HPCI システム利用研究課題　利用報告書
HPCI User Report

課題番号 Project Number	hp200321	
課題名	Organization of shallow cumulus clouds	
Project Name	Organization of shallow cumulus clouds	
課題代表者 Project Representative	Name	Fredrik Jansson
	Affiliation	TU Delft
	Country	Netherlands
キーワード [5-10 語程度]	Large eddy simulation, cloud organization, shallow cumulus	
Keywords	Large eddy simulation, cloud organization, shallow cumulus	
利用ソフトウェア Software	IN-HOUSE (DALES)	
利用枠 Project Category	「富岳」一般機動的課題：随時募集 Fugaku General Access (Small-Scale): Year-Round Opening Call	
実施期間 Periods of Use	2021/6/8 – 2022/6/7 June 8, 2021 - June 7, 2022	

利用計算資源情報 Resource Information				
機関名 Institutions	資源名 Computer Resources	資源量（通期）Resources 単位 Units	割当量 Allocated	使用量 Amount Used
理化学研究所 RIKEN	スーパーコンピュータ「富岳」 Supercomputer "Fugaku"	ノード時間 node hour	1,000,000	582,093

Project Number: hp200321

Organization of shallow cumulus clouds

Fredrik Jansson
TU Delft

1. Background and Objectives of the Research

The goal of this project is to understand how shallow cumulus clouds develop and organize under different conditions, and to find the mechanisms driving the organization. Shallow cumulus clouds and their organization have an important but not precisely known role in cloud-climate feedback [1]. We study shallow cumulus clouds using large eddy simulation (LES) with the model DALES. We have set up an LES experiment, where the large-scale conditions are described by a set of six parameters. We then performed simulations for different parameter combinations on supercomputer Fugaku, to find how the clouds and cloud patterns formed depend on the parameters.

2. Outline of Results

The main data set from this project is an ensemble of 103 LES simulations at different points in the six-dimensional parameter space. The simulations show a variety of the organization patterns which are known from satellite observations. A typical behavior is that small cumulus clouds initially form, and grow over time. If the clouds grow large enough, they start precipitating, which in turn has a strong influence on further organization due to the formation of cold pools. The simulations have a domain size 153.6 x 153.6 x 7 km at 100 m resolution horizontally and approximately 20 m vertically, and last for a time span of 60 hours. Visualizations and animations of all simulations are preliminarily available online, at http://143.178.154.95:3141/botany-7-1536/. We are working on making also the full simulation dataset openly available. A further outcome of this project is the porting, testing, and tuning of DALES to Fugaku and the A64FX architecture.

3. Calculation Model

The simulations were carried out using DALES, the Dutch Atmospheric Large Eddy Simulator. DALES is written in Fortran and available under an open-source license at https://github.com/dalesteam/dales. The latest released version, 4.4, supports use on Fugaku and contains several of the optimizations carried out during this project, in particular optional single-precision floating point support, improved loop vectorization, reduced MPI communication, and reduced memory use.

For post-processing, merging, and compressing the netCDF output files written by each MPI process of DALES, we used cdo. Initial visualizations and analysis were done with Python scripts.

4. Method and Effect (or Performance) of Parallel Computing

DALES is parallelized using MPI in the two horizontal directions (x and y). On Fugaku we have in general been running it with 48 processes per node. We have also experimented with automatic shared-memory parallelization, using the -Kparallel compiler flag. This was slower than flat MPI, however we used it for the largest simulations we tried, with 3072 x 3072 x 175 grid points. These were run with 24 processes per node and two threads per process, in order to conserve RAM memory. In scaling tests we found 64 x 32 grid columns per process to give a good balance between scaling, wall-clock time and memory required per node. The 1536 x 1536 grid column runs which make up our main dataset were thus performed with an MPI grid of 24 x 48 processes on 24 nodes.

During the project we have worked on improving DALES performance on Fugaku. One aspect of this is the RAM memory requirement, since compared to current mainstream HPC systems Fugaku has less RAM per core. For this reason we have worked to decrease the memory use of DALES. One large improvement was made with the support for using single-precision floating point variables. The total memory use decreased by about 40%, compared to using double precision variables everywhere. Using single precision gave a speed improvement as well, and testing showed no degradation of the results (however this may be case dependent). Further we found some instances where the code could be restructured to allocate fewer 3D fields. Fig. 1 shows a strong scaling measurement for the single and double precision version of the code.

DALES in general is mostly memory bandwidth limited, with no single routine being an obvious optimization target. We have however found several instances where performance could be improved by targeting the most expensive functions. Our general approach has been to use the *fipp* profiler to find the most expensive routines and loops, and then examine the compiler optimization reports for them. This was quite efficient, often the routines were expensive because they did not vectorize, sometimes for a simple reason such as an if statement inside the loop, and these instances were then re-written. These tunings for Fugaku appear beneficial also on mainstream X86, and many of them have already been merged into the official DALES in version 4.4.

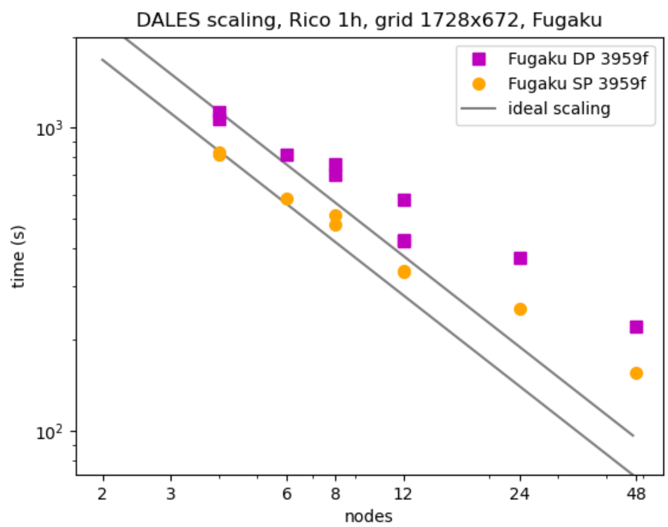

Fig. 1 Scaling test of DALES on Fugaku using one of the standard cases (RICO). The plot shows the difference between the single- (SP) and double (DP) precision version, and shows a good scaling up to 12 nodes. The 12-nodes point here corresponds most closely to the number of grid points per process used in the main simulations.

5. Research Results

In this project we simulate shallow cumulus clouds over the ocean, and their organization. We seek to determine, by LES simulations, how the large-scale physical conditions of the atmosphere influence the formation of different cloud patterns. In satellite observations four distinct patterns have been identified visually, and named *sugar*, *gravel*, *flower*, and *fish* [1]. Our aim is to create a map of the different cloud types and cloud organization patterns that our LES can produce, to study the processes leading to the formation of different mesoscale patterns, and to clarify whether the patterns are formed by small-scale processes such as aggregation of convection and precipitation or whether larger-scale structures are required for the patterns to form. The simulation domain is located over the western Atlantic, east of Barbados. This location was chosen because it was the site of the large field observation campaign EUREC4A in 2020 [2].

We have identified six parameters describing the large-scale conditions to vary in the experiment: the sea surface temperature, the potential temperature lapse rate, the surface humidity, a humidity height scale, the surface wind, and a parameter for the subsidence shape. These were chosen to span the winter conditions of the EUREC4A domain as obtained from the ERA5 reanalysis, and for their potential effects on cloud organization. These parameters span a six-dimensional parameter space in which clouds and cloud organization can be explored. For simplicity and in agreement with ERA5 data, the parameters are varied independently of each other, over ranges identified from ERA5. Thus our parameter space forms a six-dimensional hypercube.

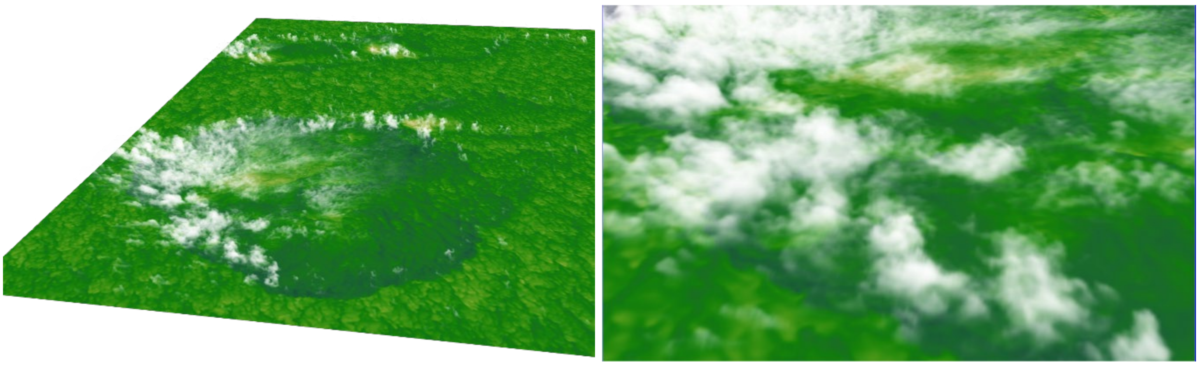

Fig. 2 3D view of one simulation (run 37) showing a large cold pool, with a zoom-in of the fine structure clouds on the right. Cloud liquid water is shown in white and total humidity near the surface in yellow (dry)–green (moist).

The main data set from this project is an ensemble of 103 large eddy simulations. The parameter space points chosen are the center and corners of the hypercube, as well as parameter sweeps through the center where one parameter at a time is varied. The simulations have a domain size 153.6 x 153.6 x 7 km and a resolution of 100 m horizontally and approximately 20 m vertically. Fig. 2 shows a 3D view of the clouds in one of the simulations, and the occurrence of a cold pool. Fig. 3 shows samples of different type of cloud organization seen in the ensemble.

We are currently working on making the full dataset publicly accessible and part of the EUREC4A intake catalog. Our aim is to store the data in the zarr format in and make it accessible with the SWIFT protocol. The zarr format is similar to netCDF but explicitly split into chunks, which makes it possible to efficiently access subsets of the data. Then it will for example be possible to access a specific slice of a specific variable online and directly form Python without downloading large datasets in advance. Examples of this, for other data sets, can be found in the *how to eurec4a* manual, https://howto.eurec4a.eu. A preliminary visualization of the ensemble, and animations of the individual runs can be found here: http://143.178.154.95:3141/botany-7-1536/index.html

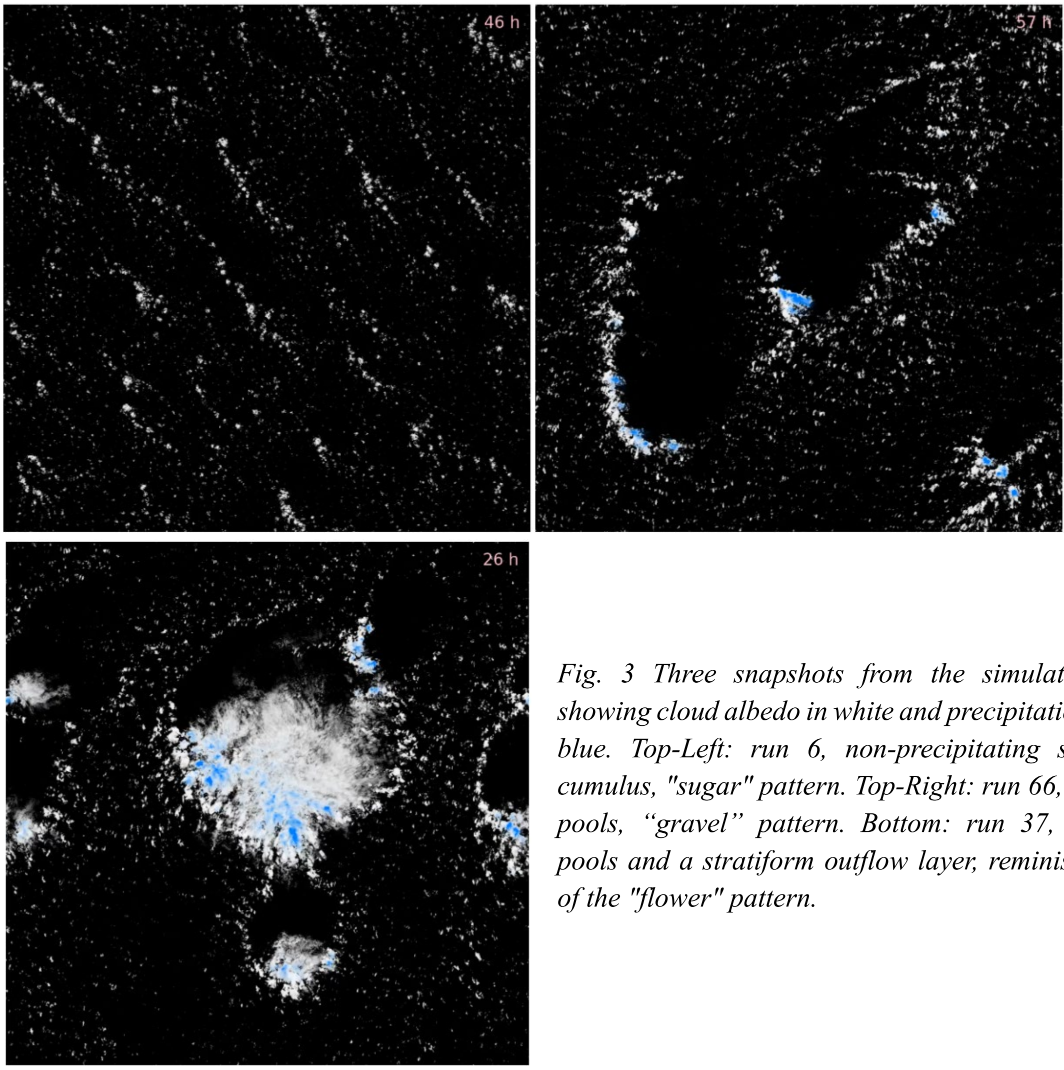

Fig. 3 Three snapshots from the simulations, showing cloud albedo in white and precipitation in blue. Top-Left: run 6, non-precipitating small cumulus, "sugar" pattern. Top-Right: run 66, cold pools, "gravel" pattern. Bottom: run 37, cold pools and a stratiform outflow layer, reminiscent of the "flower" pattern.

6. Summary and Future Subjects

Our plans for continuing this study include:

- publishing the dataset in an easily accessible format and describing it in a dataset article [3]
- further analysis of the cloud organization in the different simulations
- continuing the exploration of parameter space with more simulations
- further tuning of DALES on Fugaku. One particular optimization target is the microphysics scheme which models rain droplet creation, growth, and evaporation processes, and currently does not vectorize well.

References

[1] S. Bony et al, Geophys. Res. Lett. 48, e2019GL085988 (2020).
[2] B. Stevens et al, Earth Syst. Sci. Data, 13, 4067–4119 (2021).
[3] Cloud botany: Shallow cumulus clouds in an ensemble of idealized large-domain large-eddy simulations of the trades, Fredrik Jansson, Martin Janssens, Johanna H. Grönqvist, A. Pier Siebesma, Franziska Glassmeier, Jisk Attema, Victor Azizi, Masaki Satoh, Yousuke Sato, Hauke Schultz, and Tobias Kölling, to be submitted to Earth Syst. Sci. Data (2022).