PNNL-32009

Developing and testing capabilities for simulating cases with heterogeneous land/water surfaces in a novel atmospheric large eddy simulation code

September 2021

Kyle G Pressel
Koichi Sakaguchi

Prepared for the U.S. Department of Energy
under Contract DE-AC05-76RL01830

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

Available to DOE and DOE contractors from the
Office of Scientific and Technical Information,
P.O. Box 62, Oak Ridge, TN 37831-0062;
ph: (865) 576-8401
fax: (865) 576-5728
email: reports@adonis.osti.gov

Available to the public from the National Technical Information Service
5301 Shawnee Rd., Alexandria, VA 22312
ph: (800) 553-NTIS (6847)
email: orders@ntis.gov <https://www.ntis.gov/about>
Online ordering: http://www.ntis.gov

PNNL-32009

Developing and testing capabilities for simulating cases with heterogeneous land/water surfaces in a novel atmospheric large eddy simulation code

Final Report

September 2021

Kyle G Pressel
Koichi Sakaguchi

Prepared for
the U.S. Department of Energy
under Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory
Richland, Washington 99354

Abstract

Large eddy simulations (LES) are the primary computational tool used to simulate high Reynolds number three-dimensional turbulent flows. In the context of earth system sciences, particularly atmospheric science, LES are uniquely able to resolve the scales of atmospheric motion that are key for building process-level understanding of boundary layer turbulence, atmosphere-surface interaction, clouds, and cloud-aerosol-chemistry interaction, and are a core limited-area modeling capability. Increasing demands are being placed on LES code bases as growing high performance computing resources allow LES to address a wider range of scientific problems. In addition, LES are emerging as a source of high-quality machine learning training data. These demands necessitate an agile and extensible code base that allows the model to quickly adapt to emergent needs. However, LES have largely relied on legacy Fortran code bases that lack flexibility. A new, Python-based LES capability called Predicting INteractions of Aerosol and Clouds in Large Eddy Simulation (PINACLES) has been developed as part of the Department of Energy's Earth System Model Development (ESMD) program area's Enabling Aerosol-cloud interactions at Global convection-permitting scalES (EAGLES) project. PINACLES was developed from the ground up with a philosophy of maximizing scientific throughput, by attempting to optimize for both model throughput and software extensibility. The initial development of PINACLES delivered a state-of-the-art idealized LES capability solving the non-hydrostatic anelastic equations of motion with doubly periodic boundary conditions and idealized homogenous surface boundary conditions. Here we provide a final report on the outcomes of a fiscal year 2021 Seed Laboratory Directed Research Project that extended PINACLES in two key ways. First, PINACLES was coupled to a state-of-the-art land surface model enabling it to simulate spatially inhomogeneous land-atmosphere interactions that are known to control key atmospheric processes. Second, the dynamical core of PINACLES was modified to permit non-periodic boundary conditions. This model enhancement enables simulation of realistic cases with boundary conditions prescribed from atmospheric reanalysis and enables nested simulations conducted on a hierarchy of computational domains with increasing resolution. Together, these extensions to PINACLES make it a formidable modeling capability and expand its potential application to diverse components of DOE's atmospheric science portfolio.

Background and Motivation

PINACLES is a novel atmospheric LES code that is designed to optimize for scientific throughput. Optimization for scientific throughput (OST) involves implementation of a two-part strategy of maximizing the ability of domain experts to efficiently express new scientific ideas in a model while simultaneously maximizing traditional metrics of model throughput. The first aspect of this optimization is achieved by implementing PINACLES in Python, a language that is widely used within domain and data scientists' daily workflow due to its flexibility and ease of use. This makes it straightforward for domain scientists to customize the model to their scientific needs. The second aspect of the strategy of OST is primarily achieved in two ways. First, of the computationally intensive portions of the PINACLES Python codebase are accelerated by just-in-time compilation using Numba[1]. Second, PINACLES, like many idealized atmospheric LES codes, adopts a non-hydrostatic anelastic equation set that is solved on a horizontally doubly periodic domain[2]. The use of an anelastic system of equations allows the model, at typical LES grid resolutions of tens of meters, to take at least an order of magnitude longer explicit timestep than non-hydrostatic compressible models. This dramatically increases the throughput of anelastic models over compressible models.

While the doubly periodic implementation of PINACLES provides significant utility for both process studies and machine-learning training data generation, the use of doubly periodic domains limits the generality of the model. For example, in nature both the land and sea surface are highly heterogenous and far from periodic at the domain scales typical of LES yet idealized, doubly-periodic LES usually assume that these surfaces are homogenous to avoid the introduction of aphysical discontinuities in the surface boundary fluxes atthe domain's lateral boundaries. However, it is well established that heterogeneities in surface boundaries (for example, from land surface processes) are instrumental in mediating the initiation and development of clouds and cloud systems.[3] Moreover, physically accurate lateral and surface boundary conditions are necessary to facilitate detailed evaluation of LES against observations.

It is within this context that the overall objective of this project was established, namely, to make PINACLES more general, more broadly applicable, and less idealized while minimizing compromises to computational performance. This increase in generality is achieved by first coupling PINACLES to a state-of-the-art land surface model, which introduces surface heterogeneities. To accommodate these surface heterogeneities,and to enable simulation of more realistic cases, the dynamical core of PINACLES is modified to permit fully non-periodic momentum, pressure, and scalar lateral boundaryconditions.

Technical Approach

Coupling of a Land Surface Model into PINACLES

We selected Noah-MP[4] as a state-of-the-art land surface model to be coupled to PINACLES. Our selection of Noah-MP was based on its detailed representation of key land surface processes and its software design, which simplifies the software engineering required to couple its Fortran code base into Python-based PINACLES.

The approach we took was to first construct a standalone Python wrapper for Noah-MP that facilitated rapid testing and software prototyping. This wrapper was developed in its own git repository independent of PINACLES. The wrapper uses the Python C Foreign Function Interface (CFFI) to access C bindings to Noah-MP's Fortran subroutines. The wrapper also implemented Python data structures to facilitate allocation of memory by Python to be used by Noah-MP, and to streamline passing this data between PINACLES and Noah-MP. The advantage of this approach is that it allows arrays allocated in Python by PINACLES to be passed to Noah-MP without any memory copies and with very low overhead.

Once initial testing of the Noah-MP wrapper was completed, the Noah-MP wrapper git repo was made a submodule of the PINACLES git repo and Noah-MP was incorporated into PINACLES' build system. With this complete, PINACLES could now call the Python Noah-MP wrapper.

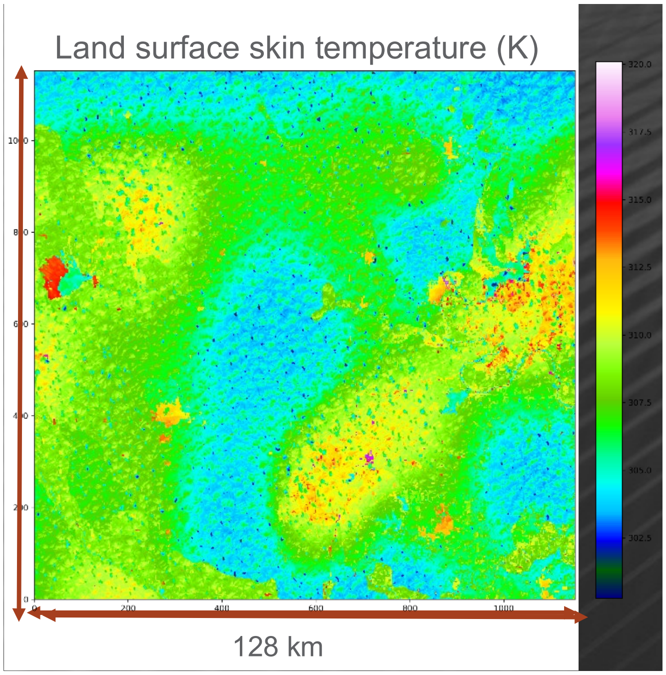


Figure 1: Skin temperature predicted by Noah-MP coupled to PINACLES in a simulation of Hi-SCALE.

Prior to this project, grids generated by PINACLES simply consisted of a regular cartesian grid of points with each grid point associated with a location in cartesian space. However, to incorporate land surface data and land surface initial conditions required by Noah-MP, each PINACLES vertical grid column needs to be associated with a latitude and longitude. To supply this information, we implemented a Lambert Conformal Conical map projection[5] to define a latitude and longitude value for each PINACLES grid point, thus facilitating interpolation of data sets defined on latitude-longitude grids onto PINACLES' cartesian grid. All interpolations are performed using standard scipy[6] interpolation functions, either cubic splines or nearest neighbor interpolation depending on whether the data being interpolated are categorical (e.g., land surface type category) or continuous (e.g., temperature or soil moisture).

With the provision of required external data in place, the remainder of the coupling consisted of providing Noah-MP with its required input from PINACLES' model state, representation of radiative transfer, and cloud microphysical parameterizations. Noah-MP then returns surface fluxes, surface skin temperature, and momentum exchange coefficients that are used to update PINACLES' state.

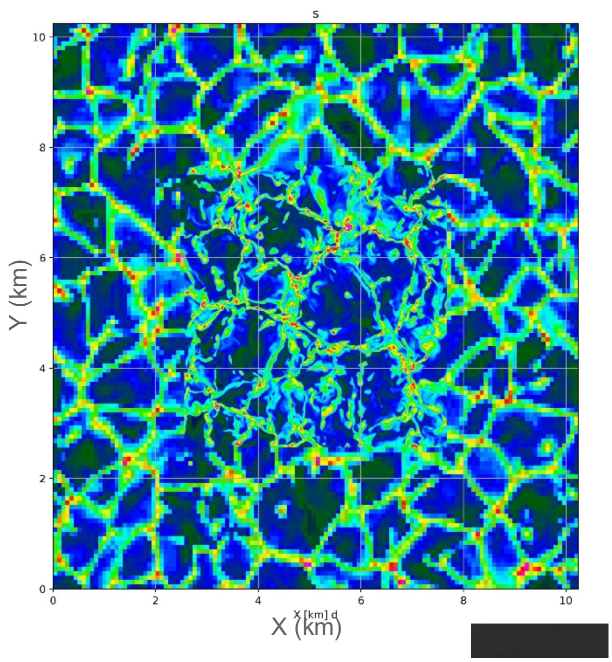
Figure 1 shows an example of a PINACLES simulation of a HI-SCALE[3] case coupled to the Noah-MP land surface model. The figure depicts the land surface skin temperature predicted by Noah-MP. The large-scale variability in skin temperature is largely attributable to variations in surface moisture and the small-scale variability largely attributable to variations in the land surface characteristics, variations in atmospheric temperature due to boundary layer processes, and cloud radiative effect. Most importantly, Figure 1 highlights the large surface heterogeneity that is potentially ignored in idealized models.

Development of non-periodic boundary conditions

The conversion of PINACLES doubly periodic solver to a non-periodic solver involves two key steps. First, components of the model's dynamical core which explicitly assume periodicity must be systematically replaced by components that allow non-periodicity. Second, alternative lateral boundary conditions must be provided.

Accomplishing the first task focused singularly on PINACLES' pressure solver, the dynamical core's most complicated and performance-critical part. The periodic pressure solver used in PINACLES is based on a highly efficient and parallel exact Fast Fourier Transfer (FFT) Method that uses parallel FFTs to reduce the anelastic system's Poisson equation for pressure to a tridiagonal system of linear equations that can be solved locally in the vertical.[2] This solver explicitly assumes that the pressure field, divergence field, and velocity field is periodic. By replacing the FFT based solver with one based on a Discrete Cosine Transform (DCT), like the approach used by the Meso-NH model[7], the Poisson solver's assumption of periodicity was eliminated with minimal reductions in computational performance.

With a non-periodic pressure solver in place, we had to provide lateral boundary conditions for all PINACLES' prognostic variables. The approach we took to specify the boundary conditions follows the approach developed by Davies[8,9] that has been used widely in atmospheric models. That approach prescribes the boundary conditions on the model's lateral boundary ghost points, while optionally relaxing a fringe of domain points proximal to the lateral boundary to specified values to avoid the formation of waves associated with discontinuities at the boundary. We developed two methods for providing the values of prognostic variables to be used at the boundaries by the Davies condition.



[redacted]

Figure 2: Near surface dry static energy in a one-way nested PINACLES simulation of a dry convective boundary layer. There is a 7x grid refinement between the inner and outer grid.

The first approach we implemented was one-way grid nesting in which multiple PINACLES simulations are run simultaneously, with an outer, larger-domain simulation providing lateral boundary conditions for an inner, smaller-domain simulation. The nesting infrastructure is sufficiently general that it allows for a hierarchy of nests running at odd integer multiples of each parent grid resolution. The outermost grid's boundary conditions must be either periodic or prescribed via some other method. Grid nesting is widely used in atmospheric modeling and affords the opportunity to downscale coarse resolution simulations run over a large area to very high resolutions over a limited area. As the lateral boundaries on inner nests are non-periodic, nesting provided an excellent opportunity to test the DCT-based pressure solver and Davies boundary conditions. The grid-nesting capability proved to be very successful with coherent structures cleanly propagating from outer nests to inner nests and with inflow turbulence rapidly generating smaller-scale turbulence on higher resolution nests. For example, Figure 2 shows the near surface dry static energy in a PINACLES simulation of a dry convective boundary layer with one-way grid nesting. The outer nest has a grid resolution of 80 m, and the inner nest has a grid resolution of approximately 11 m (a 7x nest refinement). As the figure shows, the inner domain is highly non-periodic and the coherent structures in the dry static energy field at the nest interface extend continuously across the nest boundary.

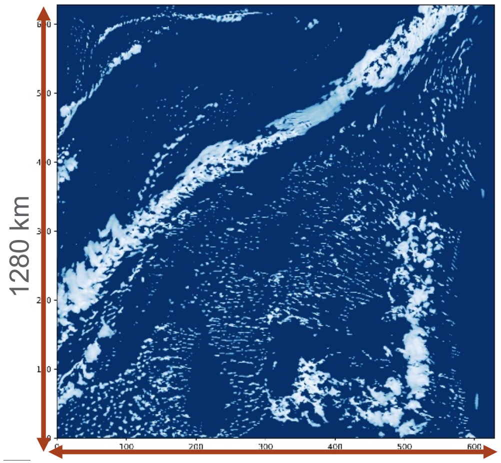


Figure 3: Liquid water path (LWP) in a PINACLES simulation of a real case with lateral boundary conditions driven by reanalysis.

The second approach we implemented was to initialize PINACLES with conditions taken from ECMWF's ERA5 reanalysis[10] and drive the lateral boundary conditions with time-evolving lateral boundary conditions also taken from ERA5. The motivation for this approach is that it enables PINACLES to simulate real cases in a capacity like WRF and affords opportunity to compare PINACLES simulations more directly to observations. Moreover, combining this approach with the nesting approach discussed above permits the performance of simulations that nest down from mesoscale model like resolutions to LES resolutions. We created a parallel online reanalysis data ingestion system that re-grids reanalysis data to PINACLES' grid for the initial and lateral boundary conditions and interpolates the lateral boundary conditions between reanalysis data output times as the model runs, thus eliminating the need for any boundary or initial condition pre-processing steps. We also included horizontal Coriolis terms in the momentum equations.

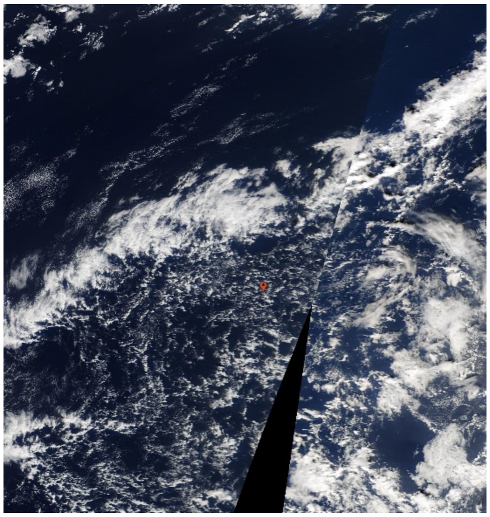


Figure 4: MODIS visible image roughly corresponding to the simulated LWP shown in Figure 4.

Figure 3 depicts the liquid water path in a PINACLES simulation of a real case with initial conditions and lateral boundaries driven by ERA5 reanalysis data. The simulation is run on a 1280 km x 1280 km x 20 km domain with 2 km horizontal resolution, essentially running PINACLES as a mesoscale or cloud resolving model. The simulation is initialized with data from January 23, 2018, at 0z and the figure shows results from the twelfth hour of the simulation. The domain is centered at 30N, 40W. An approximately corresponding MODIS visible satellite image provided by NASA Worldview is shown in Figure 4. While the two figures are exactly matched neither spatially (due to differing map projections) nor temporally (due to twice-daily satellite overpasses) the qualitative agreement between the simulated and observed cloud fields is striking, suggesting PINACLES' substantial utility for simulating real cases.

Outcomes and Impact

This seed Laboratory Directed Research Project has extended PINACLES into a substantially more general modeling capability. PINACLES now

- Is coupled to a state-of-the-art land surface model, NOAH-MP, that provides a detailed representation of land surface processes and atmosphere-land surface exchanges.
- Has a fully non-periodic dynamical core option.
- Has a grid nesting capability facilitated by the non-periodic dynamical core.
- Can be initialized and have its lateral boundaries driven by reanalysis, facilitating the simulation of real cases.

These extensions to PINACLES make it more broadly applicable to a range of scientific problems and increase the model's utility across a diverse range of topics within DOE's atmospheric science portfolio. Moreover, the ability to simulate realistic cases in a high-throughput model at very high-resolution yield unique opportunities to further leverage DOE's investments in atmospheric observations to verify and improve PINACLES, as well as to use PINACLES' model data symbiotically with observations to further process understanding and contribute to the improvement of process representations in Earth System Models.

Acknowledgements

The research described in this report was conducted under the Laboratory Directed Research and Development Program at Pacific Northwest National Laboratory, a multiprogram national laboratory operated by Battelle for the U.S. Department of Energy. A portion of the research was performed using PNNL Institutional Computing at Pacific Northwest National Laboratory. We acknowledge the use of imagery from the NASA Worldview application (https://worldview.earthdata.nasa.gov/), part of the NASA Earth Observing System Data and Information System (EOSDIS).

References

1 Lam, S. K., Pitrou, A. & Seibert, S. in *Proceedings of the Second Workshop on the LLVM Compiler Infrastructure in HPC* Article 7 (Association for Computing Machinery, Austin, Texas, 2015).

2 Pressel, K. G., Kaul, C. M., Schneider, T., Tan, Z. H. & Mishra, S. Large-eddy simulation in an anelastic framework with closed water and entropy balances. *J Adv Model Earth Sy* **7**, 1425-1456, doi:10.1002/2015ms000496 (2015).

3 Fast, J. D. *et al.* Overview of the HI-SCALE Field Campaign: A New Perspective on Shallow Convective Clouds. *B Am Meteorol Soc* **100**, 821-840, doi:10.1175/Bams-D-18-0030.1 (2019).

4 Niu, G. Y. *et al.* The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. *J Geophys Res-Atmos* **116**, doi:10.1029/2010jd015139 (2011).

5 Snyder, J. P. Map projections: A working manual. Report No. 1395, (Washington, D.C., 1987).

6 Virtanen, P. *et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python (vol 33, pg 219, 2020). *Nat Methods* **17**, 352-352, doi:10.1038/s41592-020-0772-5 (2020).

7 Lafore, J. P. *et al.* The Meso-NH atmospheric simulation system. Part I: adiabatic formulation and control simulations. *Ann Geophys-Atm Hydr* **16**, 90-109, doi:DOI 10.1007/s00585-997-0090-6 (1998).

8 Davies, H. C. Lateral Boundary Formulation for Multilevel Prediction Models. *Q J Roy Meteor Soc* **102**, 405-418, doi:DOI 10.1002/qj.49710243210 (1976).

9 Davies, H. C. Limitations of Some Common Lateral Boundary Schemes Used in Regional NWP Models. *Mon Weather Rev* **111**, 1002-1012, doi:Doi 10.1175/1520-0493(1983)111<1002:Losclb>2.0.Co;2 (1983).

10 Hersbach, H. *et al.* The ERA5 global reanalysis. *Q J Roy Meteor Soc* **146**, 1999-2049, doi:10.1002/qj.3803 (2020).

Pacific Northwest
National Laboratory

902 Battelle Boulevard
P.O. Box 999
Richland, WA 99354
1-888-375-PNNL (7665)

www.pnnl.gov