HPCI magazine
富岳百景
FUGAKU HYAKKEI

HPCI magazine vol.2

Fugaku's Record-breaking Computations Enable More Accurate Weather Forecasting

Although weather forecasts are steadily improving, it is still difficult to accurately predict the paths of tropical cyclones and where heavy rains will occur. More accurate weather forecasts require detailed simulations on supercomputers with more processing power, and also applications designed to utilize that power. To help address this issue, Hisashi Yashiro, a member of a research project[*1] under the "Program for Promoting Research on the Supercomputer Fugaku", carried out the largest-ever meteorological computations for weather forecasting use.

Hisashi Yashiro
Senior Researcher
Satellite Remote Sensing Research Section, Earth System Division,
National Institute for Environmental Studies

Better Weather Forecasts Require Massive Computations

Modern weather forecasting is based on supercomputer simulations. As supercomputers have become more powerful, the forecasts have become more accurate over the last two decades, but nevertheless sometimes fall short. Moreover, it is quite difficult to accurately predict the paths of tropical cyclones and where heavy rains may occur. Improving these simulations will require processing colossal amounts of meteorological data points known as "state values" in the weather model.

Looking back on 2014, when the development work on Fugaku first got underway, Yashiro says, "I wanted to make use of the computing power of Fugaku to help contribute to more accurate weather predictions." It was decided that Fugaku would undergo a “co-design” approach, where the hardware and software developers would collaborate in determining specifications. Yashiro contributed to this co-design effort, overseeing the weather and climate software perspective. The co-design team began work on a project, “to develop an application that could address the challenges of massive-scale calculations at a practical level when a supercomputer with the power of Fugaku became available to national meteorological organizations in the future.”

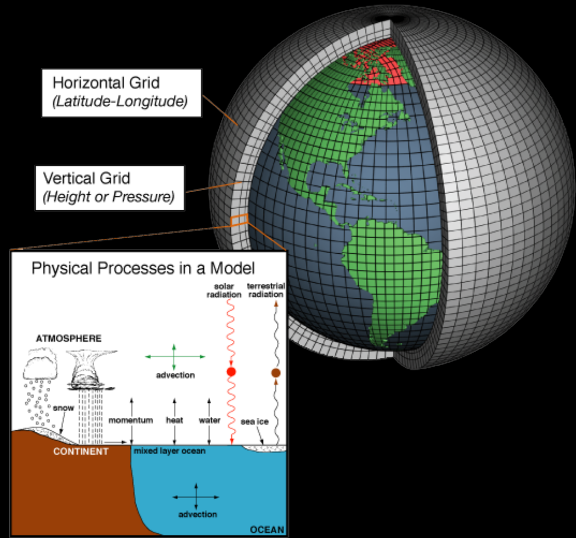
Better Forecasts Start with Better Replication of Current Weather Conditions

Weather patterns, such as clouds, rain, and snow, are the result of atmospheric conditions, such as temperature, humidity, pressure, and wind. To model these conditions for the entire planet, scientists divide the globe into a 3-dimensional grid and compute the changes expected at each grid point (or box) according to the laws of physics (Fig.1a). The challenge is choosing initial values for each grid point that accurately reflect the atmospheric conditions there. Weather pattern developments are sensitive to initial conditions. If the initial values used in the computations don't match the actual state of the atmosphere, the predictions will suffer. But, there are far more grid points in the model than actual observation points measuring atmospheric conditions. So to begin running a simulation, researchers need a way to calculate a suitable starting value for each grid point where no direct measurement data is available. Many meteorologists have been working on ways to compensate for this incomplete initial data.

Yashiro and his colleagues used an approach called “ensemble-based data assimilation”. First, they assign a temporary initial value for each grid point. Next, they run weather forecast simulations for a while to see how those values change over time. They then check those values against values acquired from actual observations, and update the value for

Fig. 1a Atmospheric Simulations and Data Assimilation

In this model, the Earth's atmosphere is divided into a grid. Changes to the atmospheric conditions at each grid point are computed according to the laws of physics. In this example, the grid is divided horizontally by latitude and longitude, and vertically by altitude with corresponding barometric pressure. The finer the mesh of the grid, the more accurate the simulation becomes, but the longer the computations and data processing take. This model is used both for weather forecasting and for data assimilation. The lower left diagram shows the physical processes modeled for each grid point, including how atmospheric heat and water interact over land and sea.



— Source: National Oceanic and Atmospheric Administration

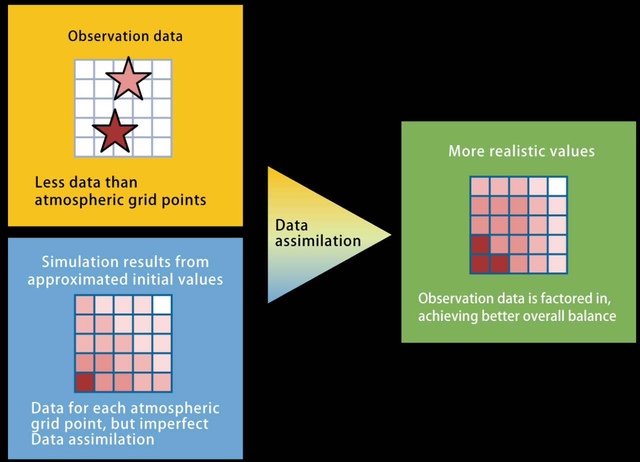


Fig. 1b

Atmospheric Simulations and Data Assimilation

Observational data is assimilated into the results of atmospheric simulations that began with approximated initial values. Each grid point is updated to more realistic values based on statistical estimates.

— Graphics based on JMA data

each grid point based on statistics and probability theory (Fig.1b). This is “data assimilation”. “Ensemble” means running multiple meteorological simulations at the same time with slightly different states. Like real weather, atmospheric simulations evolve non-linearly over time. If the weather model is incomplete, the longer the simulation runs, the more the results deviate from real-world observations. To address this problem, a number of simulations are run in parallel with different initial values. Comparing the results of multiple simulations tells the researchers which particular grid point locations are more sensitive to the initial values, and therefore more likely to contribute to a forecasting error. Those more sensitive grid points (areas) are then weighted with greater importance when the observation data is assimilated.

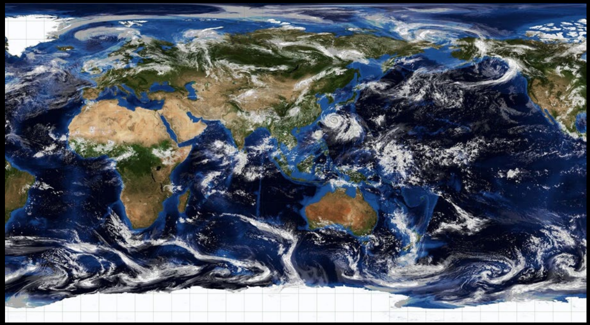
“How well data assimilation can reproduce the real atmospheric conditions will greatly affect the accuracy of the prediction results,” explains Yashiro. For this reason, meteorological agencies worldwide are attempting to assimilate data using ensembles with more members, and grids with finer meshes. Application performance and supercomputer processing power have typically allowed only dozens to hundreds of members per ensemble, with each grid point representing 10 to 50 kilometers of the Earth's surface.

The ensemble data assimilation done by Yashiro's team for this project used about 80% of Fugaku's total node count to perform computations with 1024-member ensembles, and a mesh size of approximately 3.5 kilometers. They were able to execute large-scale computations, about 500 times larger than those performed by meteorological agencies, with run times fast enough for use in weather forecasting systems.

Developing more Computationally Efficient Applications that Make Use of Fugaku's Performance

The application that Yashiro and his colleagues ran for atmospheric simulations featured optimizations by RIKEN research teams. Team Leader Hirofumi Tomita oversaw optimizations using the “Non-hydrostatic Icosahedral Atmospheric Model” (NICAM), a very high-resolution global atmospheric model. Team Leader Takemasa Miyoshi directed optimizations of the data assimilation system using a “Local Ensemble Transform Kalman Filter” (LETKF). In 2013, Tomita's team, which then included Yashiro, ran the world's highest-resolution model simulation of the global atmosphere, with a horizontal mesh

Fig. 2 Atmospheric Simulation with NICAM



The 6-hour later forecast produced by a global atmosphere simulation with a grid mesh size of 870 meters and with observation data from midnight August 25, 2012, as the initial values. Using one-quarter of the K computer's 80,000+ nodes allowed accurate reproductions of atmospheric flow and typhoon development.

— Joint research by RIKEN, the Japan Agency for Marine-Earth Science and Technology, and the University of Tokyo's Atmosphere and Ocean Research Institute (HPCI Strategic Programs for Innovative Research, Area 3).

— Visualization: Ryuji Yoshida

size of approximately 870 meters, using the K computer (Fig. 2).

This time, the mesh was less dense, with grid points every 3.5 kilometers, but with 1024 parallel NICAM calculations for each, and data assimilation throughout, making these historically intense computations. For these calculations to be practical in actual weather forecasting, however, they have to run faster than time elapses in real life. Even with the power of Fugaku, that is no simple task. Yashiro and his associates had to devise a number of modifications to make the application run faster.

In particular, a large amount of data needs to be transferred whenever shifting from the atmospheric simulation process to the data assimilation process. They therefore modified the program to minimize the time required for data transfers. They also tried to make maximum use of the solid-state drives (SSD) attached to Fugaku's computation nodes to speed up data reading and writing. SSDs have very high performance for reading and writing files. For this project, the team designed the software to take advantage of Fugaku's hardware, which was designed to include SSDs.

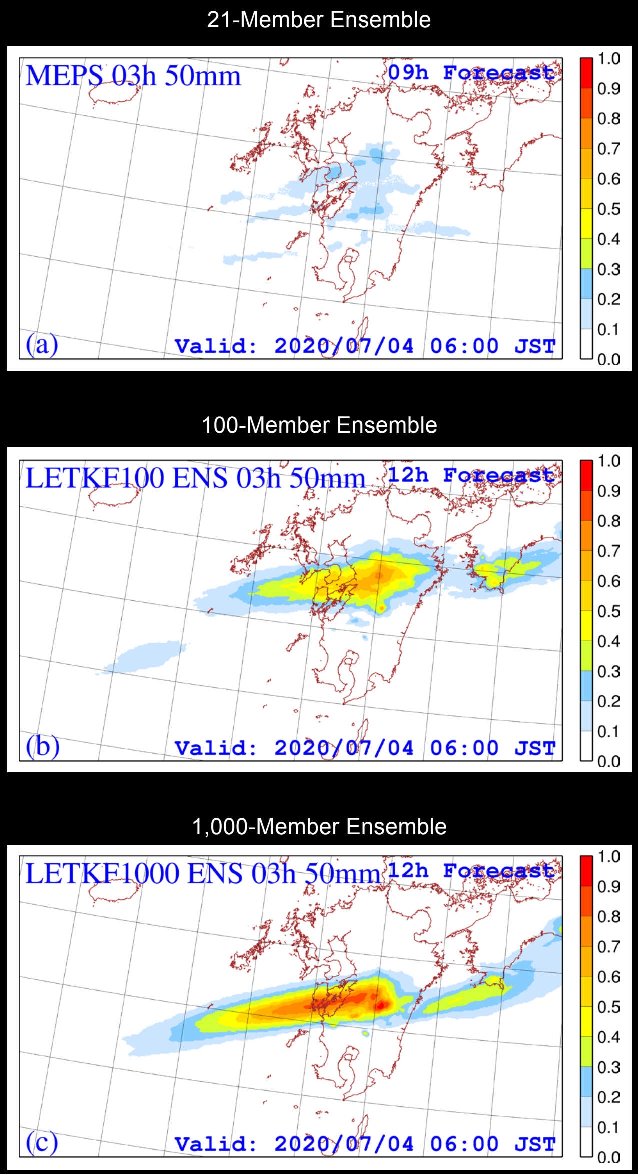
Explaining the significance of co-design, Yashiro says, “Fugaku is a world-class supercomputer that inherits the best features of the K computer. And because it was co-designed by hardware and software engineers, we could come up with ways to make full use of its performance. I think co-design will improve the results for a number of scientific fields, not only ours.” There are plans for future calculations to

Fig. 3

Increasing the number of ensemble members to improve weather forecasting.

Results from Fugaku calculations using observation data from torrential rains in July 2020 by a research project team under the “Program for Promoting Research on the Supercomputer Fugaku”. After data assimilation, the team ran ongoing ensemble calculations to produce a 12-hour later forecast predicting the likelihood of heavy rains (more than 50 mm of rain in 3 hours). The closer to red, the higher the probability. The 21-member ensemble used by the JMA's meso-ensemble prediction system did not clearly identify which locations could expect heavy rains. The research team's 1,000-member ensemble, however, was able to identify those locations, making the prediction results a strong contender for use in providing emergency sheltering advisories. The areas that the 1,000-member ensemble predicted were very likely to, did actually experience heavy rains.

— Le Duc et al. (2021): Forecasts of the July 2020 Kyushu heavy rain using a 1000-member ensemble Kalman filter, SOLA



assess the accuracy of weather forecasts based on these ensemble data assimilation results. Yashiro, however, is confident. “Seeing the results (Fig. 3) of other teams working under the 'Program for Promoting Research on the Supercomputer Fugaku', we expect the accuracy will improve considerably.” As meteorological agencies around the world develop and use software incorporating our techniques over the next ten to twenty years, they will surely be able to more accurately forecast the paths of tropical cyclones and where heavy rains will occur, which will help minimize the damage they cause.

The "Large Ensemble Atmospheric and Environmental Prediction for Disaster Prevention and Mitigation" research project received enhancement support from RIST.

The long execution time for the parallel pre-processing program for the JMA's non-hydrostatic model (NHM) was made significantly faster with enhancement support from RIST. This made it possible to execute computations including pre-processing at higher speeds with up to 10,656 Fugaku nodes, calculations not possible on the K computer.

*1 Research project name: "Large Ensemble Atmospheric and Environmental Prediction for Disaster Prevention and Mitigation" (Principal Investigator: Professor Masaki Satoh, Atmosphere and Ocean Research Institute, The University of Tokyo)
• Theme 1: "Short-range Regional Prediction"
(Principal Investigator: Takuya Kawabata, Head, Meteorological Research Institute)
• Theme 2: "Global-scale Prediction"
(Principal Investigator: Associate Professor Tomoki Miyakawa, Atmosphere and Ocean Research Institute, The University of Tokyo)
• Theme 3: "Advanced Technology of Data Assimilation"
(Principal Investigator: Hisashi Yashiro)

About the

Researcher

Hisashi Yashiro did not originally intend to become a researcher. While still a master's student, he spoke passionately about his research during a job interview. The interviewer asked him why he wasn't continuing his research, which put the idea in his head of becoming a researcher. As a student, he specialized in monitoring greenhouse gases, working with simulations only as a user of the resulting data. His research shifted to simulation-based studies and atmospheric modeling, and then to application development. In July 2019, he moved from RIKEN to the National Institute for Environmental Studies (NIES), where he is involved in research that integrates data received from the Ibuki greenhouse gas monitoring satellite into model simulations. A choral singer since his school days, he's a member of the institute's choral singing club.

Associated Research Projects

• "Large Ensemble Atmospheric and Environmental Prediction for Disaster Prevention and Mitigation" (hp200128)
Principal Investigator: Masaki Satoh, Atmosphere and Ocean Research Institute, The University of Tokyo
A paper reporting on the findings of this project was selected as a 2020 ACM Gordon Bell Prize finalist.

HPCI magazine

富岳百景

FUGAKU HYAKKEI

HPCI magazine FUGAKU HYAKKEI vol.2

Registered Institution for Facilities Use Promotion /
The Representative for HPCI Operation

**Research Organization
for Information Science and Technology**

1-5-2, Minatojimaminami-machi, Chuo-ku, Kobe, Hyogo, 650-0047 JAPAN
Tel: +81-78-599-9511

HPCI Portal Site
https://www.hpci-office.jp/folders/english

HPCI PR site
https://fugaku100kei.jp/e/

Issued: March 2021