

## Testing approach for porting 4-mode Modal Aerosol Model (MAM4) to C++/Kokkos

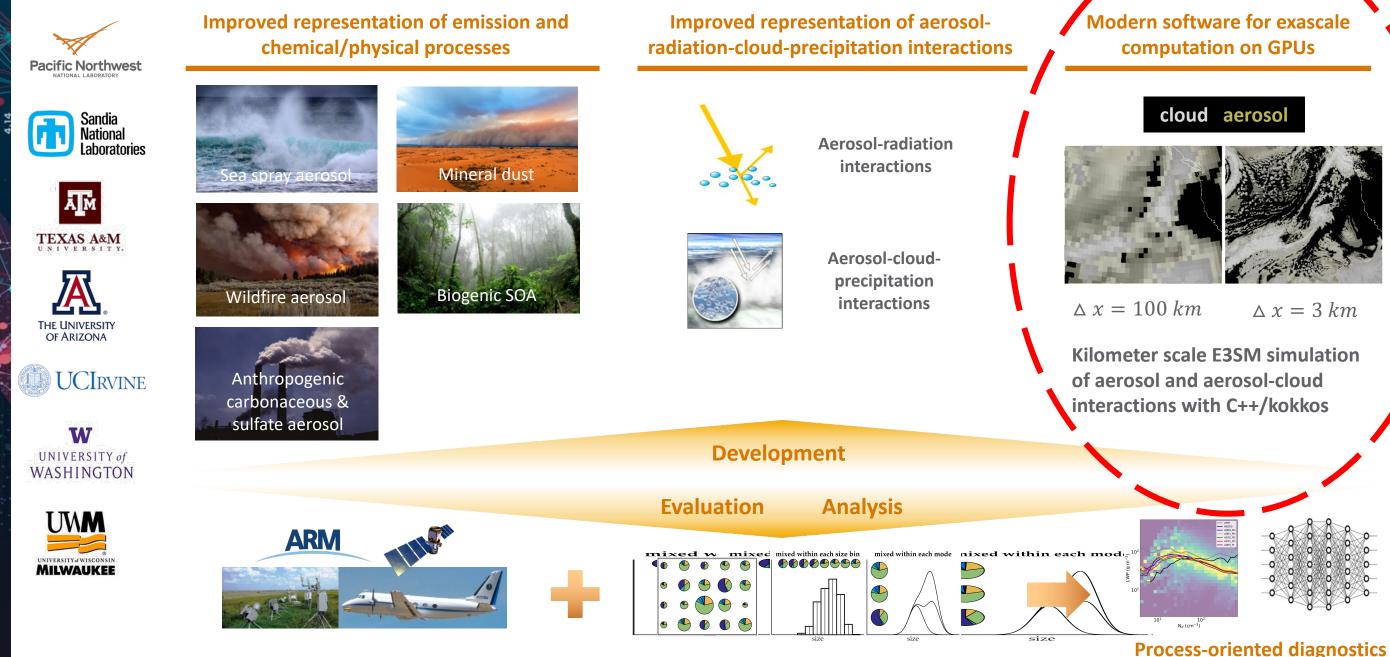
Balwinder Singh, Jeffrey Johnson, Kyle Pressel, Oscar Diaz-Ibarra, Shuaiqi Tang, Michael Schmidt, Brian Gaudet, James Overfelt, Mingxuan Wu, Hui Wan, Jaelyn Litzinger, Peter Bosler, Cameron Rutherford, Jerome Fast, Po-Lun Ma



PNNL is operated by Battelle for the U.S. Department of Energy

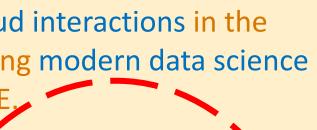


Achieve unprecedented realism in predictions of the role of aerosol and aerosol-cloud interactions in the Earth system, through the combination of advancing fundamental science and utilizing modern data science and computational methods, to address critical challenges facing the nation and DOE



Measurements

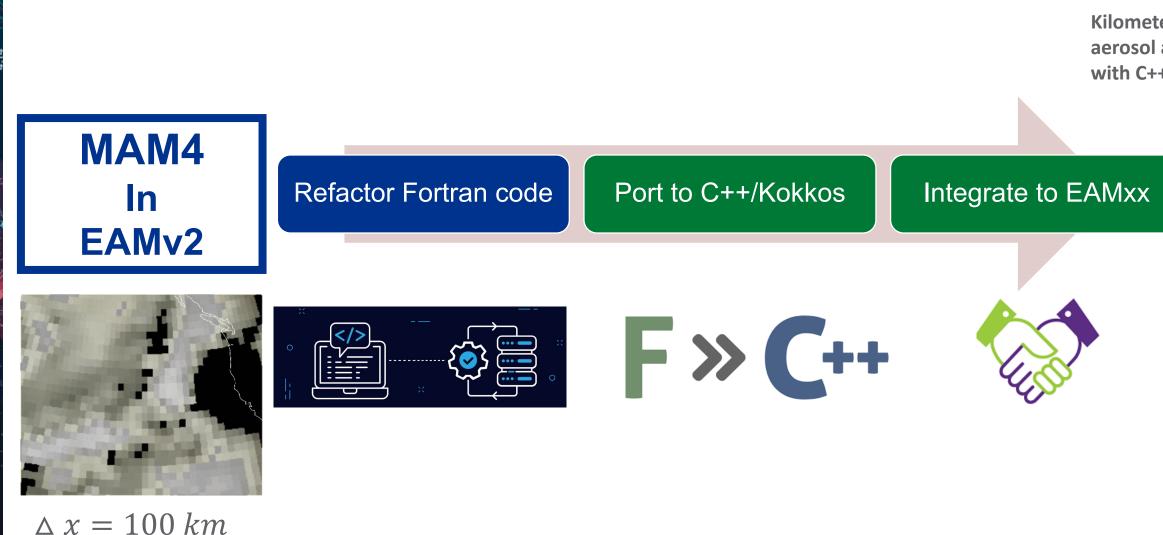
**Process model & large eddy simulations** 



Data informed parameterization



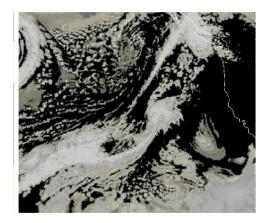
## Modern software for exascale computation on GPUs





Kilometer scale E3SM simulation of aerosol and aerosol-cloud interactions with C++/kokkos





 $\Delta x = 3 \, km$ 

C++/Kokkos



## **Refactoring MAM4 Fortran**

### Long Subroutine — Multiple short routines

- Ideally, a low-level subroutine should do only ONE task
- Logically arrange the calls clean, simplified, and easy-to-read code
- Re-write to reuse code maintainable code
- Remove dead and untested code:
  - Number of lines removed
    - Final (including boilerplate) initial = 100K+ lines removed

### Philosophy

- "True" refactor and port: Identify but **don't** fix bugs
- Remove dead and non-MAM4 codes

### **Results stay BFB – Test-driven end-to-end refactor**



# **Refactor – protected by testing**

### Before touching the code, set up a test...

• Fast-running CIME smoke test with baselines

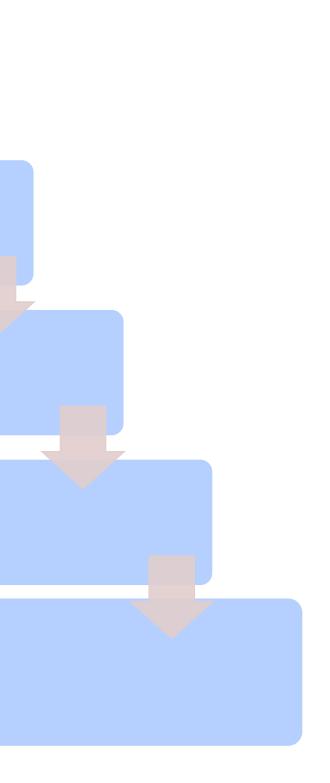
### Refactor in small chunks and run the test often

• Quick compilation and runs

Use Git frequently to check in small changes

### After refactoring: Run a suite of tests

• Protect against any non-BFB changes

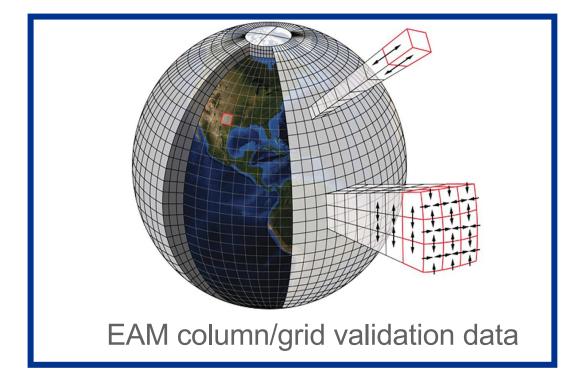


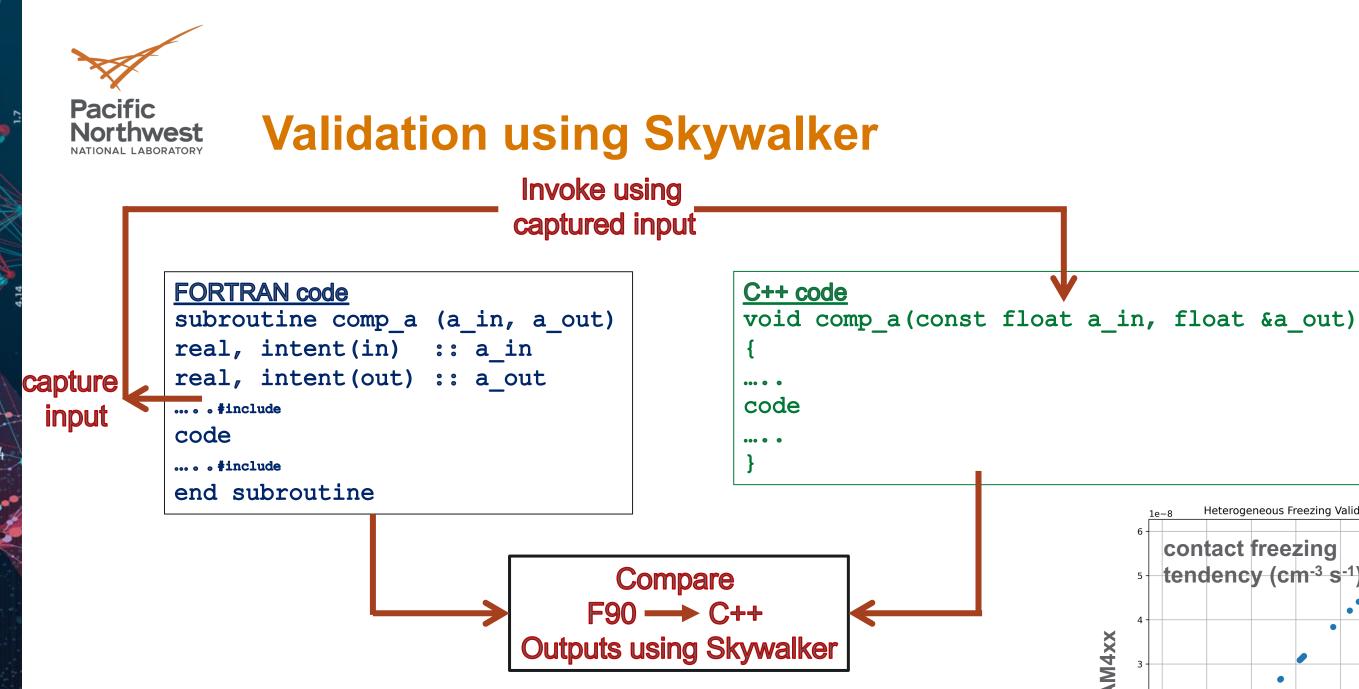


# Validation data for C++ port

Validate F90 ---- C++ port:

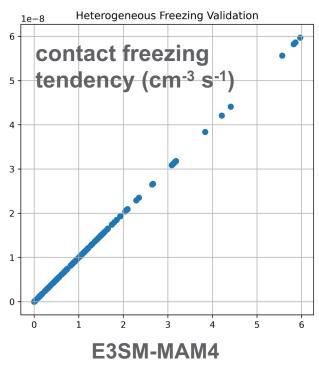
- New infrastructure for generating test data
  - Extract column data from E3SMv2
  - Store in **Skywalker** format
  - Test ported C++ routine with this data
- Flexible and extensible
  - Can extract massive amounts of data
  - Data at any column, time step, and level of the model
  - Easy to use with automated code generation





Heavily relies on the **Skywalker** tool for input and output formats and comparisons

MAM4xx

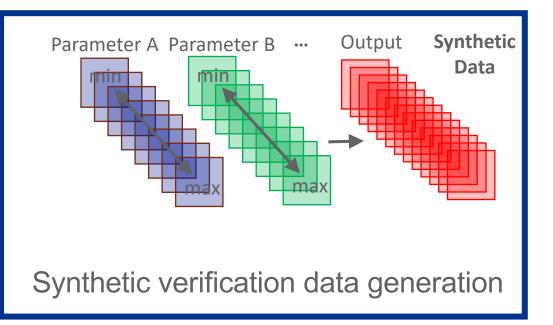




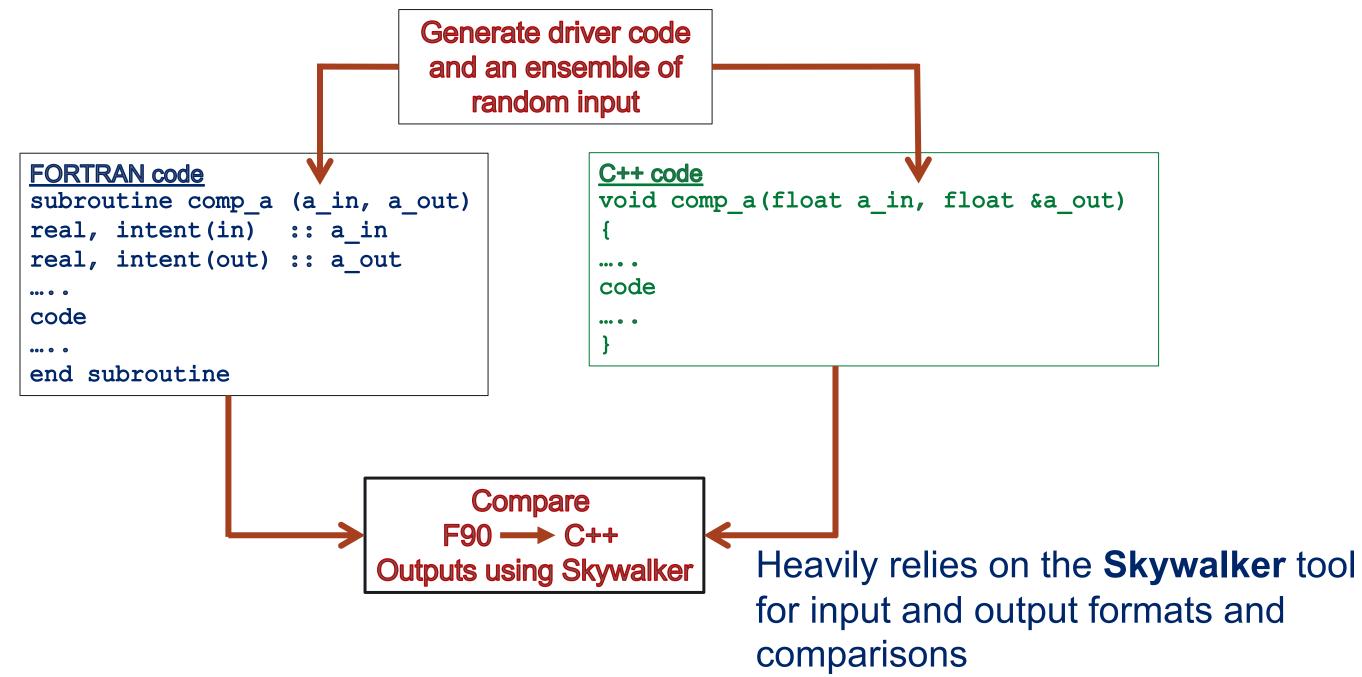
# Synthetic data tests with Skywalker

### Synthetic Ensemble data

- For complete code coverage, invoke Fortran and C++ codes with the same Skywalker synthetic data ensembles
- Compare the outputs to ensure a correct port
- Easy to use with automated code generation
- A great debugging tool!



## Validation using Skywalker



Pacific

Northwest



# **CI/CD** pipelines for automated testing

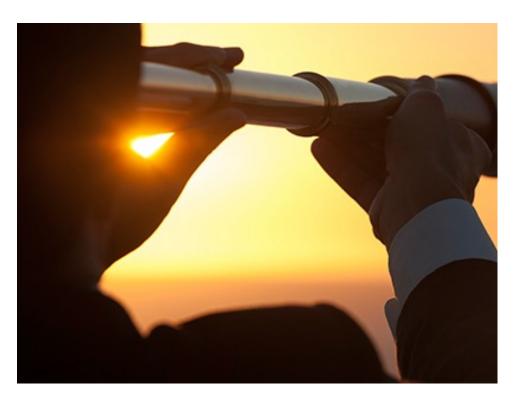
### Fortran refactored codes:

- Nightly testing on Compy using Cron
- C++/Kokkos ported MAM4xx:
  - Frequent testing of PRs on CPUs and GPUs using GitHub actions
  - The GPU architectures we are testing on:

GPU Type	Cuda Architecture
P100	60
V100	70
RTX 2080 Ti	75



# Looking ahead...



## MAM4xx in EAMxx:

- E3SMv2 and EAMxx are very different models:
  - ✓ Code architecture
  - ✓ Physics
  - $\checkmark$  No support for low-res configurations
- Testing will rely on physics-based tests and comprehensive model analysis
  - $\checkmark$  Property tests based on the first principles
  - $\checkmark$  Using diagnostics tools to look at various aspects of the aerosol lifecycle, concentrations, and feedback
- Integrating new EAGLES features into MAM4xx