Big Data Infrastructure for Marine Environment in the Arctic and Ice Silk Road

Bo Lin
National Marine Environment Forecasting Center, Ministry of Natural Resources, Beijing 100081, China
linbo@nmefc.cn

Abstract. In this paper, the big data infrastructure for the marine environment in the Arctic and Ice Silk Road from National Marine Environmental Forecasting Center is taken as the main line, the construction of the Arctic observation system, the basic data construction of the Ice Silk Road and the construction of the high-performance computer system are described in detail, which clearly outlines the technical framework and construction progress of the Marine environmental data sources, data acquisition and transmission, data warehouse of the Arctic and Ice Silk Road. And from the perspective of climate change, the future of the Arctic and Ice Silk Road and the trend of corresponding marine environmental monitoring, warning (alarm), security assurance services and their big data infrastructure construction are prospected roughly.

1. Introduction
The Arctic is the polar point of the Earth northern hemisphere, that is, the intersection point (90°N) of the earth surface and the northern end of the earth rotation axis, which is located in the Arctic Ocean and generally refers to the region contained within the Arctic Circle (66°33′N), while the surface in the Arctic Circle includes the entire Arctic Ocean \[1,2,3\]; located in the northern hemisphere, China is an important stakeholder in Arctic affairs, which is geographically 'near-Arctic country' and one of the countries closest to the Arctic circle on land. The natural conditions and changes in the Arctic have a direct impact on China's climate system and ecological environment, which in turn are related to China's economic interests in agriculture, forestry, fishery and Marine fields \[4\]. The Ice Silk Road refers to the Silk Road on ice proposed by the Russian government and promoted by the Chinese government and jointly built by the two sides \[5\]. Starting from the end of the East Ocean Route \[6\] of the ancient Maritime Silk Road (Sea of Japan), it enters the Arctic Ocean via the Okhotsk Sea and the Bering Sea, and then reaches the coastal port of northwest Europe westward along the northern fairway of Russia, which is the shortest sea distance from northeast Asia to northwest Europe. Ice Silk Road got its name because vast majority of its roads its roads freeze perennially (or have sheets and floating ice). The average sailing period of the freighter is less than 3 months/year, coupled with frequent storms throughout the year, and the sailing period is also shrouded by fog, which is the most dangerous sea route on the earth with the natural conditions parallel to the Antarctic fairway. The construction of big data infrastructure for the Arctic and Ice Silk Road marine environment is a strategic focus of scientific research and technical services of China's National Marine Environment Forecasting Center (NMCEF).
2. Construction of observation system for Arctic expedition

Since 1999, China has carried out 11 Arctic expeditions. While actively completing the prediction and support tasks of the past Arctic expeditions, the Forecasting Center has organized teams to participate in each Arctic expedition. The expedition team members shoulder three major tasks: scientific investigation, field support and infrastructure construction, which includes two aspects: spatial expansion and time accumulation of basic data, and location and construction of fixed and sustainable observation stations. The observation of basic data was carried out throughout the whole voyage of the research ship (team) from its home port in China to the target operation area in the Arctic (Fig.1). The spatial expansion of basic data is realized by expanding the excursion area, while the time accumulation is realized by designing and constructing the corresponding data warehouse; the site selection and construction of the observation station refers to the observation station established in the Arctic Ocean in the Arctic region, which is progressive from the Arctic Circle to the North Pole and can obtain the local environmental elements in a long-term and stable way. At present, China has built and put into operation two large-scale integrated observatories in the Arctic, one of which is the China Arctic Yellow River Observatory (2004.07), located in New Orsson (78°55′N, 11°56′E) in the Svalbar Islands, Norway \[7\], halfway between the Arctic Circle and the North Pole. Since 2013, the Forecasting Center began to participate in glacier meteorological observation at the Yellow River Station, and an automatic weather station was established in Glacier A and Glacier P in the Yellow River Station area (Fig.2), which is also the first time for human to conduct meteorological monitoring in Glacier P \[8\]; the other is the China-Iceland North Pole Station (2018.10) jointly built by China and Iceland, located in the Al-Karkh Village in the north of Iceland (65°7′N, 17°36′W) \[9\]. Located near the Arctic Circle, the station is upgraded from the “China-Iceland Joint Aurora Observatory” jointly built by China and Iceland in 2013, which is the first-level ferry for the Arctic investigation from the Arctic Circle to the Arctic point (the Yellow River Station is the intermediate ferry); the expedition has also opened up long-term observation ice station \[10\], ice-based automatic observation drifting ice station \[11\], ice-based automatic drifting weather station \[12\], short-term (1-year) buoy relay ocean and meteorological observation system \[13\] in the main functional areas of the Arctic Ocean, together with China's series of marine satellites, which constitute China's independent Arctic ocean and meteorological observation network system and become a real-time data source for the continuous research and production of Arctic and Maritime Silk Road ocean numerical forecasting. MARIS, China's key national research and development program and a key inter-governmental program for international scientific and technological innovation cooperation participated by the National Forecasting Center, has carried out comprehensive cooperation with INTAROS, the largest observation project of the EU – the EU Horizon 2020 Program \[14\], which will further enhance China's research capability and participation in the Arctic, greatly improve the overall scientific and technological level of research and observation in the Arctic, and continuously expand the shared space for observation data and information collection.

Fig.1 Schematic Diagram of Voyage Route of China's Arctic Expedition (Forecasting Center Picture)
3. Basic data construction of Ice Silk Road

Russia is the world’s largest country with a long coastline winding across Asia and Europe, accompanied by the Ice Silk Road from the North Pacific through the Arctic Ocean to the North Atlantic (Fig.3). Its Far East has a vast territory and rich natural resources, making it one of the few unexplored human settlements on earth. After China and Russia proposed to jointly build a Silk Road on ice, scientists from the two countries took the lead in responding by jointly carrying out two multidisciplinary surveys (oceanographic surveys) in the Arctic Ocean. The target areas for the first joint expedition (2016.08.19-09.20) were the Chukchi Sea and East Siberian Sea in the Arctic Ocean off the coast of Russia \[15\]; the target areas for the second joint expedition (2016.09.06-10.22) were the Chukchi Sea, East Siberian Sea and the Laptev Sea, which is also off the coast of Russia \[16\]; it is expected that in the future, China and Russia will continue to advance westwards to organize more joint investigations on the Arctic Ocean between China and Russia. The Chukchi Sea, the East Siberian Sea, the Laptev Sea, the Kara Sea and the Barents Sea on the ice Silk Road (the Arctic Ocean segment) involve the Russian territorial waters and the exclusive economic zone, which is also the first time for Chinese expeditioner to enter the waters of Russia in the Arctic Ocean. The results of the two joint expeditions have cleared up the cognitive blind area of this sea area in the history of China’s Arctic expeditions and filled up the deficiency of space in the excursion area and the blank of scientific survey data. The Chukchi Sea, the East Siberian Sea and the Laptev Sea are three key sea areas of the Arctic Northeast Passage (Russia's Northern Passage), among which the Laptev Sea is known as the "ice factory" of the Arctic Ocean. Multidisciplinary comprehensive survey data provide valuable scientific understanding and basic data support for human beings to correctly assess the marine environment, bioecology and climate evolution \[17\].

Scientific expedition and oceanographic survey are the pilots of establishing fixed stations for long-term observation on the Ice Silk Road. Ice Silk Road runs across the Atlantic fan of the Arctic Ocean. In cooperation with Norway, Iceland and other Arctic countries, China has also made remarkable achievements in ocean observation stations and basic data construction. Norway and Iceland are the sites of China's Arctic Yellow River Station and China-Iceland Arctic Research Station respectively, and the waters in the Norwegian Sea between these two countries are the last section of the Ice Silk Road in the Arctic Ocean after it goes west out of Russian waters. The two stations are both a platform for basic data acquisition and scientific research, and are also security guards to escort ships sailing on the Ice Silk Road. There are two important basic data support in the construction of the big data infrastructure of the Arctic and the ice Silk Road: one is the global ocean environment observation (monitoring and exploration) data combination (Fig.4) provided by the China Ocean Satellite Series (HY-1 satellite * 4 / HY-2 satellite B / CFOSAT); the other is Global Ocean Observing System (GOOS) data, including the Northeast Asia Regional Pilot Project of the Global Ocean Observing System (Near-
GOOS (1996) [18] and the Ocean Observing Project of the ARGO Global Ocean Observatory Network (2001) [19]. As a member country, China not only has its own online ocean observation stations, but also can share the ocean observation data of the two platforms.

Fig.3 Ice Silk Road Voyage Chart of China Oceangoing Freighter "Yongsheng" (Forecasting Center Chart)

Fig.4 (2021020112) Reanalysis of Real-time Global Sea Surface Wind Data from Satellite Data (Forecasting Center Chart)

4. Construction of high-performance computer system
The acquisition, transmission, integrated processing, storage computing, analysis and management of the fundamental data on marine environment in the remote Arctic and Ice Silk Road need to be realized by computer, communication network and other information technology means. The construction of high-performance computer [20] system is an important part of the marine environmental infrastructure of the Arctic and Ice Silk Road. The observation data of Arctic and Ice Silk Road marine environment are the original data (real time data) for the development and production of numerical forecast [21] products of Arctic and Ice Silk Road marine environment. Modern Marine forecasting has entered a new era of comprehensive technology application such as big data, cloud computing and artificial intelligence, which cannot be separated from high-performance computing [22]; the sustainable mining, analysis and machine learning of Marine environmental data in the Arctic and Ice Silk Road are more dependent on high-performance computer systems. After entering the 21st century, driven and promoted by polar scientific investigation, ocean deep-sea investigation and resource development, the construction of the Maritime Silk Road in the 21st century, the scientific research and technological innovation of global marine environment prediction and the demand for products (services), the construction of high-performance computer system in the Forecasting Center ushered in
an explosive growth in quantity and quality. He scales of high-performance computer system is rapidly expanding (Fig.5), and the high-performance computing ability is greatly improved (Fig.6). The construction of basic software such as data warehouse and numerical model of ocean forecast is also steadily advancing. The data warehouse has been rapidly reformed and innovated from the early database to the data pool [23] and data lake [24], and the polar data also has its own independent data pool; after being continuously optimized and upgraded, the numerical model is gradually developed into today's modern ocean forecasting numerical model system, which has a single-factor forecasting numerical model, multi-factor forecasting numerical model, multi-factor composite, multi-model coupling comprehensive forecasting numerical model and other models. In the Arctic and the Arctic Silk Road, and even across the Arctic Ocean, China has many other available (industry / institution) basic data sources related to the marine environment, including thematic and special data system from the China Meteorological Administration, China Seismological Bureau, Maritime Safety Bureau/Salvage Bureau/Research Institute of Water Transportation Science of Ministry of Transport of China, Fisheries Administration Bureau of Ministry of Agriculture and Rural Affairs of China, Hydrology and Water Resources Monitoring and Forecasting Center, Ministry of Water Resources of China, Marine Environment Monitoring Center, Ministry of Ecology and Environment of China, Chinese Academy of Sciences, China National Oceanographic Survey Fleet, Ocean University of China and other relevant colleges and universities. Of course, they need to go through the dedicated ground lines to be connected to the high-performance computer system (hub) of the forecast center. Based on high performance computer system, a big data visualization platform for the global marine environment in the Forecasting Center has been built, including the Arctic and the Ice Silk Road, which contains multi-data sources, multi-types and multi-dimensional massive real-time data of the marine environment, can carry out distributed storage, distributed computing and real-time monitoring of the data, and can be read and paged at random.

| System     | First time | Performance | Expansion time | Performance | Total         |
|------------|------------|-------------|----------------|-------------|---------------|
| IBM Cluster| 2012       | 18.6 TFlops 104 TB | 2014           | 28 TFlops 111 TB | 46.6 TFlops 215 TB |
| Lenovo Cluster | 2017       | 258 TFlops 480 TB | 2019           | 420 TB       | 258 TFlops 900 TB |
| GPU Cluster | 2018       | 218.4 TFlops 192 TB |                |             | 218.4 TFlops 192 TB |

Fig.5 The Composition of a High-Performance Computer System in the Forecasting Center

Fig.6 The Computing Power of the High-Performance Computer in the Forecasting Center
5. Conclusion
The latest scientific research data show that a warmer, more rainy, less ice-covered 'New Arctic' \[25\] is coming to mankind. If this scientific prediction is true, the result will not just be a reconfiguration of the Earth's climate system, but a further deterioration of the global ecosystem. Nevertheless, this is a 'Gospel' for the ice Silk Road, because the ice-free Arctic Ocean will make the Ice Silk Road a new 'maritime highway' with infinite future. Let's put it 10,000 steps back, even if this is not the case in the future, the prosperity of theiced silk road is just a matter of time to come, thanks to China's increasingly advanced shipbuilding technology and its superb ice breaking capacity. At the same time, high-performance computer systems will become a veritable supercomputing centre, shoultering the responsibility of data science and information technology services for monitoring, forecasting, early warning and security of the global marine environment, including the Arctic and the Arctic Silk Road, to make new contributions to building a community of shared future for mankind. It is one of the central tasks of the Forecasting Center for the world, for the future and for the forefront of science and technology to deepen the exploration and understanding of the Arctic, to actively participate in the development and utilization of Arctic waterways, and to ensure the safety of marine environment along the Arctic and the Ice Silk Road. The big data infrastructure construction of the Marine environment in the Arctic and Ice Silk Road is always on the way.

Acknowledgments
Thanks for the support of fund projects such as National Key R&D Program of China - National Key Project of High Performance Computing "Development of Application Software for High Performance Numerical Simulation of Marine Environment" [2016YFB0201100] and National Key R&D Program of China, Intergovernmental International Science and Technology Innovation Cooperation Key Program "Research on Multi-element Space Observation and Information Service of the Arctic Environment" [2017YFE0111700]; thanks for China National Marine Environmental Forecasting Center and its related departments, experts and teachers for their data support.

References
[1] China National Science and Technology Terms Approval Committee (Termonline) : Arctic Geography Terms (2nd Edition /2006), https://www.termonline.cn/index, Beijing, January 2021.
[2] China National Science and Technology Terms Approval Committee (Termonline) : Marine Science and Technology Terms (2nd Edition /2007), https://www.termonline.cn/index, Beijing, January 2021.
[3] China National Science and Technology Terms Approval Committee (Termonline) : Law Terms (2020), https://www.termonline.cn/index, Beijing, January 2021.
[4] The Information Office of the State Council of China, China's Arctic Policy white paper (2018.01), http://www.scio.gov.cn/zfbps/ndhf/37884/Document/1618193/1618193.htm, Beijing, January 2018.
[5] China's Belt and Road Network, The Ice Silk Road "Heats" the Arctic Circle (2020.01.21), https://baijiahao.baidu.com/s?id=1656336569122455784, Beijing, January 2020.
[6] Qinhuai Yao, Xiaopeng Hu et al. "Maritime Silk Road in the 21st Century" and New Mechanism of Regional Cooperation. [M], Shanghai Academy of Social Sciences Press, June 2018, p. 003.
[7] The Polar Exploration Office of the State Oceanic Administration of China, Organization - Yellow River Station, http://chinare.mnr.gov.cn/, Beijing, July 28, 2020.
[8] China National Marine Environmental Forecasting Center, Center News, http://www.nmefc.cn/, Beijing, May 15, 2019.
[9] Xinhuanet, (Science and Technology) The Sino-Ice Arctic Science Station Is Officially into Operation, http://www.xinhuanet.com/tech/2018-10/18/c_1123579959.htm, Beijing, January 31, 2021.
[10] China National Marine Environmental Forecasting Center, Center News, http://www.nmefc.cn/,
Beijing, September 20, 2016.

[11] China National Marine Environmental Forecasting Center, Center News, http://www.nmefc.cn/, Beijing, August 16, 2017.

[12] China National Marine Environmental Forecasting Center, Center News, http://www.nmefc.cn/, Beijing, September 5, 2018.

[13] China National Marine Environmental Forecasting Center, Center News, http://www.nmefc.cn/, Beijing, August 16, 2017.

[14] China National Marine Environmental Forecasting Center, Center News, http://www.nmefc.cn/, Beijing, October 26, 2018.

[15] The First Institute of Oceanography, Ministry of Natural Resources of China, Successful Conclusion of the First Sino-Russian Joint Arctic Scientific Expedition, http://www.fio.org.cn/news/media-detail-284.htm, Beijing, October 12, 2016.

[16] Wanming Su, The 2018 Sino-Russian Arctic Joint Scientific Expedition Has Achieved a Number of Achievements, http://www.xinhuanet.com/2018-10/30/c_1123637355.htm, Beijing, October 30, 2018.

[17] Mingju Gai, Introduction to the Global Ocean Observation System. [J] Tianjin, Marine Information, 1997, 12 (12) : 11-13.

[18] Wanming Su, The 2018 Sino-Russian Arctic Joint Scientific Expedition Has Achieved a Number of Achievements, http://www.xinhuanet.com/2018-10/30/c_1123637355.htm, Beijing, October 30, 2018.

[19] China ARGO Real-time Data Center, ARGO Global Ocean Observation Network, http://wwwargo.org.cn/index.php?m=content&c=index&a=lists&catid=42, Hangzhou, Feb. 03, 2021.

[20] Thomas Sterling, Matthew Anderson, Maciej Brodowicz, Modern System and Application Practice of High Performance Computing. [M].Translated by Zhibin Huang, Bangcheng Ai, Wubing Yang, Xiuxiao Li. Beijing, China Machine Press, March 2020, P 4-6.

[21] Jian Shi, Wenjing Zhang, Fundamentals of Ocean Numerical Forecast [M]. Beijing, China Ocean Press, June, 2019 (preface) : 001- (text) : 003.

[22] Thomas Sterling, Matthew Anderson, Maciej Brodowicz, Modern System and Application Practice of High Performance Computing. [M].Translated by Zhibin Huang, Bangcheng Ai, Wubing Yang, Xiuxiao Li. Beijing, China Machine Press, March 2020, P 1-3.

[23] Alex Gorelik, Best Practices for Big Data Lakes. [M].Translated by Jianqing Fu, Qiting Cui, Weijie Zheng, Nao Xu. Beijing, China Electric Power Press, July 2020, pp. 89-98.

[24] Alex Gorelik, Best Practices for Big Data Lakes. [M].Translated by Jianqing Fu, Qiting Cui, Weijie Zheng, Nao Xu. Beijing, China Electric Power Press, July 2020, pp. 89-98.

[25] Polar Survey Office of the State Oceanic Administration, (Information Dynamics -- Media Dynamics) 10°C Warming in 30 Years: "New Arctic" Is Forming, http://chinare.mnr.gov.cn/, Beijing, September 23, 2020.

[26] The State Council Information Office of China, China's Arctic Policy white paper, http://www.scio.gov.cn/zfbps/ndhf/37884/Document/1618193/1618193.htm, Beijing, January 2018.