

Review

High-Performance Computing in Meteorology under a Context of an Era of Graphical Processing Units

Tosiyuki Nakaegawa 

Meteorological Research Institute, Tsukuba 305-0052, Japan; tnakaega@mri-jma.go.jp; Tel.: +81-29-853-8622

Abstract: This short review shows how innovative processing units—including graphical processing units (GPUs)—are used in high-performance computing (HPC) in meteorology, introduces current scientific studies relevant to HPC, and discusses the latest topics in meteorology accelerated by HPC computers. The current status surrounding HPC is distinctly complicated in both hardware and software terms, and flows similar to fast cascades. It is difficult to understand and follow the status for beginners; they need to overcome the obstacle of catching up on the information on HPC and connecting it to their studies. HPC systems have accelerated weather forecasts with physical-based models since Richardson's dream in 1922. Meteorological scientists and model developers have written the codes of the models by making the most of the latest HPC technologies available at the time. Several of the leading HPC systems used for weather forecast models are introduced. Each institute chose an HPC system from many possible alternatives to best match its purposes. Six of the selected latest topics in high-performance computing in meteorology are also reviewed: floating points; spectral transform in global weather models; heterogeneous computing; exascale computing; co-design; and data-driven weather forecasts.

Keywords: meteorology; graphical processing unit (GPU); high-performance computing (HPC); heterogeneous computing; floating point; spectral transform; co-design; data-driven; physical-based; weather forecast model

**Citation:** Nakaegawa, T.High-Performance Computing in Meteorology under a Context of an Era of Graphical Processing Units. *Computers* **2022**, *11*, 114. <https://doi.org/10.3390/computers11070114>

Academic Editor: Paolo Bellavista

Received: 20 June 2022

Accepted: 10 July 2022

Published: 13 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Numerical weather predictions were one of the first operational applications of scientific computations [1] since Richardson's dream in 1922, which attempted to predict changes in the weather by a numerical method [2]. The current year of 2022 is the hundred year anniversary. After so-called programable and electronic digital computers were developed, computers such as ENIAC have been used for weather predictions. High-performance computing (HPC) enables atmospheric scientists to better predict weather and climate because it allows them to develop numerical atmospheric models with higher spatial resolutions and more detailed physical processes. Therefore, they develop numerical weather/climate models (hereafter referred to as weather models) so that the models can be run on any HPC system available to them.

In the 1980s, supercomputers composed of vector processors such as CRAY-1 and -2 dominated computing fields because computing was primarily used in the business and scientific fields [3]. These computations focused on floating point operations per second (FLOPS); the supercomputers had central processing units (CPU) that implemented an instruction set where its instructions were designed to operate efficiently and effectively on large one-dimensional arrays of data called vectors (https://en.wikipedia.org/wiki/Vector_processor, accessed on 3 May 2022). Since the middle of the 1980s, CPUs for personal computers (PCs) such as the Intel x86 CPU became popular and the cost performance became higher for CPUs for PCs because of the economics of scale and lower CPUs with vector processors (Figure 1). In the late 1990s, many vendors built low supercomputers with the free operating system Linux with many CPUs for PCs as a cluster. This concept is

called the Beowulf approach; it faded out the supercomputer and faded in HPC clusters. Multi- and many-core CPUs were released in the late 2000s. This concept was designed to break the technological limitation of a higher operating frequency and a higher density of large integrated circuits to align with Moore's Law. Since the late 2000s, HPC clusters have usually been composed of multi- and multi-/many-core CPUs. In parallel to the dawn of multi-/many-core CPUs, graphical processing units (GPU) have become faster with more generalized computing devices or general purpose GPUs (GPGPUs; GPGPU is hereafter referred to as GPU) [4]. Deep learning has also attracted people to the emerging technology because deep learning behaves similar to human intelligence. Capable GPUs play an important role in deep learning. A few HPC machines have been composed of multi-/many-core clusters and GPU clusters since the late 2010s because a GPU is fast and cost-effective in HPC businesses and science. However, the two components do not always work with each other and often work independently. Exascale computing of 1018FLOPS has been a hot topic in HPC technology and science since the late 2010s; heterogeneous computers composed of different CPUs and GPUs are essential for exascale computing and broad users with different computing purposes. Since the late 2010s, processing units for deep learning or artificial intelligence (AI-PU, often referred to as an AI accelerator) have been developed because of the high demand for AI applications in many industrial sectors. An AI-PU has similar features to a GPU, but only has special matrix processing chips with a high volume of low-precision computation, which is considered to have an important role in the near future.

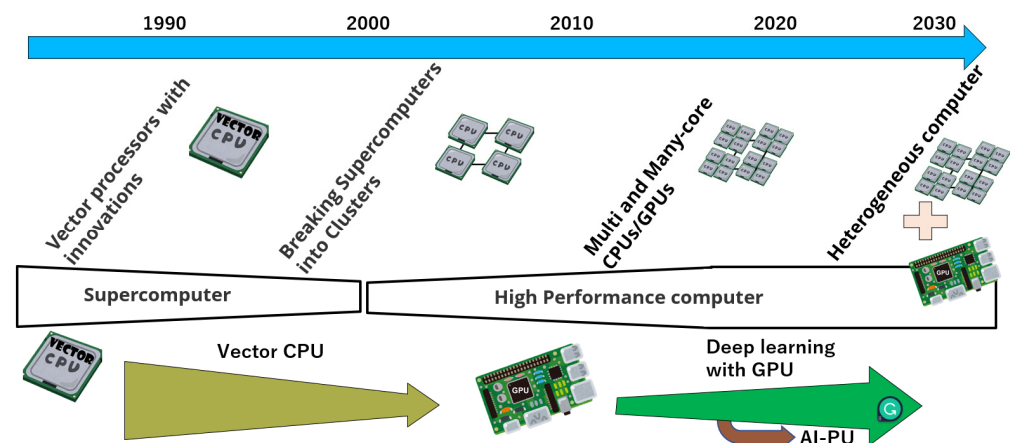


Figure 1. History of high-performance computing with the evolution of hardware since the 1980s.

The current status surrounding HPC is distinctly complicated in both hardware and software terms, and flows similar to fast cascades. It can be difficult to understand and follow the status for beginners. A short review is required to overcome the obstacles to be fully aware of the information on HPC and to connect it to studies. This short review shows how innovative processing units—including GPUs—are used in HPC computers in meteorology, introduces the current scientific studies relevant to HPC, and discusses the latest topics in meteorology accelerated by HPC computers.

2. Utilization of GPUs in Weather Predictions

There are two approaches to the usage of GPUs in meteorology as well as general science: the data-driven approach, and the physical-based approach (Figure 2). Deep learning is most popular in the data-driven approach. Deep learning can provide weather forecasts even with large ensemble members in a short computing time once a deep learning model has been trained with a large amount of data [5–7]. So-called big data or a large amount of data are required to train the deep learning model. Many deep learning models have open-source frameworks and several of them are freely available (e.g., TensorFlow and Keras; see Appendix A for further information). The deep learning model can be

developed with less effort than a physical-based model. A few of the physical-based models introduced below are freely available from downloads, but it is difficult to understand all codes from the top to the bottom. GPUs can effectively accelerate the training and are always used in deep learning studies. This approach prohibits areas without data to develop a deep learning model in most cases, and contributes little to the progress of the science of meteorology because one cannot tell how input data are mapped to output data through physical processes in the atmosphere.

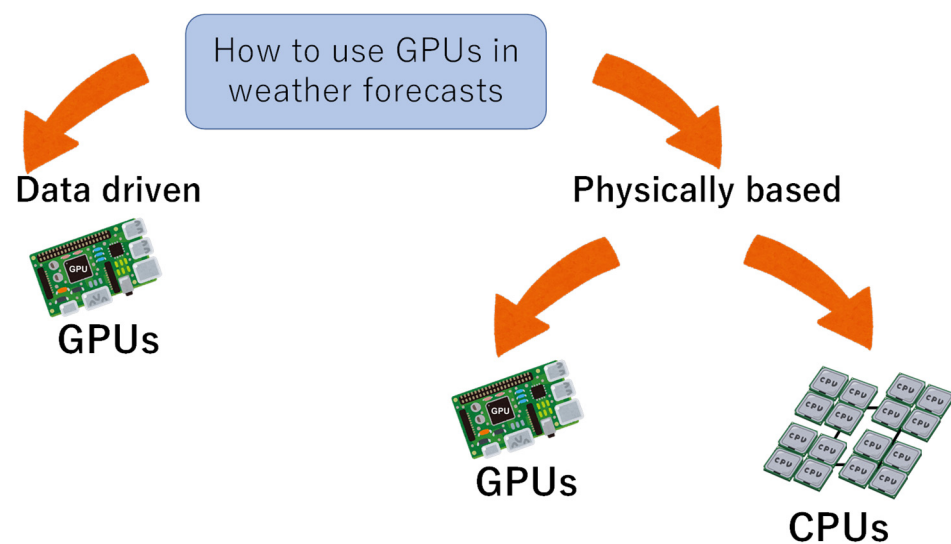


Figure 2. Schematics for the usage of GPUs in a weather forecast.

The physical-based numerical weather model has been the traditional approach for predicting future weather since Richardson’s dream [2]. The model is described in discretized mathematical formulas of the physical processes and can implement progress in science in a relatively easy manner. The usage of GPUs brings fast predictions and high cost-performance to weather predictions. However, the original codes of the model are written for CPU-based computers and need to be rewritten or converted to run on GPUs. Software for exporting codes to GPUs (e.g., Hybrid Fortran) [8] has been developed to lower these obstacles. In the following sections, HPC computers for physical-based weather models are discussed.

3. HPC in Meteorological Institutes and Centers

3.1. European Centre for Medium-Range Weather Forecasts

The European Centre for Medium-Range Weather Forecasts (ECMWF) is the leading center of weather forecasts and research in the world. ECMWF predicts global numerical weather and other data for the members and co-operating states of the European Union as well as the rest of the world. ECMWF has one of the fastest supercomputers and largest meteorological data storage facilities in the world. The fifth generation of the ECMWF re-analysis, ERA5 [9], is one of their distinguished products. Climate re-analyses such as ERA5 combine past observations with models to generate a consistent time series of multiple climate variables. ERA5 is used for quasi-observations in climate predictions and sciences, especially in regions where no observation data are available [10].

ECMWF is scheduled to install a new HPC computer in 2022. It has a total of 1,015,808 AMD EPYC Rome cores (computer processor microarchitecture, Zen2) without GPUs and has a total memory of 2050 TB [11]. Its maximal LINPACK performance (Rmax) is 30.0 PFLOS. This HPC computer was ranked 14th on the Top 500 Supercomputer List as of June 2022 (<https://www.top500.org/lists/top500/list/2022/06/>, accessed on 3 May 2022). The choice of an HPC computer without GPUs was probably because stable operational forecasts are prioritized. Similar choices have been seen at the United Kingdom

Meteorological Office (UKMO), Japan Meteorological Agency (JMA), and others. ECMWF is, however, working on a weather model on GPUs with hardware companies such as NVIDIA under a project called ESCAPE [12] (<http://www.hpc-escape.eu/home>, accessed on 3 May 2022) and ESCAPE2 (<https://www.hpc-escape2.eu/>, accessed on 3 May 2022).

3.2. Deutscher Wetterdienst

Deutscher Wetterdienst (DWD) is responsible for meeting the meteorological requirements arising from all areas of the economy and society in Germany. The operational weather model, ICON [13], has an icosahedral grid system to avoid the so-called pole problem, which produces too small grids in polar regions to set long time steps in numerical integrations due to the Courant–Friedrichs–Lewy Condition [14].

DWD installed a new HPC system in 2020. It is supposed to have a total of 40,200 AMD EPYC Rome cores (Zen2) without GPUs and has a total memory of 429 TB (256 GB \times 1675 CPUs; [15]; https://www.dwd.de/DE/derdwd/it/_functions/Teasergroup/datenverarbeitung.html, accessed on 3 May 2022) when all computers are facilitated. Its Rmax is 3.9 and has 3.3 PFLOS. This HPC computer was ranked 130th and 155th on the Top 500 Supercomputer List in June 2022. The unique feature of this system is vector engines, which are exiting in the development of HPC computers because of the small economic scale as mentioned in the Introduction. The vector engines are an excellent accelerator for weather models that were written for CPUs with vector processors. Vector engines and GPUs share a similar design from the ground-up to handle large vectors, but a vector engine has a wider memory bandwidth. This choice also prioritizes a stable operational HPC system for weather forecasts.

3.3. Swiss National Supercomputing Centre

The Swiss National Supercomputing Centre (CSCS) develops and operates cutting-edge high-performance computing systems as an essential service facility for Swiss researchers. These computing systems, Piz Daint, are used by scientists for a diverse range of purposes from high-resolution simulations to the analysis of complex data (<https://www.cscs.ch/about/about-cscs/>, accessed on 3 May 2022). MeteoSwiss operates the Consortium for Small-Scale Modelling (COSMO) with a 1.1 km grid on a GPU-based supercomputer, which was the first fully capable weather and climate model to become operational on a GPU-accelerated supercomputer (<http://www.cosmo-model.org/content/tasks/achievements/default.htm>, accessed on 3 May 2022). This required the codes to be rewritten from Fortran to C++ [16]. Fuhrer et al. [17] demonstrated the first production-ready atmospheric model, COSMO, at a 1 km resolution for near-global areas. These simulations were performed on the Piz Daint supercomputer. This simulation opens a door for less than 1 kilometer-scale global simulations.

CSCS installed an HPC system in 2016. It has a total of 387,872 Intel E5-2969 v3 cores with 4888 NVIDIA P100 GPUs and has a total memory of 512 TB (CSCS 2017; <https://www.cscs.ch/computers/decommissioned/piz-daint-piz-dora/>, accessed on 3 May 2022) [18]. Its Rmax is 21.23 PFLOS. This HPC computer was ranked 23rd on the Top 500 Supercomputer List in June 2022. This choice was due to their leading efforts on exporting to GPUs as well as multi-user environments.

3.4. National Center for Atmospheric Research

The National Center for Atmospheric Research (NCAR) was established by the National Science Foundation, USA, in 1960 to provide the university community with world-class facilities and services that were beyond the reach of any individual institution. NCAR has developed a weather model, the Weather Research and Forecasting (WRF), which is widely used in the world. WRF is a numerical weather prediction system designed to serve both atmospheric research and operational forecasting needs. WRF is also used for data assimilation and has derivations such as WRF-Solar and WRF-Hydro. A GPU-accelerated version of WRF was developed by TQI (<https://wrfg.net/>, accessed on 3 May 2022). NCAR

is developing a Model for Prediction Across Scales (MPAS) for global simulations. MPAS has a GPU-accelerated version as well as WRF.

The Centre will install a new HPC computer, DERECHO, at the NCAR Wyoming Supercomputing Center (NWSC) in Cheyenne, Wyoming, in 2022 [19] (https://arc.ucar.edu/knowledge_base/74317833, accessed on 3 May 2022). It has a total of 323,712 AMD EPYC Milan 7763 cores (Zen3) with a total of 328 GPUs of NVIDIA A100 and a total memory of 692 TB [20]. Its Rmax is 19.87 PFLOS. The high-speed interconnection bandwidth in a Dragonfly topology reaches 200 Gb/s. The HPC computer was ranked 25th on the Top 500 Supercomputer List in June 2022. The choice of an HPC computer with GPUs was because NCAR is the leading atmospheric science institution in the world, seeking the potential of GPU acceleration in weather models such as WRFg and GPU-enabled MPAS.

3.5. Riken or Institute of Physical and Chemical Research in Japan

The Institute of Physical and Chemical Research (RIKEN) Center for Computational Science [20] is the leading center of computational science, as the name suggests. There are three scientific computing objectives: the science of, by, and for computing. One of the main research fields by computers is atmospheric science such as large ensemble atmospheric predictions. One thousand ensemble members of a short-range regional-scale prediction were investigated for disaster prevention and evacuation. A reliable probabilistic prediction of a heavy rain event was achieved with mega-ensemble members of 1000 [21].

The Centre installed a new HPC computer, Fugaku—named after Mt. Fuji—in 2021 (<https://www.r-ccs.riken.jp/en/fugaku/>, accessed on 3 May 2022). It has a total of 7,630,848 Fujitsu A64FX cores without GPUs and has a total memory of 4850 TB [22]. Its maximal LINPACK performance (Rmax) is 997 PFLOS, which is almost that of an exascale supercomputer. This HPC computer was ranked second on the Top 500 Supercomputer List in June 2022, having retained the first position for two years. The choice of an HPC computer without GPUs was probably because A64FX is a fast processor with a scalable vector extension, which allows for a variable vector length and retains its speed for actual versatile programs.

3.6. Japan Agency for Marine–Earth Science and Technology

The Japan Agency for Marine–Earth Science and Technology (JAMSTEC) is working for our society in achieving this goal by developing new scientific and technological capabilities that contribute to the sustainable development and responsible maintenance of a peaceful and fulfilling global society [23]. Oceans play an important role in climate change as well as variabilities through the interaction between the atmosphere and oceans and through the huge heat capacitor of the Earth. Therefore, JAMSTEC engages in future climate projections and contributes to adaptation plans in Japan. Future climate projections with a super-high horizontal resolution of 20 km grid spacing have been performed for 20 years [24] and JAMSTEC has initiated one of the CMIP6 experiments, the High Resolution Model Intercomparison Project (HighResoMIP) [25], which serves as a more reliable source for assessing climate risks that are associated with small-scale weather phenomena such as tropical cyclones and line-shaped heavy precipitation [26].

JAMSTEC installed a new HPC system in 2021. It is composed of three computers. The first one has a total of 43,776 AMD EPYC Rome cores (Zen2) without GPUs and has a total memory of 2050 TB (JAMSTEC 2021; <https://www.jamstec.go.jp/es/en/>, accessed on 3 May 2022). Its Rmax is 9.99 PFLOS. This HPC computer was ranked 51st on the Top 500 Supercomputer List in June 2022. This computer has vector engines the same as in DWD to make most of the conventional optimized codes written for vector engines. The other two are a system without GPUs and vector engines, and a system with GPUs. This type of HPC system is called a heterogeneous system and is mentioned below. The choice of a heterogeneous HPC system was probably because JAMSTEC promotes studies on artificial intelligence in Earth system science as well as the conventional ones. This choice is similar to NCAR.

3.7. Institutes in China and the United States

China has the largest HPC system share of 34.6% in the Top 500 Supercomputer List as of June 2022 and the United States has the second largest of 25.6% when the HPC performance share or the total of Rmax is selected as the comparison category. The United States has the largest HPC performance share of 47.7% and China has the third largest of 12%. The China Meteorological Administration has an HPC system composed of two computers: one has a total of 50,816 Xeon Gold cores with an NVIDIA Tesla P100 with an Rmax of 2.55 PFOPS; the other has a total of 48,128 Xeon Gold cores without GPUs with an Rmax of 2.44 PFOPS. The United States is leading HPC developments and operations and has 32% in the top 100 of the Top 500 Supercomputer List as of June 2022, including the first rank HPC system mentioned below. The Department of Energy operates 10 out of 32 HPC systems in the United States. They are used for multiple purposes, including weather and climate research. The National Oceanic and Atmospheric Administration (NOAA), as the national meteorological agency, installed an HPC system in 2018. It has a twin with a total of 327,680 AMD EPYC 7742 cores (Zen2) and 48,128 without GPUs.

Figure 3 shows the schematics for the positions of each type of HPC system introduced above on a simplified plane expressed as with/without CPUs and/or GPUs. There are three different types. ECMWF is positioned in the CPU-only field of Figure 3; most meteorological agencies are positioned in the same field if their HPC systems are drawn because stable operational weather forecasts are prioritized. CSCS and NCAR are positioned in the CPU–GPU field whereas DWD and R-CCS are positioned in the vector field. JAMSTEC is positioned mainly in the vector field and secondarily in the CPU field as well as in the CPU–GPU field.

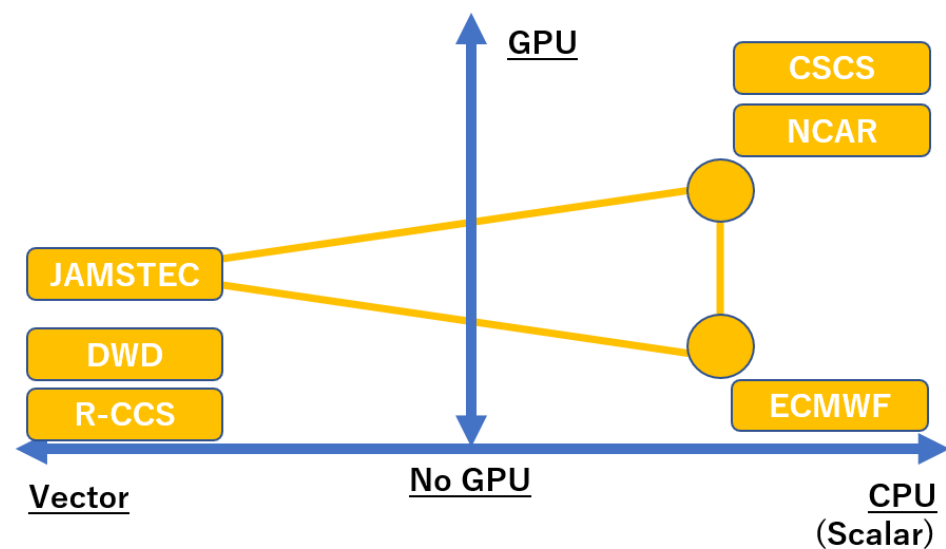


Figure 3. Schematics for positions of each type of HPC system introduced in Section 3 on a simplified field expressed as with/without CPUs and/or GPUs.

4. Latest Topics in High-Performance Computing in Meteorology

4.1. Floating Point

Several topics relevant to GPUs, especially accelerated AI-PU, are emerging in HCP. The double-precision floating point and numerical integration scheme have been prerequisites in weather prediction. However, the usability of single-precision floating points and even half-precision floating points has been investigated for a few segments of whole computations for two reasons [27,28]. First, the volume of variables is reduced by one second and can be stored in the cache and memory, which the CPUs can rapidly access. Second, several GPUs have a fast processing unit for single- and half-precision floating

points in comparison with a double one. Moreover, NVIDIA introduced mixed-precision (TensorFloat-32) running on the A100 GPU [29], reaching 156 TFLOPS [30].

4.2. Spectral Transform in Global Weather Models

The spectral transform method used in the dynamical core of several global weather models requires large computing resources such as $O(N^3)$ calculations and $O(N^3)$ memory, where N is the number of grids in the longitudinal direction representing a horizontal resolution because it usually uses Legendre transformation in the latitudinal direction. There are two reduction methods: fast Legendre transformation $O(N^2(\log N)^3)$ calculations and $O(N^3)$ memory [31], and the double Fourier series, which uses fast Fourier transform (FFT) for both longitudinal and latitudinal directions with $O(N^2 \log N)$ calculations and $O(N)$ or $O(N^2)$ memory [32]. The latter method is superior in both calculations and memory to the former one, especially for a high horizontal resolution of about 1 km or less.

FFT is widely used in scientific and engineering fields because it rapidly provides the frequency and phase of the data. Hence, several GPU-accelerated libraries are available (<https://docs.nvidia.com/cuda/cuda-c-std/index.html>, accessed on 3 May 2022) by calling the FFT library application programming interface, cuFFT [33]. Hence, the double Fourier series mentioned above has another merit in this availability.

4.3. Heterogeneous Computing

Different users utilize computers, including the HPC system, for different purposes. Different hardware such as CPUs, GPUs, and AI-PUes each have strong features. Under these circumstances, a multi-user system may be heterogeneous when it is optimized for different users. A plain heterogeneous system is composed of single-type CPUs and single-type GPUs, which are widely in operation; another heterogeneous system is composed of multi-type CPUs and multi-type GPUs by a natural extension. The recent exponential growth of data and intelligent devices requires a more heterogeneous system composed of a mix of processor architectures across CPUs, GPUs, FPGAs, AI-PUes, DLPs (deep learning processors), and others such as vector engines, collectively described as XPUes [34,35]. The vector-GPU field is vacant in Figure 3 and will be filled with the institute(s) as heterogeneous computing prevails in HPC.

A plain heterogeneous system allows us to develop codes/software in single/multiple languages such as C++ and CUDA. However, multiple languages cannot describe the codes/software properly with high optimization for the heterogeneous system. To overcome such an obstacle, a high-level programming model, SYCL, has been developed by the Khronos group (<https://www.khronos.org/sycl/>, accessed on 3 May 2022), enabling code for heterogeneous processors to be written in a single-source style using completely standard C++. Data Parallel C++ (DPC++) is an Intel open-source implementation [36]. DPC++ has additional functions, including unified shared memory and reduction operations (<https://www.alcf.anl.gov/support-center/aurora/sycl-and-dpc-aurora>, accessed on 3 May 2022).

ECMWF is leading the research into heterogeneous computing for weather models designed to run on CPU-GPU hardware based on the outcomes of ESCAPE and ESCAPE2 as mentioned above. Weather model developments under literally heterogeneous computing composed of three XPUes or more cannot be realized without collaborative efforts among not only model developers, but also semiconductor vendors. Heterogeneous computing as HPC is now a preparation stage in meteorology [37].

4.4. Co-Design

When we design an HPC system and environment, multiple scientists, software developers, and hardware developers need to discuss an optimal system with collaborations; hardware based on existing technology and software running on it as well as new hardware and software not yet emerging result in challenging scientific goals (Cardwell et al.) [38–40]. The best ideal HPC system requires the shared development of actionable information that

engages all communities in charge and the values guiding their engagement. Figure 4 depicts the co-design of an HPC system, which is drawn from the inspiration of co-producing regional climate information in the Intergovernmental Panel on Climate Change Sixth Assessment Report [41]. These three communities are essential for a better HPC system and keep the co-design process ongoing. Participants in the community represented by the closed circles have different backgrounds and perspectives, meaning a heterogeneous community. Based on the understanding of these conditions, sharing the same concept can inspire trust among all the communities and promote the co-design process. Both co-designed regional climate information and an HPC system share the same idea in terms of collaborations across the three different communities.

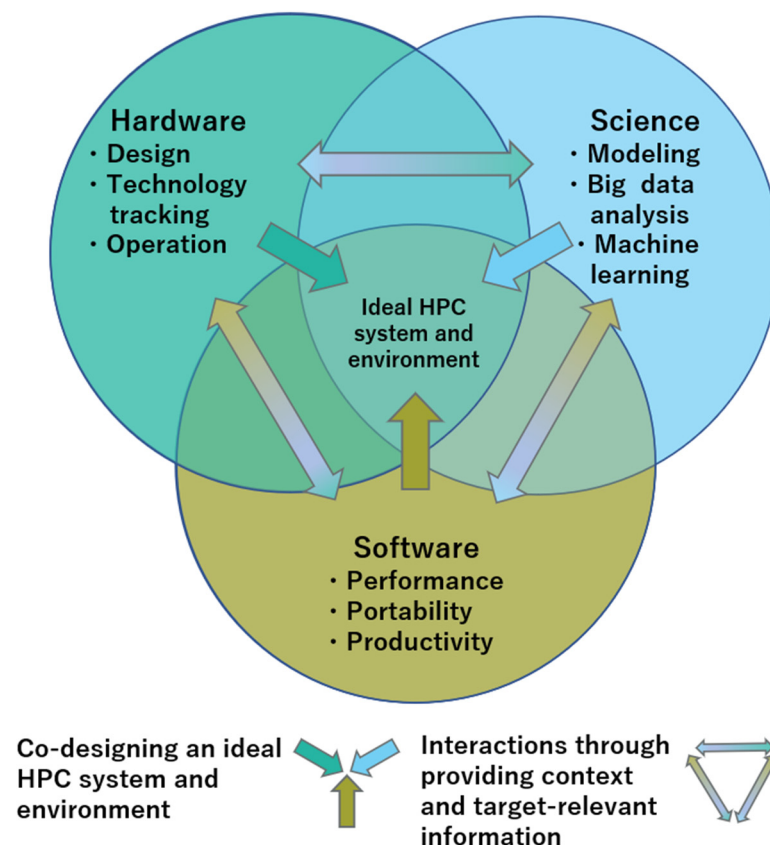


Figure 4. The co-design of an HPC system. Participants in the community—represented by the closed circles—have different backgrounds and perspectives, meaning a heterogeneous community. The arrows connecting those communities denote the distillation process of providing context and sharing HPC-relevant information. The arrows that point toward the center represent the distillation of the HPC information involving all three communities. This figure is drawn based on Figure 10.17 in the Intergovernmental Panel on Climate Change Sixth Assessment Report [41].

4.5. Resource Allocation of an HPC System

A higher horizontal resolution of a physical-based weather model can resolve more detailed atmospheric processes than a lower one [42,43]. Operational meteorological institutes/centers do not always allocate their resources to the HPC systems to enhance a horizontal resolution because they have many options to raise their forecast skills such as ensemble forecasts, detailed physical cloud processes, and initial conditions produced by four-dimensional data assimilations. Figure 5 shows the time evolutions of the theoretical performance of the HPC systems and the horizontal resolutions of the global weather model of JMA [44]. As a long trend, the horizontal resolution increases as the HPC system has a higher performance. However, the timing of the operational starts between the improved

horizontal resolutions and the HPC systems is different owing to the system optimization (mentioned above), stable operations, and others.

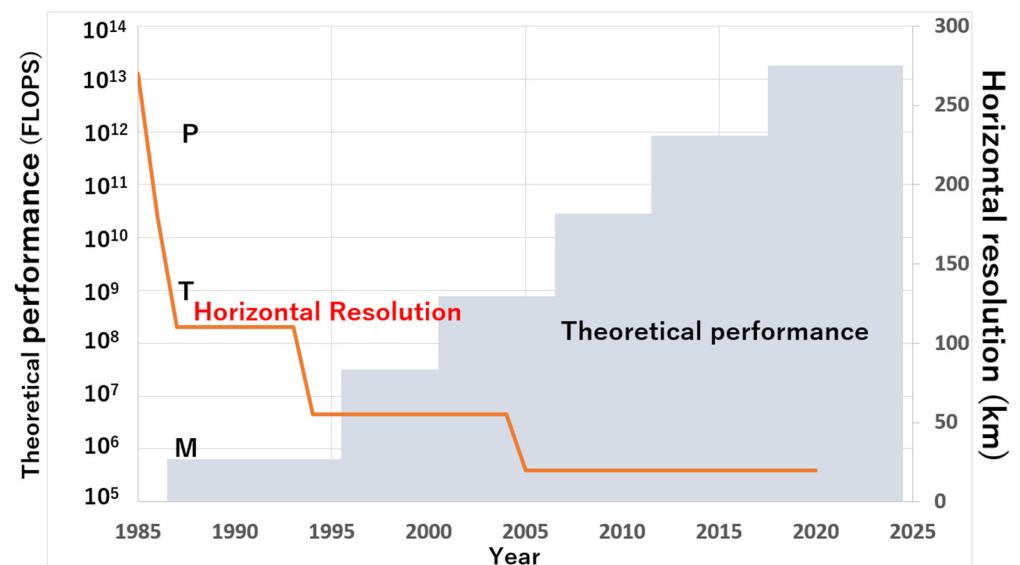


Figure 5. Time evolutions of the theoretical performance of the HPC systems and of the horizontal resolutions of the global weather model of JMA. The data were obtained from JMA [44]. M, T, and P in this figure represent the unit of Mega, Tera, and Peta, respectively.

4.6. Data-Driven Weather Forecast

This review focused on physical-based weather models. Data-driven weather forecasts are becoming popular and will be essential components of weather forecasts, enabled by technological advancements in GPUs and DLPs. Here, we briefly introduce two studies. Pathak et al. [7] developed a global data-driven weather forecast model, FourCastNet, at 0.25° horizontal resolutions with a lead time from 1 day to 14 days. FourCastNet forecasts large-scale circulations with a similar level to the ECMWF weather forecast model and outperforms small-scale phenomena such as precipitation with the ECMWF weather forecast model. The forecast speed is about 45,000 times faster than the conventional physical-based weather model. Espenholt et al. [6] developed a 12 h precipitation model, MetNet-2, and investigated the fundamental shift in forecasts from a physical-based weather model to a deep learning model. This investigation demonstrated how neural networks learn forecasts. This study may open a door to co-research between the physical-based model and the deep learning model.

5. Discussion and Summary

This short review reviews how GPUs are used in HPC computers in meteorology to support beginners to choose an on-premise HPC system for each of their laboratories or teams relevant to meteorology, especially forecasts. HPC systems have accelerated weather forecasts with physical-based models since Richardson's dream [2]. Meteorological scientists and model developers have written the codes of the models by making the most of the latest HPC technologies available at the time. Several of the leading HPC systems used for weather forecast models have been introduced. Each institute chose its HPC system from many possible alternatives to best match its purposes.

The six latest topics in high-performance computing in meteorology were also overviewed: floating points; spectral transform in global weather models; heterogeneous computing; exascale computing; co-design; and data-driven weather forecasts. Each of the topics was limited to an introduction owing to the objectives of this short review. Readers interested in them are expected to gather further information from the many references that are available.

The HPC system introduced in this short review is the world's leading one. Small on-premise HPC systems can be set up for a laboratory or on-premise deployment because the latest single systems (such as NVIDIA DGX A100) have a quarter speed [45] as much as the Earth Simulator that was ranked first on the Top 500 Supercomputer List for 2.5 years from 2002 to 2004. This suggests that a small on-premise HPC system may pave the way for weather forecast studies in a laboratory or team.

ECMWF organizes a series of HPC in meteorology every two years. The latest information about state-of-the-art HPC in meteorological fields such as recent experience and achievements as well as future plans and demands can be obtained from the web [46]. The information gives hints of small HPC systems as well as studies on weather forecast models.

Funding: This work was supported by JSPS KAKENHI, grant number 20K12154, and by the Advanced Research Program for Climate Change Prediction (SENTAN), grant number JPMXD 0722680734, from the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan. In addition, it was supported by Secretaría Nacional de Ciencia, Tecnología e Innovación: EIE18-16.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data described in this short review are available from web pages on the internet and their uniform resource locators or URLs are seen in either the main body of this review or the references.

Acknowledgments: This short review is based on an invited keynote speech in the kick-off international workshop of Lanzamiento del Proyecto SENACYT, EIE18-16: “Equipamiento e Instrumentación de un Laboratorio de Investigación y Simulación Asistida por Computadoras a Diferentes Escalas y Fenomeno” at Lugar Hotel le Meridien, Panamá. I acknowledge giving an opportunity in this speech. I would like to thank anonymous reviewers who gave constructive comments to improve this paper.

Conflicts of Interest: The author declares no conflict of interest.

Appendix A. Deep Learning for Further Information

Deep learning is widely used for data-driven weather forecast models as seen in Figure 2. Table A1 tabulates the major deep learning frameworks. Further information can be obtained from the URLs. Several software and codes developed by using these frameworks are freely downloadable from GitHub (<https://github.com/>, accessed on 3 May 2022). Therefore, developers can start software based on the frameworks, not from scratch.

arXiv (<https://arxiv.org/>, accessed on 3 May 2022), a preprint server is an open access repository of electronic preprints and postprints, and provides the latest research outcomes and a major scientific competition, especially in research fields relevant to deep learning.

Table A1. List of AI and deep learning frameworks. All link sites were accessed on 3 May 2022.

Deep Learning Frameworks	URL
Tensorflow	https://www.tensorflow.org/
Keras	https://keras.io/
PyTorch	https://pytorch.org/
MXNet	https://aws.amazon.com/jp/mxnet/
CNTK	https://cntk.ai
Caffe	https://caffe.berkeleyvision.org/
PaddlePaddle	www.paddlepaddle.org/
Scikit-learn	https://scikit-learn.org/stable/
R	https://www.r-project.org/
Weka	https://www.cs.waikato.ac.nz/ml/weka/

References

1. Michalakes, J. HPC for Weather Forecasting. In *Book Parallel Algorithms in Computational Science and Engineering*; Grama, A., Sameh, A.H., Eds.; Birkhäuser Cham: Basel, Switzerland, 2020; pp. 297–323. [CrossRef]
2. Richardson, L.F. *Weather Prediction by Numerical Process*, 2nd ed.; Cambridge University Press: London, UK, 1922; p. 236. ISBN 978-0-521-68044-8.
3. Eadline, D. The evolution of HPC. In *The insideHPC Guide to Co-Design Architectures Designing Machines Around Problems: The Co-Design Push to Exascale, insideHPC*; LLC: Moscow, Russian, 2016; pp. 3–7. Available online: <https://insidehpc.com/2016/08/the-evolution-of-hpc/> (accessed on 3 May 2022).
4. Li, T.; Narayana, V.K.; El-Ghazawi, T. Exploring Graphics Processing Unit (GPU) Resource Sharing Efficiency for High Performance Computing. *Computers* **2013**, *2*, 176–214. [CrossRef]
5. Dabrowski, J.J.; Zhang, Y.; Rahman, A. ForecastNet: A time-variant deep feed-forward neural network architecture for multi-step-ahead time-series forecasting. In Proceedings of the International Conference on Neural Information Processing, Bangkok, Thailand, 18–22 November 2020; Springer: Cham, Switzerland, 2020; pp. 579–591.
6. Espeholt, L.; Agrawal, S.; Sønderby, C.; Kumar, M.; Heek, J.; Bromberg, C.; Gazen, C.; Hickey, J.; Bell, A.; Kalchbrenner, N. Skillful twelve hour precipitation forecasts using large context neural networks. *arXiv* **2021**, arXiv:2111.07470.
7. Pathak, J.; Subramanian, S.; Harrington, P.; Raja, S.; Chattopadhyay, A.; Mardani, M.; Kurth, T.; Hall, D.; Li, Z.; Azizzadenesheli, K.; et al. FourCastNet: A Global Data-driven High-resolution Weather Model using Adaptive Fourier Neural Operators. *arXiv* **2022**, arXiv:2202.11214. [CrossRef]
8. Müller, M.; Aoki, T. Hybrid Fortran: High Productivity GPU Porting Framework Applied to Japanese Weather Prediction Model. In *International Workshop on Accelerator Programming Using Directives*; Springer: Cham, Switzerland, 2017; pp. 20–41.
9. Hersbach, H.; Bell, B.; Berrisford, P.; Hirahara, S.; Horanyi, A.; Muñoz-Sabater, J.; Nicolas, J.; Peubey, C.; Radu, R.; Schepers, D.; et al. The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* **2020**, *146*, 1999–2049. [CrossRef]
10. Nakaegawa, T.; Pinzon, R.; Fabrega, J.; Cuevas, J.A.; De Lima, H.A.; Cordoba, E.; Nakayama, K.; Lao, J.I.B.; Melo, A.L.; Gonzalez, D.A.; et al. Seasonal changes of the diurnal variation of precipitation in the upper Río Chagres basin, Panamá. *PLoS ONE* **2019**, *14*, e0224662. [CrossRef]
11. ECWMF 2021a: Fact Sheet: Supercomputing at ECMWF. Available online: <https://www.ecmwf.int/en/about/media-centre/focus/2021/fact-sheet-supercomputing-ecmwf> (accessed on 3 May 2022).
12. Müller, A.; Deconinck, W.; Kühnlein, C.; Mengaldo, G.; Lange, M.; Wedi, N.; Bauer, P.; Smolarkiewicz, P.K.; Diamantakis, M.; Lock, S.-J.; et al. The ESCAPE project: Energy-efficient Scalable Algorithms for Weather Prediction at Exascale. *Geosci. Model Dev.* **2019**, *12*, 4425–4441. [CrossRef]
13. Jungclaus, J.H.; Lorenz, S.J.; Schmidt, H.; Brovkin, V.; Brüggemann, N.; Chegini, F.; Crüger, T.; De-Vrese, P.; Gayler, V.; Giorgetta, M.A.; et al. The ICON Earth System Model Version 1.0. *J. Adv. Modeling Earth Syst.* **2022**, *14*, e2021MS002813. [CrossRef]
14. Courant, R.; Friedrichs, K.; Lewy, H. Über die partiellen Differenzengleichungen der mathematischen Physik. *Math. Ann.* **1928**, *100*, 32–74. (In German) [CrossRef]
15. Schättler, U. Operational NWP at DWD Life before Exascale. In Proceedings of the 19th Workshop on HPC in Meteorology, Reading, UK, 20–24 September 2021; Available online: https://events.ecmwf.int/event/169/contributions/2770/attachments/1416/2542/HPC-WS_Schaettler.pdf (accessed on 3 May 2022).
16. Fuhrer, O.; Osuna, C.; Lapillonne, X.; Gysi, T.; Cumming, B.; Bianco, M.; Arteaga, A.; Schulthess, T.C. Towards a performance portable, architecture agnostic implementation strategy for weather and climate models. *Supercomput. Front. Innov.* **2014**, *1*, 45–62. Available online: <http://superfri.org/superfri/article/view/17> (accessed on 17 March 2018).
17. Fuhrer, O.; Chadha, T.; Hoefler, T.; Kwasniewski, G.; Lapillonne, X.; Leutwyler, D.; Lüthi, D.; Osuna, C.; Schär, C.; Schulthess, T.C.; et al. Near-global climate simulation at 1 km resolution: Establishing a performance baseline on 4888 GPUs with COSMO 5.0. *Geosci. Model Dev.* **2018**, *11*, 1665–1681. [CrossRef]
18. CSCS. Fact Sheet: “Piz Daint”, One of the Most Powerful Supercomputers in the World. 2017. p. 2. Available online: https://www.cscs.ch/fileadmin/user_upload/contents_publications/factsheets/piz_daint/FSPizDaint_Final_2018_EN.pdf (accessed on 3 May 2022).
19. Hosansky, D. *New NCAR-Wyoming Supercomputer to Accelerate Scientific Discovery*; NCAR & UCAR News: Boulder, CO, USA, 27 January 2021. Available online: <https://news.ucar.edu/132774/new-ncar-wyoming-supercomputer-accelerate-scientific-discovery> (accessed on 3 May 2022).
20. R-CCS. RIKEN Center for Computational Science Pamphlet. 2021, p. 14. Available online: https://www.r-ccs.riken.jp/en/wp-content/uploads/sites/2/2021/09/r-ccs_pamphlet_en.pdf (accessed on 3 May 2022).
21. Duc, L.; Kawabata, T.; Saito, K.; Oizumi, T. Forecasts of the July 2020 Kyushu Heavy Rain Using a 1000-Member Ensemble Kalman Filter. *SOLA* **2021**, *17*, 41–47. [CrossRef]
22. R-CCS. About Fugaku 2022. Available online: <https://www.r-ccs.riken.jp/en/fugaku/about/> (accessed on 3 May 2022).
23. JAMSTEC 2013. JAMSTEC Vision. Available online: <https://www.jamstec.go.jp/e/about/vision/> (accessed on 3 May 2022).
24. Mizuta, R.; Oouchi, K.; Yoshimura, H.; Noda, A.; Katayama, K.; Yukimoto, S.; Hosaka, M.; Kusunoki, S.; Kawai, H.; Nakagawa, M. 20-km-mesh global climate simulations using JMA-GSM model—mean climate states—. *J. Meteorol. Soc. Jpn. Ser. II* **2006**, *84*, 165–185. [CrossRef]

25. Haarsma, R.J.; Roberts, M.J.; Vidale, P.L.; Senior, C.A.; Bellucci, A.; Bao, Q.; Chang, P.; Corti, S.; Fučkar, N.S.; Guemas, V.; et al. High Resolution Model Intercomparison Project (HighResMIP v1.0) for CMIP6. *Geosci. Model Dev.* **2016**, *9*, 4185–4208. [\[CrossRef\]](#)
26. Kawase, H.; Imada, Y.; Sasaki, H.; Nakaegawa, T.; Murata, A.; Nosaka, M.; Takayabu, I. Contribution of Historical Global Warming to Local-Scale Heavy Precipitation in Western Japan Estimated by Large Ensemble High-Resolution Simulations. *J. Geophys. Res. Atmos.* **2019**, *124*, 6093–6103. [\[CrossRef\]](#)
27. Hatfield, S.; Düben, P.; Chantry, M.; Kondo, K.; Miyoshi, T.; Palmer, T. Choosing the Optimal Numerical Precision for Data Assimilation in the Presence of Model Error. *J. Adv. Model. Earth Syst.* **2018**, *10*, 2177–2191. [\[CrossRef\]](#)
28. Nakano, M.; Yashiro, H.; Kodama, C.; Tomita, H. Single Precision in the Dynamical Core of a Nonhydrostatic Global Atmospheric Model: Evaluation Using a Baroclinic Wave Test Case. *Mon. Weather Rev.* **2018**, *146*, 409–416. [\[CrossRef\]](#)
29. NVIDIA 2020a. Training, HPC up to 20x'. Available online: <https://blogs.nvidia.com/blog/2020/05/14/tensorfloat-32-precision-format/> (accessed on 3 May 2022).
30. NVIDIA 2022a. NVIDIA A100 Tensor Core GPU. pp3. Available online: <https://www.nvidia.com/content/dam/en-zz/Solutions/Data-Center/a100/pdf/nvidia-a100-datasheet-nvidia-us-2188504-web.pdf> (accessed on 3 May 2022).
31. Suda, R. Fast Spherical Harmonic Transform Routine FLTSS Applied to the Shallow Water Test Set. *Mon. Weather Rev.* **2005**, *133*, 634–648. [\[CrossRef\]](#)
32. Yoshimura, H. Improved double Fourier series on a sphere and its application to a semi-implicit semi-Lagrangian shallow-water model. *Geosci. Model Dev.* **2022**, *15*, 2561–2597. [\[CrossRef\]](#)
33. NVIDIA 2022b cuFFT Library User's Guide. p. 88. Available online: https://docs.nvidia.com/cuda/pdf/CUFFT_Library.pdf (accessed on 3 May 2022).
34. Intel 2020, Intel Executing toward XPU Vision with oneAPI and Intel Server GPU, Intel Newsroom. Available online: <https://www.intel.com/content/www/us/en/newsroom/news/xpu-vision-oneapi-server-gpu.html> (accessed on 3 May 2022).
35. Bailey, B. What Is An XPU? Semiconductor Engineering 11 November 2021. Available online: <https://semiengineering.com/what-is-an-xpu/> (accessed on 3 May 2022).
36. Reinders, J.; Ashbaugh, B.; Brodman, J.; Kinsner, M.; Pennycook, J.; Tian, X. *Data Parallel C++: Mastering DPC++ for Programming of Heterogeneous Systems Using C++ and SYCL*; Springer Nature: Berlin, Germany, 2021; p. 548.
37. Vetter, J.S. Preparing for Extreme Heterogeneity in High Performance Computing. In Proceedings of the 19th Workshop on High Performance Computing in Meteorology, Reading, UK, 20–24 September 2021. Available online: https://events.ecmwf.int/event/169/contributions/2738/attachments/1431/2578/HPC_WS-Vetter.pdf (accessed on 3 May 2022).
38. Barrett, R.F.; Borkar, S.; Dosanjh, S.S.; Hammond, S.D.; Heroux, M.A.; Hu, X.S.; Luitjens, J.; Parker, S.G.; Shalf, J.; Tang, L. On the Role of Co-design in High Performance Computing. In *Transition of HPC Towards Exascale Computing*; IOS Press: Amsterdam, The Netherlands, 2013; pp. 141–155. [\[CrossRef\]](#)
39. Cardwell, S.G.; Vineyard, C.; Severa, W.; Chance, F.S.; Rothganger, F.; Wang, F.; Musuvathy, S.; Teeter, C.; Aimone, J.B. Truly Heterogeneous HPC: Co-Design to Achieve what Science Needs from HPC. In Proceedings of the Smoky Mountains Computational Sciences and Engineering Conference, Oak Ridge, TN, USA, 26–28 August 2020; Springer: Cham, Switzerland, 2020; pp. 349–365.
40. Sato, M.; Kodama, Y.; Tsuji, M.; Odajima, T. Co-Design and System for the Supercomputer “Fugaku”. *IEEE Micro* **2021**, *42*, 26–34. [\[CrossRef\]](#)
41. IPCC. *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*; Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S.L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M.I., et al., Eds.; Cambridge University Press: Cambridge, UK; New York, NY, USA, 2021. [\[CrossRef\]](#)
42. Nakaegawa, T.; Kitoh, A.; Murakami, H.; Kusunoki, S. Annual maximum 5-day rainfall total and maximum number of consecutive dry days over Central America and the Caribbean in the late twenty-first century projected by an atmospheric general circulation model with three different horizontal resolutions. *Arch. Meteorol. Geophys. Bioclimatol. Ser. B* **2013**, *116*, 155–168. [\[CrossRef\]](#)
43. Mizuta, R.; Nosaka, M.; Nakaegawa, T.; Endo, H.; Kusunoki, S.; Murata, A.; Takayabu, I. Extreme Precipitation in 150-year Continuous Simulations by 20-km and 60-km Atmospheric General Circulation Models with Dynamical Downscaling over Japan by a 20-km Regional Climate Model. *J. Meteorol. Soc. Jpn. Ser. II* **2022**, *100*, 523–532. [\[CrossRef\]](#)
44. JMA 2020: Changes in Numerical Forecasts over the Past 30 Years, in Sixty Years of Numerical Forecasting, 1–11. Available online: https://www.jma.go.jp/jma/kishou/known/whitep/doc_1-3-2-1/all.pdf (accessed on 3 May 2022). (In Japanese).
45. NVIDIA, 2020b: NVIDIA CEO Introduces NVIDIA Ampere Architecture, NVIDIA A100 GPU in News-Packed ‘Kitchen Keynote’. Available online: <https://blogs.nvidia.com/blog/2020/05/14/gtc-2020-keynote/> (accessed on 3 May 2022).
46. ECMWF. In Proceedings of the 19th Workshop on High Performance Computing in Meteorology, Reading, UK, 20–24 September 2021. Available online: <https://events.ecmwf.int/event/169/> (accessed on 3 May 2022).