Toward the Marine-Earth Informatics

*Keiko Takahashi[1]

1. Japan Agency for Marine and Earth Science and Technology

In this presentation, I would like to show you the importance and possibility why it is necessary what the Marine-Earth Informatics is. I shall introduce key concept and technologies that accelerate development of the academic field and industrial applications in the Marine-Earth Informatics. In addition, I would like to argue about the future prospects and ripple effects.

Keywords: Big Data, Informatics, Machine Learning

Mission, potential, and prospects of d4PDF

*Izuru Takayabu[1]

1. Meteorological Research Institute

When stakeholders make decision how to adapt to the disaster caused by climate change, PDF information of extreme events is needed. The mission of d4PDF is to get such PDF. For this purpose, at least order hundred ensemble experiments have been done. Driving hundred ensembles by using Earth Simulator, we could set the resolution of the model moderately, not so coarse. In this case, we drive AGCM with equivalent grid size of 60km (MRI-AGCM3.2H), and downscale the calculation results by using 20km grid regional climate model (NHRCM20) around the Japanese Archipelago. Here the ensemble number has been increased by adopting many kinds of SST as the lower boundary condition. The perturbation is partly caused by uncertainty comes from the accuracy of observation data. It is clearly shown that hundred ensembles have potential to produce PDF of extreme events. However, the size of the dataset is over 2PB, which makes it difficult to be used in many kinds of adaptation issues.
All the calculation in d4PDF has been done by using the Earth Simulator, under the "strategic project with special support" of the center for earth information science and technology (CEIST) / JAMSTEC. Also, data integration and analysis system (DIAS) helped us archiving the calculated data. Fundamental support has been done with program for risk information on climate change (SOUSEI), sponsored by ministry of education, culture, sports, science and technology –Japan (MEXT).

Keywords: earth simulator, ensembles, extreme event, downscale, climate change

Earthquake and tsunami forecasting procedures using large number of simulation scenarios

*Takane Hori[1], Yasuhiko Igarashi[2], Masato Okada[2]

1. R&D Center for Earthquake and Tsunami, Japan Agency for Marine-Earth Science and Technology, 2. Graduate school of frontier science, the University of Tokyo

Recent progress in earthquake and tsunami simulations, we can calculate many scenarios for earthquake generation patterns and tsunami propagation patterns. For earthquake forecasting procedure, we propose a method that is based on spatio-temporal variation in slip velocity on the plate interface, which causes interplate earthquakes (Hori et al., 2014). Model outputs are not only information about the occurrence of great earthquakes (time, place, and magnitude) but also information about the physical state evolution that causes earthquakes. To overcome the difficulty in forecasting earthquake generation resulting from uncertainty both in the physical model and in the observation data, we introduce a type of sequential data assimilation. In this method, we compare observed crustal deformation data to simulations of several great interplate earthquake generation cycles. We are currently constructing a prototype, applying this forecasting procedure to the Nankai Trough, Southwest Japan, where great interplate earthquakes have occurred and are anticipated. On the other hand, for tsunami hight forecasting, we constructed a model to forecast the maximum tsunami height by a Gaussian process (GP) that uses pressure gauge data from the Dense Oceanfloor Network System for Earthquakes and Tsunamis (DONET) in the Nankai trough (Igarashi et al., 2016). We studied the relationship between offshore and coastal tsunami heights with the aim of using DONET1 ocean-bottom pressure gauges for tsunami prediction. We assumed various tsunami models, including fault models and tsunami sources, and created a large number of simulations (more than 1,500) to reveal the relationship between DONET1 ocean bottom pressure gauge measurements and coastal tsunami heights. We found a greatly improved generalization error of the maximum tsunami height by our prediction model. The error is about one third of that by a previous method, which tends to make larger predictions, especially for large tsunami heights (>10 m). These results indicate that GP enables us to get a more accurate prediction of tsunami height by using pressure gauge data. We will further develop both earthquake and tsunami forecasting procedure using data driven science and high performance computing technology.

Smart weather forecasting by computing network comprising the Earth Simulator and edge servers

*Ryo ONISHI[1], Koya Mori[2], Fumiaki Araki[1], Ken'ichi Itakura[1], Shintaro Kawahara[1], Ryota Nakada[2], Noriyuki Takahashi[2], Keiko Takahashi[1]

1. Japan Agency for Marine-Earth Science and Technology, 2. NTT Network Innovation Laboratories

Japan Agency for Marine-Earth Science and Technology (JAMSTEC) and NTT Network Innovation Laboratories (NTT) have started a new collaborative research on the development of smart weather forecasting system that can bring new information infrastructure for future smart society. As illustrated in Fig.1, the system relies on a computing network, including the Earth Simulator at JAMSTEC and edge servers consisting the edge computing platform lead by NTT. The big data obtained from huge number of various IoT sensors are utilized to improve the forecasting in an efficient way in terms of both computation and network communication. The edge servers, in the vicinity of the users and devices, can provide detailed local weather information that can be of use for many local services. We will present the possible social applications as well as the new system itself.

Keywords: weather forecasting simulation, Earth Simulator , Edge Computing
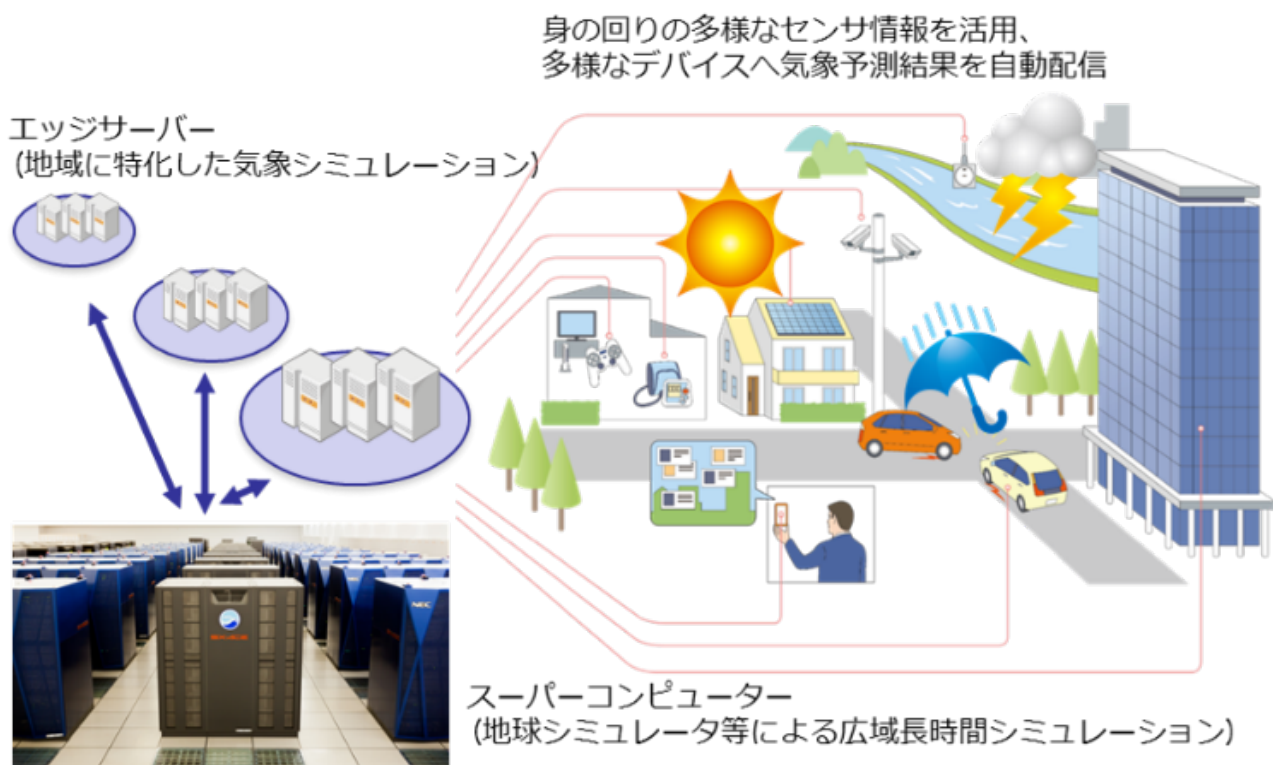
Quasi-realtime forecast using global nonhydrostatic icosahedral atmospheric model for the observation plan on the Earth Simulator.

*Mikiko Ikeda[1], Tomoe Nasuno[1], Satoru Okura[1], Hitoshi Uehara[1]

1. Japan Agency Marine-Earth Science and Technology

The Earth Simulator was once top ranked in "TOP500" supercomputers between 2002 and 2004. Its third generation introduced in 2015 performs 32 times more FLOPS than the original one, and provides irreplaceable amount of computing resources to the earth science as well as many other fields. JAMSTEC manages and operates the Earth Simulator and supports its users with technical assistance. We run near real-time forecasts using global nonhydrostatic icosahedral atmospheric model (NICAM) on the Earth Simulator during field campaigns led by JAMSTEC. The close collaboration between in-situ observation and numerical simulation is one of major tasks of Marine Earth Informatics. In Pre-YMC (November - December 2015) campaign, we conducted real-time forecasts and provided the simulation results to the observational sites, including Research Vessel "Mirai", via internet. Now we are working on the field campaign "the Years of the Maritime Continent (YMC)" (July 2017 - July 2019). We make improvement in the simulation setups, execution procedures, and the job scheduling method to efficiently run the forecast system on the Earth Simulator under a close collaboration between the research division and the operational division. We aim for better performances of forecasts both in physical accuracy and in computational performance.

Keywords: supercomputer, weather forecast, global atmospheric model

Large Scale Simulations of Dynamic Rupture Propagation to Investigate the Fault Behavior of Mega-Thrust Earthquakes

*Kenichi Tsuda[1], Jun'ichi Miyakoshi[1], Jean Paul Ampuero[2], Yoshiyuki Imato[3], Daisuke Sugiyama[3], Seiji Tsuboi[3]

1. Institute of Technology, Shimizu Corporation, 2. Seimo Lab, California Institute of Technology, 3. Japan Agency for Marine-Earth Science and Technology

The investigation of dynamic rupture propagation is very important to understand the seismic behavior of mega-thrust earthquakes such as the 2011 Tohoku earthquake. The shallow parts of the fault (near the trench) hosted large slip and long period seismic wave radiation, whereas the deep parts of the rupture (near the coast) hosted smaller slip and strong radiation of short period seismic waves. Understanding such depth-dependent feature of the rupture process of the Tohoku earthquake is necessary as it may occur during future mega-thrust earthquakes in this and other regions, such as the Nankai Trough. In order to achieve such understanding, dynamic rupture modeling is an important tool (e.g., Galvez *et al.,* 2014).

In this study, we have simulated the dynamic rupture propagation for models of the Tohoku earthquake. Our large-scale simulations used the 3D spectral element method on unstructured grids (Galvez *et al.,* 2014) with performance tuning for the Earth Simulator at JAMSTEC. The number of elements in the mesh is 4,300,000 with 2 km size and polynomial order 4. The simulation takes around 10 hours of wall-clock time on 512 cores. The effective period for the simulation is longer than 1.2 sec.

Our model reproduced the depth-dependency of the rupture process of the Tohoku earthquake. We also examine the sensitivity of the results to model parameters and assumptions, for instance to the value of the slip weakening distance (Dc). We find that the value of Dc does not affect the final slip distribution, as long as it is small enough to allow the rupture to develop and propagate to the trench. A long Dc (order of 10 m) is reasonable in terms of fracture energy and promotes the generation of long period seismic waves on the shallow part of the fault.

Keywords: Dynamic Rupture Propagation, Spectral Element, Earth Simulator, Slip Weakening Distance

The Big Data Repository at the IRIS Data Management Center: Developing Standards for Services and Federated Access for Globally Distributed Data Centers

*Timothy Keith Ahern[1], Chad Trabant[1], Mick Van Fossen[1], Rob Casey[1]
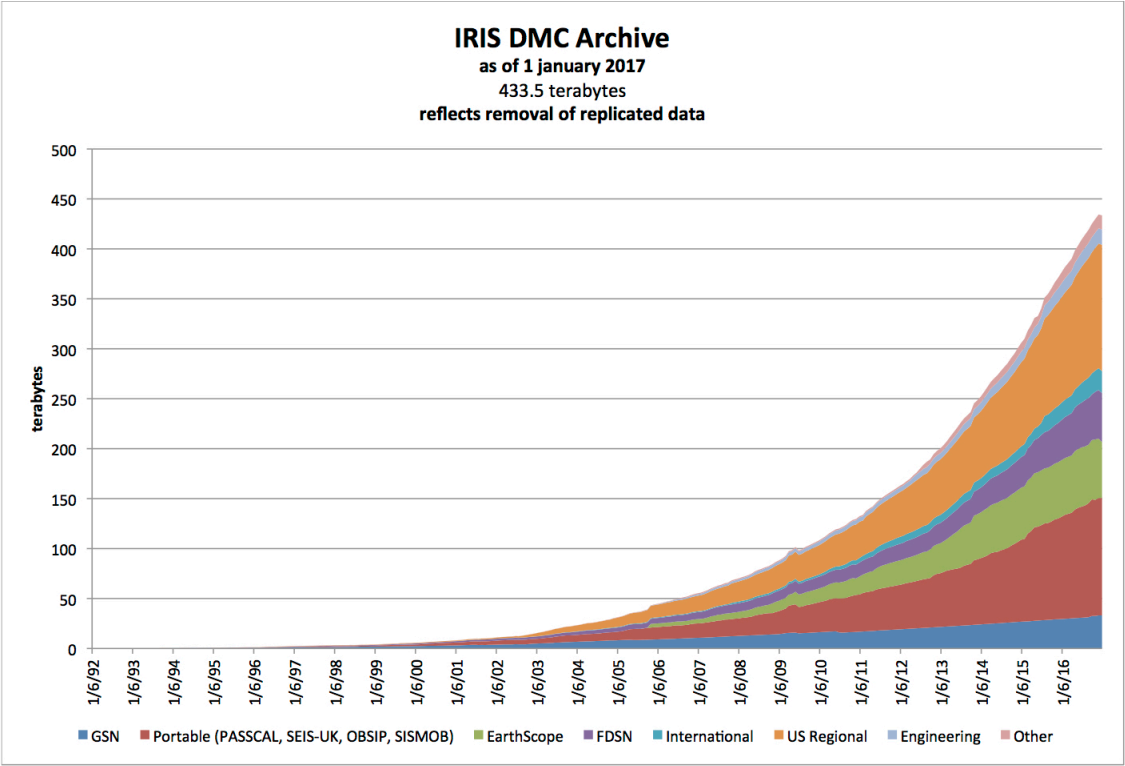
1. Incorporated Research Institutions for Seismology

The IRIS DMC is the largest repository of seismological data from permanent observatories in the world. The archive at IRIS is nearly 1/2 petabyte in size and distributes nearly 1 petabyte per year to the scientific and monitoring communities. IRIS works closely with the International Federation of Digital Seismograph Networks (FDSN) in coordinating standards for data exchange formats, metadata descriptions, standardized web services, standardization of request parameters and federated concepts for data centers.

IRIS currently operates a large analytics engine that automatically assesses quality of data from seismic observatories. This "Big Data" problem has resulted in a powerful system that will enable researchers to receive data that have been filtered by their quality in addition to the specification of space-time constraints. As the volume of data increases in the expanding federated system of seismological data centers, IRIS is also developing a system called Research Ready Data Sets (RRDS) that will allow a users data request to be filtered by their quality as measured by the quality assurance system at the DMC that is now fully operational.

IRIS is currently testing the concept of operating its data center infrastructure in both High Performance Computing (HPC) as well as Cloud Computing environments with the ultimate goal of this project to identify an appropriate environment in which to run a data center.

This presentation will describe several aspects of current IRIS DMC systems, how web services have totally changed our ability to service huge numbers of user requests, how have adopted new access mechanisms, and how quality assurance systems are now being used to improve seismic network performance as well new tools being developed that will make scientific researchers more efficient in conducting their research.

Keywords: Seismology, Web Services, Federation, Quality Assurance

IRIS DMC Archive
as of 1 january 2017
433.5 terabytes
reflects removal of replicated data

Legend: GSN | Portable (PASSCAL, SEIS-UK, OBSIP, SISMOB) | EarthScope | FDSN | International | US Regional | Engineering | Other

Science Metadata Management, Interoperability and Data Citation of the National Institute of Polar Research, Japan

*Masaki Kanao[1]

1. National Institute of Polar Research

The Polar Data Centre (PDC) of the National Institute of Polar Research (NIPR) has a responsibility to manage the polar science data as one of the National Antarctic Data Centre (NADC) under the Science Committee on Antarctic Research (SCAR). During the International Polar Year (IPY 2007–2008), a remarkable number of data/metadata involving multi-disciplinary science activities had been compiled. Although a long-term stewardship of those accumulation of metadata falls to the data centre of NIPR, the efforts have been in great collaboration with the Global Change Master Directory (GCMD), the Polar Information Commons (PIC), the World Data System (WDS) and other data science bodies / communities under the International Council for Science (ICSU). In addition, linkages of metadata interoperability with other data centers, such as the Data Integration and Analysis System Program (DIAS) of the Global Earth Observation System of Systems (GEOSS) , the Polar Data Catalogue of Canada have been initiated in 2014 by using the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH). The compiled metadata by the PDC of NIPR, moreover, are recently equipped an automatically attributing system of the Digital Object Identifier (DOI) by requesting to the DataCite through the Japan Link Center (JaLC).

Keywords: Metadata management, Interoperability, Data citation

BISMaL as an integrated information system for marine organism: a current achievement and the next step

*Takashi Hosono[1], Hideaki Saito[1], Katsunori FUJIKURA[1], Akira Sonoda[1], Seiji Tsuboi[1]

1. Japan Agency for Marine-Earth Science and Technology

In response to the concerns to marine environment issues, comprehensive grasping the information on marine organism has become more important. We started to develop an information system BISMaL (Biological Information System for Marine Life; http://www.godac.jamstec.go.jp/bismal/j/) that everyone can easily get various information on marine species around Japan (Tanaka et al. 2010). In 2016, BISMaL is nearly completed as one-stop online service with many functions: visualizing function for occurrence records of marine species as a map, browsing function for taxonomy, videos and images for a given target species, and data management service by data providers themselves. When focusing on the individual functions, there are several online databases similar to BISMa; OBIS as a global database for occurrence data of marine species, or WoRMS as a huge taxonomic database for marine species. Apparent from these online databases, BISMaL is a unique system from the viewpoint that the useful functions/services are tightly connected within the system (therefore, a one-stop service). BISMaL becomes a famous system specialized for marine biological information in Japan, but we are exploring what service we could provide as a next step of BISMaL. If the system could be connected directly to a system/database that store environmental data such as temperature, current or salinity, it would be possible to estimate causal relationships between distribution pattern of marine species and environmental factors on the next BISMaL. Alternatively, by using a large assimilated environment dataset instead of a simple observation-based dataset, it would be possible to refer environmental condition in any time and space even if biological data have no information of their habitat. We will review the functions/services in the current BISMaL, and discuss what type of services in the next BISMaL could contribute to progress of marine science or to resolving global environmental issues. [Tanaka et al. (2010) Data integration system for marine biodiversity constructed in JAMSTEC. JpGU 2010, Chiba.]

Keywords: BISMaL, Integrated information system

Construction of long-term stability archive System for Ocean observation data of large capacity

*Tomoaki Kitayama[1], Takehumi Kishira[2], Shoko Matsuda[2], Shin Miyagi[2], Shun Ishiguro[1], Kazuhiro Irei[1], Tetsuo Toyomura[1], Hideaki Saito[1]

1. Japan Agency for Marine-Earth Science and Technology, 2. Marine Works Japan LTD.

The Ocean observation data obtained from manned/unmanned research submersibles of the Japan Agency for Marine-Earth Science and Technology (JAMSTEC) continue to diversity and increasing for development of new observation technique and improvement of instrument. For example, Multi-Beam Echo Sounder (MBES) for measurement of submarine topography capable of high-resolution measurement by shortening the measuring interval and collecting vertical beam more than ever before. JAMSTEC's Ocean Observation data is particularly high academic value because it was observed under special environment of deep sea that cannot be easily accessed. It is widely used by various fields, it was necessary to keep without losses in the future. Accordingly, it is necessary to construct data archive system combining network and Large Capacity Storage that can be stored long-term and extract data from storage available if need.

The Center for Earth Information Science and Technology (CEIST) of JAMSTEC is possible to a long-term stability archive supposed the disaster recovery by making to increase the size of a storage volume and install the wideband/high speed network. The Ocean observation data is recorded in both of the hard disk and the magnetic tape that because to maintain data access and lower the operation coast. In the magnetic tape storage, by combine Hierarchical Storage Management (HSM) and Linear Tape File System (LTFS), the flexibility of the operation is secured and lower the operation coast about power consumption or the like. Furthermore, we constructed the wideband/high speed network to introduce the data transfer system by User Datagram Protocol (UDP) and install the wideband network between bases in accordance with update the Science Information NETwork (SINET) of National Institute of Informatics (NII).

In this a nnouncement, we report on the activity of data archive system in JAMSTEC.

Keywords: Ocean Observation Data, Large Capacity Storage, Disaster Recovery, High-speed data transfer, Data Archive System

Classification and visualization of simulated clouds using machine learning

*Daisuke Matsuoka[1]

1. Japan Agency for Marine-Earth Science and Technology

High-resolution atmospheric general circulation models reproduce realistic behavior of atmosphere in global scale. The data set generated by such simulation contains a large amount of information. One of the most important variable of the simulation results is a cloud. In order to understand such simulation results, it is necessary to visualize individual clouds and their physical properties. In the present study, we propose a new visualization method which enables scientists to classify and visualize them based on ten type cloud classification proposed by World Meteorological Organization (WMO). The proposed method is divided into two steps. In the first step, individual clouds are classified into six types (low clouds, middle clouds, high-clouds, low-middle clouds, cumulus, cumulonimbus) based on their vertical flow and altitude of top and bottom of them. In the final step, their clouds are further classified more finely into ten types (cirrus, cirrostratus, cirrocumulus, altocumulus, altostratus, nimbostratus, stratocumulus, stratus, cumulus, cumulonibmus) by their appearance using deep learning which is one of machine learning techniques. Here, we used photographs of these clouds, which we can easily download on the web, as training data. As a result, we succeeded in effectively visualizing three-dimensional cloud and their temporal behavior during complex atmospheric phenomena such as development of cumulonimbus and generation of tropical cyclone. The proposed method is beneficial to intuitive understand information-rich simulation data.

Keywords: Visualization, atmospheric simulation, machine learning

Co-analysis of observations and LESs in MM21 district

*Tooru Sugiyama[1], Sadohara Satoru[2]

1. Japan Agency for Marine-Earth Science and Technology Center for Earth Information Science and Technology, 2. Yokohama National University Institute of Urban Innovation

We have performed thermal and wind environment LESs in MM21 district in Yokohama. The used simulation model is MSSG (Multi-Scale Simulator for the Geo-environment). The spatial resolution is about 5m in horizontal and vertical axis. We have also made observations at the Grand moll park located center of MM21 district. At the center of the park, there observed some characteristic wind which blows to anti-direction of the area averaged wind. Here we will report the collaboration of these numerical and observational analysis in the micro-meteorology in the urban area, especially in the point of view that the number of data are highly un-balancing.

Keywords: thermal environment, micro-meteorology, Minato Mirai 21