copilot for a while, and uses it for FORTRAN, C, C++, Julia, Javascript and markdown (documentation)
- ➔ His words: “There are some tasks where it can do all of the work, or the majority with only minor changes by a human (me). These tend to be very small, and focused without ambiguity.”

Reorganize READMEs for Python, JavaScript, and Fortran interfaces
https://github.com/NCAR/musica/pull/726

- Four markup files changed with documentation and code examples for three different languages
- Four copilot commits and one from Kyle
- Much less time required than writing this documentation by hand.

Add Julia formatting to CI workflow
https://github.com/NCAR/musica/pull/752

- Yaml and json files changed to update the workflow steps.
- One commit from copilot and two small changes in another commit from Kyle.
- This PR conversation includes the original prompt, which was a pretty simple outline of acceptance criteria and idea.

Expose solver parameters through all language bindings
https://github.com/NCAR/musica/pull/783

- This one done with Claude
- Larger PR with 23 files changed and about 1000 lines of code added. Kyle made a one line change by hand.
- Fortran, Js, and C++ code and header files all included in one commit.
- Claude reviewed its own PR and gave the code 15 comments that it (mostly) fixed itself.

Kyle and AI Agents: Negative Notes

➔ Some more complex tasks were started by the agents but Kyle ended up taking over.
➔ Kyle: “The commit history (when I didn't rewrite the history with a rebase) can be telling”

Refactor conditions system with unified DataFrame-based API

https://github.com/NCAR/music-box/pull/413

- Currently 48 files changed and ended up being a rather major refactor.
- PR is still in review and spinning off other issues.
- Tried to break it down into smaller tasks that copilot could handle more easily but gave up after two of those.

Type erasure

https://github.com/NCAR/musica/pull/771

- PR Description: “Much of this written with claude and copilot, though much had to be corrected by hand still”
- 56 files changed and 35 commits.
- A lot of the commits are Kyle saying “correcting nvidia dockerfile” or “correcting map type” or “micm conditions correction”

Add initial Julia wrapper for MUSICA

https://github.com/NCAR/musica/pull/736

- Should have been a (relatively) small PR that just “Adds a minimal Julia package and CxxWrap C++ bindings exposing `get_version()`”
- Comments on the PR from Kyle: “@copilot there's a build error in the github actions file”, “@copilot everything built, but the tests don't pass”, “@copilot the library still cannot be found”

CheMPAS-A with Multiple Claude AI Agents

David Filmore and Victor Weeks are both using multiple Claude AI agents to write code that interfaces with MPAS atmosphere.

- David described a workflow with three competing and antagonistic agents competing with each other to write the best solution to a problem.
- Victor used three agents with different roles to develop code for emissions handling in MPAS.
  - The three agents are a code architect, a tester (quality control), and a developer.
  - They review each other's proposals and have implemented the new feature but it hasn't been integrated or tested yet.

My few projects with Claude…

- CI Compass intern last summer named Tamera Segal who worked with Claude (last year) to convert a bash-Fortran-ncl project into python
- She had the best luck "role-playing" with the AI, using prompts like "Now pretend you are an extremely skilled software engineer with expertise in converting scripts from one programming language to another; you know how to take into account all the nuances of each language and make sure that no functionality or precison or components are lost during the conversion process. Create a plan (and show me that plan) for how you will go about approaching converting the script provided in this prompt into python."
- In the end, it got something that completed the task without errors, but we are not sure if it did it correctly yet.

My few projects with Claude…

- Recently, I did a review of a massive, complicated PR for CISM (the reproducible sums with changing processor counts PR). I spent a full day (8 hours) reading through the code changes, and I glazed over some of the worst files, but found a few things that would be nice to change for better performance.
- Out of curiosity, I asked Claude to review the PR. And it came up with 5 suggestions that were very similar to mine! :)
- Even with all of that, we are getting runtime errors in testing. I spent a little time trying to debug but decided to ask Claude if it could find the bug, and it found a missing halo update.
- However, the error persisted, so I'll have to go back and ask again? Or find it myself….

In Summary - Please Snow