The 16th Symposium on Polar Science

2－5 December 2025

National Institute of Polar Research

Research Organization of Information and Systems

Session ID

Polar Data Science

Abstracts

Conveners: Masaki Kanao, Yoshimasa Tanaka, Kunio Takahashi,
Jun‘ichi Okuno, Masayoshi Kozai, Naohiko Hirasawa,
and Akira Kadokura, (NIPR/ROIS-DS)

IDo1

Cross-Domain Data Governance

David Castle[1]
[1]*University of Victoria; World Data System*

Arctic data is generated primarily by countries in the region, for obvious reasons. For less obvious reasons, finding ways to make that data accessible has been challenging. Focusing on the data governance matters, rather than the geopolitical, it takes concerted effort to conceive of, and deliver on, means by which several countries invested in aligned research programs can nevertheless find ways to share data. For example a program of federated Arctic data search was launched over a decade ago. The World Data System International Technology Office created a stream of work that lead to POLDER (Polder Data Discovery Enhancement Research). POLDER is now under the banner Polar Data Search, with continuing work that is now focused on both polar regions. A success story in its own right, there is a relatively unknown connection between POLDER and the Super Dual Auroral Radar Network (SuperDARN) which operates radars monitoring the near-Earth space environment. As a global scientific network, SuperDARN has its own challenges in creating network conditions for its multi-country members that enable data sharing and overall interoperability. POLDER, with a slight temporal lead on the SuperDARN revisiting of its own challenges, was able to model governance options that turned out to be useful to SuperDARN. The conclusion that will be reached is that, along side technical considerations related to data stewardship, there are intertwined data governance, and matters of governance more broadly, that can be shared between scientific domains.

Arctic Data Committee's international data stewardship and coordination

Chantelle Verhey[1], Vanessa Raymond[2]
[1]*World Data System - International Technology Office*
[2]*University of Alaska Fairbanks*

The Arctic Data Committee (ADC), operating under the International Arctic Science Committee (IASC) and Sustaining Arctic Observing Networks (SAON), continues to advance open, interoperable, and ethically grounded data management practices across the polar regions. This presentation provides updates on recent initiatives and outcomes that collectively strengthen polar data stewardship.

Contributions to the International Conference on Arctic Research Planning (ICARP IV) have informed priority areas for interoperability, data accessibility, and Indigenous data sovereignty across all research priority teams. The ongoing development of the ADC Data Statement highlights the application of FAIR and CARE principles to ensure responsible and equitable data sharing. Outcomes from the Polar Data Forum and the Arctic Observing Summit (AOS) illustrate practical progress in harmonizing metadata standards, improving visualization tools, and fostering cross-network coordination. Engagement during the International Data Week (IDW) Polar Session strengthened linkages between polar research communities and the broader global data science community. The 6th Polar Data Forum included discussions from a 2 day conference and 2 day workshop marathon discussing important topics such as Ship data, Semantics and vocabularies, data storytelling, and international collaboration. Finally, lessons to be revisited by the International Polar Year (IPY) Data Task Force continue to guide best practices for upcoming large-scale initiatives. In addition, recent enhancements to the Polar Data Search platform have improved data discoverability and multidisciplinary access in a time of the advanced need for science diplomacy.

Collectively, these initiatives demonstrate the ADC's role in aligning polar data systems with global frameworks while maintaining a focus on the specific needs of Arctic and Antarctic communities. The outcomes presented underscore progress toward sustainable, equitable, and interoperable polar data infrastructures that support both scientific advancement and societal relevance.

IDo3

The Antarctic Meteorological Research and Data Center Data Repository

David Mikolajczyk[1], Matthew Lazzara[1,2], Matthew Noojin[1,2], Lee Welhouse[1,2], Karissa Shannon[1,2], Matthew Lazzara[1,2], Matthew Lazzara[1,2], Matthew Lazzara[1,2], Matthew Lazzara[1,2], and Jerry Robaidek[1,2]

[1]*Antarctic Meteorological Research and Data Center, Space Science and Engineering Center, University of Wisconsin-Madison, Madison, WI, USA*

[2]*Department of Physical Sciences, School of Sciences, Madison Area Technical College, Madison, WI, USA*

The Antarctic Meteorological Research and Data Center (AMRDC) at the University of Wisconsin-Madison (UW-Madison) recently launched its new data repository (Lazzara et al. 2025). The data repository can be found here: https://amrdcdata.ssec.wisc.edu/. The purpose of the AMRDC Data Repository (ADR) is to provide an access point for all Antarctic meteorological data. The ADR follows FAIR principles (findable, accessible, interoperable, and reusable), with each submitted dataset containing appropriate metadata and issued a Digital Object Identifier (DOI). Currently, there are over 5500 datasets available on the ADR. These datasets include scientific campaign data such as the Antarctic Circumnavigation Expedition (ACE) and the Year of Polar Prediction – Southern Hemisphere (YOPP-SH), as well as ongoing datasets like UW-Madison Automatic Weather Station (AWS) network observations, satellite composite imagery, and United States Antarctic Program (USAP) field camp data. This presentation will explore how the ADR hardware and software are managed, discuss real-time data relay of Antarctic meteorological datasets, and further discuss some aspects of the UW-Madison AWS Program.

References

Lazzara, M. A., M. G. Noojin, K. J. Shannon, D. E. Mikolajczyk, and L. J. Welhouse, 2025: An Antarctic Meteorological Data Repository. *Bull. Amer. Meteor. Soc.*, 106, E1434-E1438, https://doi.org/10.1175/BAMS-D-24-0178.1.

International collaboration activities for open sciences and data sciences by the Polar Environment Data Science Center

Masaki Kanao[1]

[1] *National Institute of Polar Reseaarch / Polar Environment Data Science Center,*

Joint Support-Center for Data Science Research, Research Organization of Information and Systems

The Polar Environment Data Science Center (PEDSC) of the Joint Support-Center for Data Science Research (DS), the Research Organization of Information and Systems (ROIS) aimed to promote opening and sharing scientific data obtained by research activities in polar regions. One of its purposes is to strengthen collaboration with universities and other communities, and to support creation of further scientific outputs and advancement of polar research. PEDSC is also expected to play a role of the national data center for polar science in Japan. In this presentation, several international collaborative activities regarding open and data sciences conducted by PEDSC are introduced.

1) Invitation to the International Strategic Advisors

A total of five famous researchers / data managers from UK, China and Australia were invited by PEDSC to serve as the ROIS International Strategy Advisor in FY2022, 2023 and 2024 (+2025 as planned). They stayed at ROIS and had fruitful meetings and discussions with the staff of ROIS-DS and gave precious advice and suggestions in terms of data management and operation of ROIS and PEDSC. They also made related research presentations at institutional seminars and international symposiums such as DSWS. These invitations made strength the collaboration not only with two institutions where they are belongings and ROIS, but also among open and data science communities in the areas of Asia and Oceania, as well as global data initiatives such as WDS and CODATA.

2) Organizing International Data Science Symposium

PEDSC organized several international conferences hosted by Japan, with many participants form overseas. 1) International Workshop on Sharing, Citation and Publication of Scientific Data across Disciplines (DSWS-2017, December 2017, Tachikawa, Tokyo), 2) International Workshop on Data Science - Present & Future of Open Data & Open Science - (DSWS-2018, November 2018, Mishima, Shizuoka), 3) International Symposium on Data Science - Global Collaboration on Data beyond Disciplines - (DSWS-2020, September 2020, online), 4) International Symposium on Data Science 2023 - Building an Open-Data Collaborative Network in the Asia-Oceania Area - (DSWS-2023, December 2023, Science Council of Japan (hybrid conference), https://ds.rois.ac.jp/article/dsws_2023). Special issues for 3) and 4) are published as the Special Collection in the CODATA Data Science Journal (https://datascience.codata.org/collections/open-data-collaborative-network).

3) SCAR - Standing Committee on Antarctic Data Management (SCADM)

The Scientific Committee on Antarctic Research (SCAR) under the International Science Council (ISC) has established the Standing Committee on Antarctic Data Management (SCADM) to discuss data management and publication in the Antarctic and exchange information regarding data activities in polar regions. SCADM has developed a data policy called the Data and Information Management Strategy of SCAR. SCADM also requests that each country involved in Antarctic observations should establish a National Antarctic Data Center (NADC). In addition to the in-person meetings once a year, including the SCAR General Meeting every two years, SCADM conducts monthly online meetings to facilitate close information exchange.

PEDSC has been served as the NADC in Japan and participates in the SCADM related activities. As a part of NADC activities, PEDSC publishes metadata and data obtained from polar regions through a Polar Science Database (http://scidbase.nipr.ac.jp/). The information registered is also forwarded to the Antarctic Master Directory (AMD) within the Global Change Master Directory (GCMD) of NASA. The GCMD aggregates metadata from SCADM countries. As related activities, SCADM collaborates with the International Arctic Science Committee (IASC) under ISC and WDS, CODATA and has been continuously holding international symposiums (Polar Data Forum; PDF) since 2013. PEDSC has been involved in organizing the PDFs from the first symposium (2013 at the National Museum of Nature and Science in Tokyo). The last PDF-VI was successfully held at Hobert, Tasmania, Australia in October 2025.

IDo5

From Local Publishing to Global Impact: Amplifying Polar Biodiversity Data through Data Standards and International Networks

Yi-Ming Gan[1], Charlie Plasman[1], Pablo Deschepper[1] and Anton Van de Putte[1]
[1]*Royal Belgian Institute of Natural Sciences, Brussels, Belgium*

Biodiversity observations from the polar regions are critical for detecting and understanding ecosystem change, yet many remain inaccessible to the wider research community. Publishing datasets using the Darwin Core standard to global repositories such as GBIF and OBIS which adhere to the FAIR (Findable, Accessible, Interoperable, Reusable) principles ensures that data are preserved, interoperable, and discoverable across disciplines. Importantly, a dataset only needs to be published once for it to be reused across multiple global infrastructures, maximizing both efficiency and impact..

Through GBIF, datasets receive persistent identifiers (DOI) and citation tracking, providing recognition for data providers in scientific publications and assessments. Within OBIS, published datasets contribute directly to the development of Essential Ocean Variables (EOVs) related data products, enabling polar biodiversity data to inform global ocean observing efforts. The same datasets are also made accessible through our data portal SCAR Antarctic Biodiversity Portal (biodiversity.aq).

For researchers, publishing in Darwin Core not only increases the visibility and longevity of their work, but also supports collaboration across disciplines and countries. By contributing to these global systems, polar scientists help ensure their data inform both science and policy. This presentation will highlight the benefits of publishing polar biodiversity datasets in FAIR manner and show how local contributions strengthen global understanding of a rapidly changing region.

IDo6

Enhancing India's Polar Data Stewardship through OSF–NPDC Integration

V Sakthivel Samy[1]

Veena Thenkanidiyoor[2]

[1]National Centre for Polar and Ocean Research, Goa, India

[2]National Institute of Technology, Goa, India

The integration of the Open Science Framework (OSF) with the Indian National Polar Data Center (NPDC) represents a critical advancement in polar data stewardship, aligning with global efforts toward open, interoperable, and transparent scientific data infrastructures. As polar regions become focal points in understanding climate change and environmental variability, the need for robust, scalable, and FAIR-aligned (Findable, Accessible, Interoperable, Reusable) data management systems has become increasingly urgent. The NPDC, operated by the National Centre for Polar and Ocean Research (NCPOR) under the Ministry of Earth Sciences (MoES), serves as the central archival and dissemination platform for scientific data generated from Indian expeditions to the Antarctic, Arctic, Himalayas, and the Southern Ocean. By integrating with OSF—a free, open-source platform that supports all stages of the research lifecycle—the NPDC significantly enhances its capabilities for data versioning, metadata standardization, provenance tracking, and long-term accessibility.

This integration provides a unified digital workspace for organizing and linking research assets such as datasets, protocols, field reports, preprints, and peer-reviewed publications. It facilitates real-time collaboration, ensures traceability of data transformations, and promotes transparency across the full spectrum of scientific workflows. Furthermore, the use of persistent identifiers (e.g., DOIs) and harmonized metadata schemas under OSF improves dataset citation and reuse while supporting automated compliance with international data-sharing standards.

From a scientific perspective, this enhanced infrastructure benefits multiple research domains, including glaciology, polar ecology, oceanography, geophysics, and atmospheric sciences. Researchers now have streamlined access to both real-time and historical datasets, enabling integrative and interdisciplinary research across national and institutional boundaries. The system's design also supports the ingestion of sensor-based data, remote sensing outputs, and observational records, allowing for dynamic updates and scalable growth of polar data holdings. Importantly, this integration positions NPDC as a leading contributor within global data governance frameworks, fostering deeper collaboration with organizations such as the Scientific Committee on Antarctic Research (SCAR), the Standing Committee on Antarctic Data Management (SCADM), CODATA, and the World Data System (WDS). It underscores India's commitment to the principles of Open Science, scientific reproducibility, and international cooperation in addressing critical Earth system challenges.

Overall, the NPDC–OSF integration serves as a model for modern data infrastructure in polar science. It demonstrates how technology-driven approaches to data management can enhance scientific transparency,

foster cross-disciplinary collaboration, and strengthen national contributions to global environmental monitoring and policy frameworks.

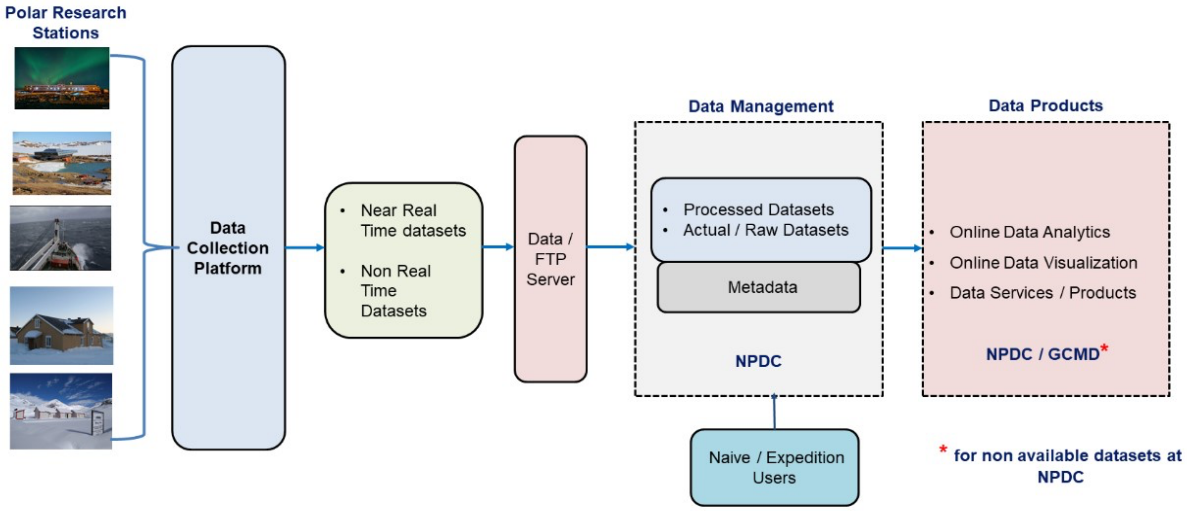


Figure 1: Overall functional view of NPDC

Figure 1 provides a detailed representation of the process through which polar datasets are collected from research stations, including Maitri, Bharati, Himadri, Himansh, and Southern Ocean expeditions. The datasets, encompassing both near real-time and non-real-time data, are transmitted to the Data/FTP server and are systematically archived at NPDC to facilitate efficient data dissemination and future accessibility for research purposes. Figure 2 provides a brief overview of the dataset submission process at NPDC, outlining the steps from the initial proposal submission by expedition members to the final submission of datasets and metadata by the principal investigator (PI) after the expedition's completion.

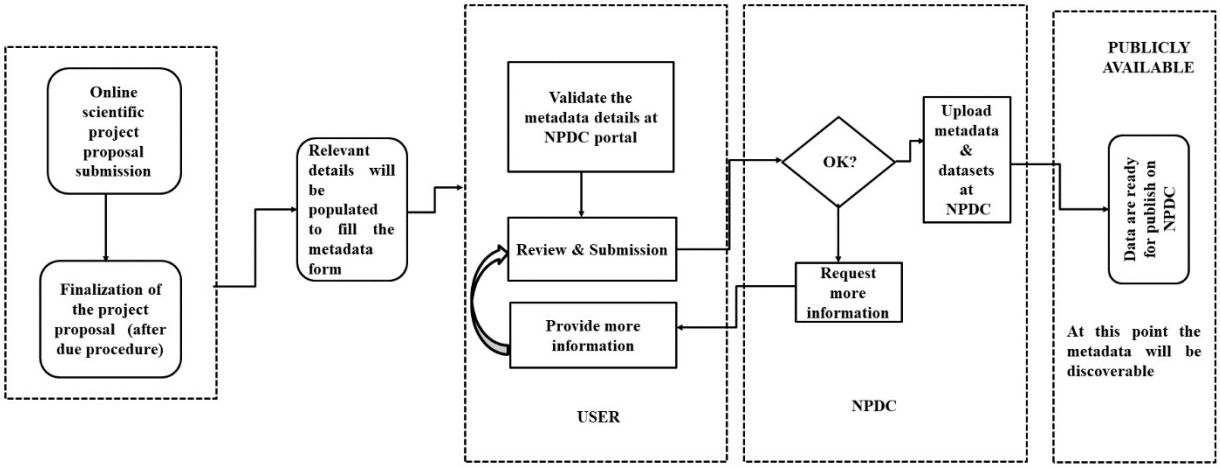


Figure 2: Flow diagram for datasets submission at NPDC

NIPR's Short-term Sea Ice Forecast Service for the Whole Arctic Ocean Based on Ice-Ocean Coupled Simulations

Y. Niwa[1*§], M. Oyama[2], T. Sugimura[1], and H. Yabuki[1]

[1] *National Institute of Polar Research (NIPR), Japan*
[2] *Graduate School of Engineering, Kogakuin University, Japan*
[§]*Corresponding author: niwa.yoshihiro@nipr.ac.jp*

With the rapid melting of sea ice in the Arctic Ocean due to global warming, the demand for accurate sea ice forecasts has been increasing. In particular, short-term sea ice forecasts – predicting sea ice distribution a few to 10 days in advance– are crucial for ship navigation, as they help identify safe and efficient routes that avoid ice-covered areas in the Arctic Ocean. Since the summer of 2024, the Arctic Sea Ice Information Center of National Institute of Polar Research (NIPR) has been providing a short-term sea ice forecast service covering the whole Arctic Ocean based on simulations using an ice-ocean coupled numerical model.

The numerical model used has a horizontal grid resolution of approximately 5 km. It is initialized using the sea ice data, the ocean temperature and salinity data from the analysis products of RIOPS (Regional Ice Ocean Prediction System), managed by Environment and Climate Change Canada. Then, the 10-day forecasts are conducted daily using the atmospheric forcing data from the forecast products of ECMWF (European Centre for Medium-Range Weather Forecasts). The resulting daily sea ice forecast maps can be viewed on the NIPR's website, "ADS VENUS for Mirai" (https://ads.nipr.ac.jp/venus.mirai/#/mirai) (Fig. 1). These maps were also provided to the R/V *Mirai* of JAMSTEC (Japan Agency for Marine-Earth Science and Technology) to support navigation and research operations during its summer Arctic research cruise.

In this presentation, we will report on sea ice forecast performance over the one-year period from summer 2024, with a focus on ice edge position errors, which is an essential metric for ensuring safe navigation in the Arctic Ocean. Across the entire Arctic Ocean, ice edge errors are estimated to range from 10 to 20 km for a 5-day forecast lead time and from 15 to 30 km for a 10-day lead time. However, in specific regions, notably the Pacific sector of the Arctic Ocean, the forecast errors occasionally become exceptionally large, highlighting the importance of real-time estimation of forecast errors to better support the Arctic maritime traffic. We will discuss the factors that contribute to the growth of such forecast errors.

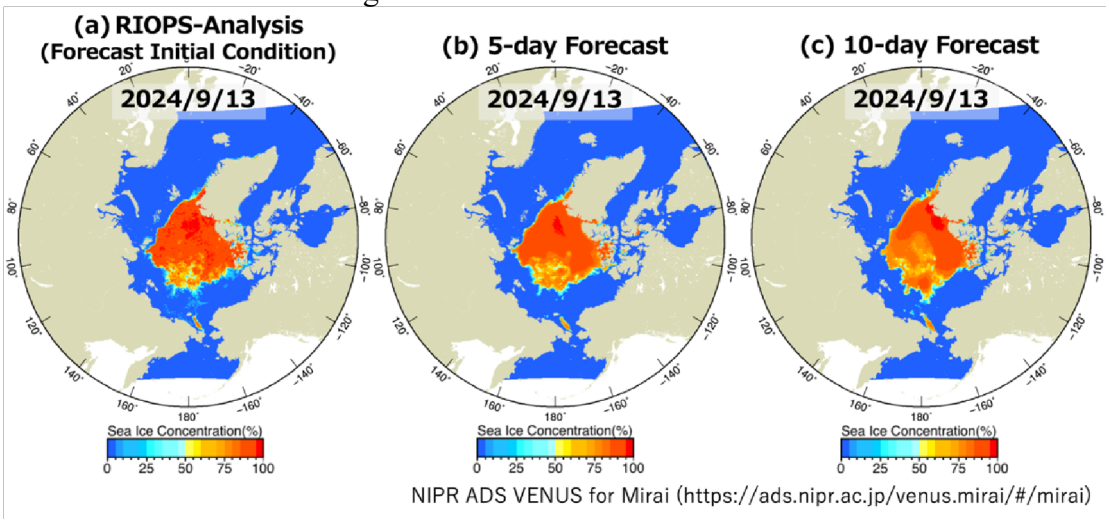


Figure 1. Examples of short-term sea ice forecast services for the whole Arctic Ocean available on the NIPR's website, "ADS VENUS for Mirai" (https://ads.nipr.ac.jp/venus.mirai/#/mirai).

15 Years of the GLISN Seismic Network in Greenland: Data Analyses and Insights

Genti Toyokuni[1], Dapeng Zhao[1], Masaki Kanao[2], Seiji Tsuboi[3], and Hiroshi Takenaka[4]
[1]*Department of Geophysics, Tohoku University, Japan*
[2]*National Institute of Polar Research, Japan*
[3]*Japan Agency for Marine-Earth Science and Technology, Japan*
[4]*Department of Earth Sciences, Okayama University, Japan*

The Greenland Ice Sheet Monitoring Network (GLISN) was launched in 2009 as an international collaboration among 11 countries to investigate Greenland's lithosphere, mantle dynamics, and cryospheric processes. Since then, the network has expanded to 34 seismic stations, including six installed and maintained by the Japanese team between 2011 and 2018. These efforts enabled long-term broadband seismic monitoring in one of the world's most inaccessible regions, yielding new insights into both solid Earth and ice sheet processes.

Observations

A major characteristic of GLISN is the installation of three new stations on the Greenland Ice Sheet (GrIS): DY2G, ICESG, and NEEM (Toyokuni et al., 2014). The installation and maintenance of these stations have been carried out by a joint USA–Japan observation team. The team was dispatched annually from 2011 to 2018 and was involved in maintaining not only the three stations on the GrIS but also three stations on outcrops. At all three newly installed stations on the GrIS, broadband seismometers (Güralp CMG-3T) designed for cold temperatures were equipped, recording three-component seismograms at a sampling rate of 100 Hz. At the DY2G and NEEM stations, borehole seismometers (Güralp CMG-3TB) were also installed at a depth of ~300 m using the vertical shaft of the ice core drilling. Additionally, GPS stations were installed near the seismometers to monitor ice flow and snow accumulation. Another characteristic of GLISN is the immediate release of data. By using satellite communication, transfer of seismic waveform data (20 sps) from the GrIS was successfully achieved in 2014, which was the first transfer of broadband seismogram from ice sheet. This made it possible to download data from all GLISN observation points in near real-time.

Achievements on Ice Sheet

By comparing the observed Rayleigh waveforms extracted from ambient noise with theoretical waveforms, high attenuation (Qp, Qs) = (20, 20) of the ice sheet was revealed for the first time through ultra-long-distance (hundreds to thousands of kilometers) propagation (Toyokuni et al., 2021). Furthermore, the detection of seasonal and long-term variations in Rayleigh-wave phase velocity suggested the possibility of basal melting of the ice sheet in northern Greenland (Toyokuni et al., 2018). We also theoretically predicted seismic phases trapped within the GrIS, and referred to it as "*Le*" wave (Toyokuni et al., 2015).

Achievements on Upper Mantle

Regional *P*-wave tomography (Toyokuni et al., 2020a) revealed that regions with basal melting of the ice sheet are located at the intersection of the thermal tracks of the Iceland plume and the Jan Mayen plume, where the crustal heat flow may singularly be elevated. A hot plume was discovered beneath the Svalbard Islands in the upper mantle, which is called the Svalbard plume. A large high-velocity body (= the northeastern rock body) was revealed in the upper mantle off the northeastern coast of Greenland. This body is thought to be the remnant of the Iapetus Ocean lithosphere, which closed 490-390 Ma, and may have influenced the spreading style of the Mid-Atlantic Ridge. The region is located at the ridge where the Farallon slab, subducted beneath North America, and the Izanagi slab, subducted beneath Eurasia, were located close to each other. A long flat slab, formed by the accumulation of lighter ridge material in the mantle transition zone (MTZ), was identified (Toyokuni & Zhao, 2023). The splitting of Greenland and Canada, along with related volcanic activity on the west coast, are considered to be caused by geodynamic processes in a big mantle wedge formed above this flat slab. Additionally, regional *P*-wave anisotropic tomography (Toyokuni & Zhao, 2021) estimated the mantle flow field around the Iceland plume.

Achievements on Lower Mantle

Global *P*-wave tomography (Toyokuni et al., 2020b) identified a hot plume (= the Greenland plume) rising from the core-mantle boundary (CMB) to the base of the MTZ beneath Greenland (Figure 1). The Jan Mayen and Svalbard plumes are considered to be branches of the Greenland plume in the upper mantle, split by the northeastern rock body. The Iceland plume is thought to be connected to the Greenland plume at two points and to another plume in Western Europe at one point. The existence of multiple heat supply routes likely explains the large number of active volcanoes in Iceland compared to its surrounding regions.

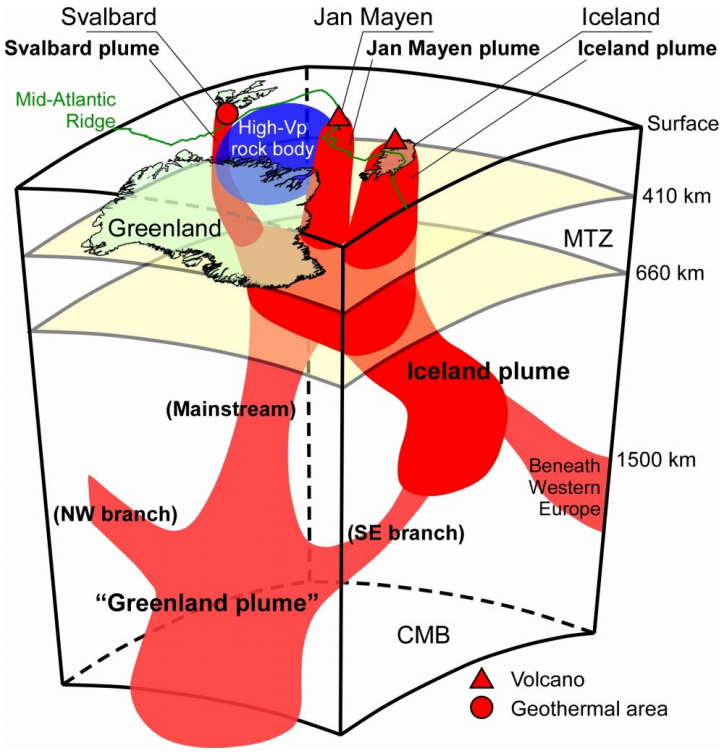


Figure 1. Schematic diagram showing relationship between the newly discovered "Greenland plume" and other plumes (Toyokuni et al., 2020b)

Through the GLISN observations, a detailed structure beneath Greenland and surrounding regions, from the ice sheet to the CMB, has been revealed. Volcanic activity, geothermal activity, and ice sheet melting are now understood as part of a unified thermal system driven by the hot plume in the lower mantle and its branches. With continued data accumulation through network maintenance, further deepening of understanding in this region can be expected.

References

Toyokuni G, Kanao M, Tono Y, Himeno T, Tsuboi S, Childs D, Anderson K, Takenaka H, Monitoring of the Greenland ice sheet using a broadband seismometer network: the GLISN project. Antarctic Record, 58(1), 1-18, 2014.

Toyokuni G, Takenaka H, Kanao M, Tsuboi S, Tono Y, Numerical modeling of seismic waves for estimating the influence of the Greenland ice sheet on observed seismograms. Polar Science, 9(1), 80-93, 2015.

Toyokuni G, Takenaka H, Takagi R, Kanao M, Tsuboi S, Tono Y, Childs D, Zhao D, Changes in Greenland ice bed conditions inferred from seismology. Physics of the Earth and Planetary Interiors, 277, 81-98, 2018.

Toyokuni G, Matsuno T, Zhao D, P wave tomography beneath Greenland and surrounding regions: 1. Crust and upper mantle. Journal of Geophysical Research: Solid Earth, 125(12), e2020JB019837, 2020a.

Toyokuni G, Matsuno T, Zhao D, P wave tomography beneath Greenland and surrounding regions: 2. Lower mantle. Journal of Geophysical Research: Solid Earth, 125(12), e2020JB019839, 2020b.

Toyokuni G, Zhao D, P-wave tomography for 3-D radial and azimuthal anisotropy beneath Greenland and surrounding regions. Earth and Space Science, 8(12), e2021EA001800, 2021.

Toyokuni G, Komatsu M, Takenaka H, Estimation of seismic attenuation of the Greenland Ice Sheet using 3-D waveform modeling. Journal of Geophysical Research: Solid Earth, 126(4), 2021.

Toyokuni G, Zhao D, Ancient slabs beneath Arctic and surroundings: Izanagi, Farallon, and in-betweens. Progress in Earth and Planetary Science, 10(1), 64, 2023.

IDo9

Inversion of antipodal PKPab waves using the adjoint method and implications for upwelling mantle plumes in polar regions

Seiji Tsuboi[1] and Rhett Butler[2]

[1]*Japan Agency for Marine-Earth Science and Technology*
[2]*University of Hawaii at Manoa*

The base of the Earth's lower mantle is characterized by large seismic velocity anomalies, known as large low-velocity provinces (LLVPs) (Garnero et al., 2016). There are several hypotheses related to the origin of LLVPs, such as remnants of Earth's early differentiation (Labrosse et al., 2007) and buried relics of proto-Earth's mantle after the Moon-forming giant impact (Yuan et al., 2023). However, the geodynamical implications, such as the role of LLVPs as driving mechanisms of plumes or subducted slabs, are not well resolved because some observations of the polarization of seismic velocity at LLVPs use the azimuthal anisotropy of shear wave splitting. Here, we combine new observations of antipodal PKPab seismic waves with the adjoint method to perform an inversion of the radially anisotropic Vp structure at the base of the lower mantle (Tsuboi and Butler, 2025). We have carefully examined antipodal stations with sufficient signal to noise ratios for both the vertical and horizontal components over the past 30 years and selected 23 source–receiver pairs with epicentral distances greater than 178.0 degrees and Mw values less than 7.0. We calculate synthetic seismograms with an accuracy of 6.9 s and perform an inversion of the radially anisotropic Vp structure at the base of the lower mantle by the adjoint method. The results of our inversion reveal that vertically polarized Vp is dominant within the LLVPs of the Pacific and African regions (Figure 1). These features are characterized by relatively small spots of high vertically polarized Vp anomalies, which may be interpreted as the locations of ascending mantle plumes inside LLVPs in the Pacific region. A similar feature is observed below Greenland, where an upwelling hotspot appears to root beneath the west coast of Greenland. This aligns with the findings of P-wave travel time tomography by Toyokuni et al (2020).

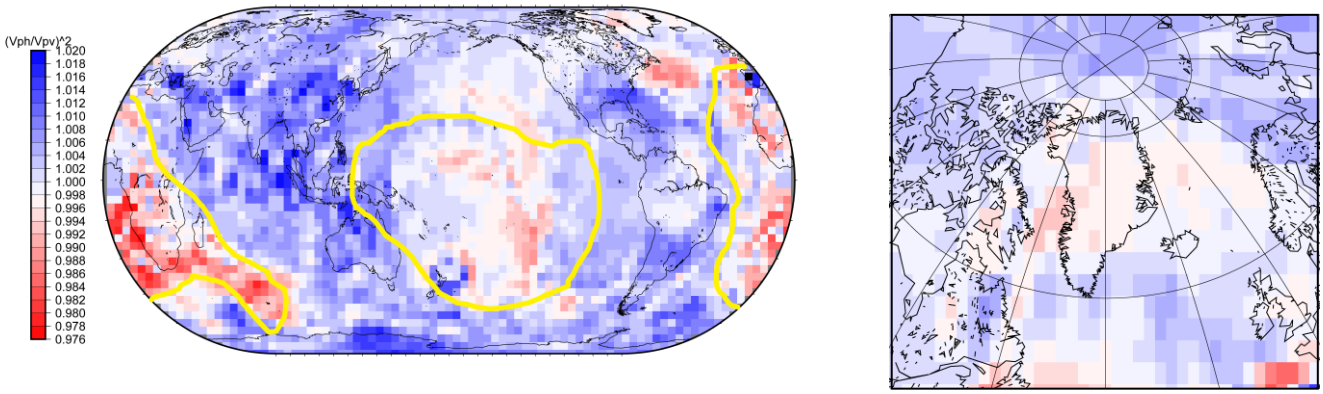


Figure 1. (left) Distribution of the radially anisotropic Vp structure at the base of the lower mantle. (Vph/Vpv)2 is plotted. The approximate locations of the LLVPs are shown as yellow lines. (right) Close up of Greenland region.

References

Garnero, E. J., McNamara, A. K. & Shim, S. H. Continent-sized anomalous zones with low seismic velocity at the base of Earth's mantle. Nat. Geosci. 9, 481–489, 2016.

Labrosse, S., Hernlund, J. W. & Coltice, N. A crystallizing dense magma ocean at the base of the Earth's mantle. Nature 450, 866–869, 2007.

Toyokuni, G., Matsuno, T., Zhao, D., P wave tomography beneath Greenland and surrounding regions: 2. Lower mantle. Journal of Geophysical Research: Solid Earth 125(12) e2020JB019839, 2020.

Tsuboi, S. and Butler, R. Inversion of antipodal PKPab waves by the adjoint method to reveal the locations of the upwelling mantle plume at the base of the mantle. Geophys. J. Int. 242, https://doi.org/10.1093/gji/ggaf251, 2025.

Yuan, Q. et al. Moon-forming impactor as a source of Earth's basal mantle anomalies. Nature 623, 95-99 2023.

IDo10

Toward Future Icy-Moon Icequake Observations: Syowa-Area Icequake Analysis and a Reproducible Catalog-Building Workflow for Penetrator-Mounted Single-Axis Seismometers.

Yasuhiro NISHIKAWA[1,2†], Satoshi TANAKA[3], Kazuto SAIKI [4], Kodai YAMAMOTO[2], Takamasa HIRATSUKA[2], Taichi KAWAMURA[5], Keisuke ONODERA[6], Masahiro MINOWA[7], Ken KONDO[8] , Naoyuki KURITA[8]

[1] *Osaka Kyoiku University*

[2] *Kochi University of Technology*

[3] *Department of Solar System Sciences, Institute of Space and Astronautical Science (ISAS), JAXA*

[4] *Ritsumeikan University*

[5] *Université Paris Cité, Institut de Physique du Globe de Paris*

[6] *Okayama University*

[7] *Hokkaido University*

[8] *Nagoya University*

† *Visiting Lecturer, Kochi University of Technology*

Seismic observations play a crucial role in planetary exploration, and upcoming missions to icy moons are no exception. One of the major challenges in extraterrestrial seismology is the deployment of a reliable seismic network on the surface of a planetary body. Penetrators have been proposed as a promising solution to this issue, enabling the installation of compact, single-axis seismometers directly into the subsurface. These penetrator-mounted instruments could serve as key components in future missions to icy worlds, facilitating the study of internal structures and geophysical processes under extreme environmental conditions. Under such mission realities, on-board compute and downlink are severely constrained, so eventlevel extraction and transmission are essential. In this study, we investigate onboard event detection algorithms optimized for capturing icequakes in future icy-moon missions. We focus on the STA/LTA method, which is widely used in terrestrial seismology, and identify an optimal set of parameters to detect icequakes with high precision. Using Antarctic analog seismic data, we develop and present a reproducible workflow for catalog generation tailored to penetrator-mounted single-axis seismometers. This workflow incorporates lightweight STA/LTA optimization, making it suitable for systems with limited computational resources.

Our analog site is the Syowa area (Telen Glacier). We analyze vertical-only continuous records, deliberately mirroring a single-axis penetrator payload. Icequake arrivals are picked with PhaseNet, then quality-controlled by two mission-compatible steps: (i) temporal non-maximum suppression (±5 s) to merge near-duplicate triggers while retaining the strongest one, and (ii) SNR in the 1–8 Hz band computed from RMS in ±2 s windows around each pick. On a ~3-day interval (mid-January 2024), a PhaseNet score threshold ≥ 0.5 yields 1,545 picks; time-domain merging consolidates them to 1,271 (−17.7%), removing all < 5 s duplicates (minimum spacing 5.05 s). We establish two sets of catalog criteria tailored to different purposes: a Base threshold (score ≥ 0.5, SNR ≥ 3) used for statistical rate analyses, and a Strict threshold (score ≥ 0.6, SNR ≥ 5) applied to highconfidence exemplars for publication-quality figures and detector calibration. Exemplars are visualized on a relative-time axis (pick = 0 s) to compare onset shape and coda behavior (Fig. 1). Because PhaseNet favors onset shape over amplitude, the combination of SNR + time merging effectively suppresses high-score/low-amplitude cases, yielding a defensible single-axis catalog-building process.

A harmonic analysis reveals strong diurnal constituents P1 (24.07 h) and K1 (23.93 h) and significant semidiurnal constituents M2 (principal lunar semidiurnal constituent ~12.42 h), N2 (12.66 h), and S2 (12.00 h) (O1 is small), indicating tidal modulation of icequake activity (Fig. 2). This modulation, together with simple single-axis spectral features (band-limited SNR, peak frequency, coda decay), guides source-family subcatalogs (tidal rift/crack, basal stick-slip, surface crevasse) and paves the way for condition-aware STA/LTA thresholds (e.g., tide-phase-dependent sensitivity) without inflating the on-board computational budget.

Our ultimate goal is a lightweight STA/LTA-based icequake detector that runs in situ under Mars/icy-moon constraints, extracting events with bounded false-alarm rates while preserving recall. Using the Strict set as a pseudo ground truth, we systematically explore STA/LTA window lengths and on/off thresholds to optimize the recall–false-alarm trade-off (FP per hour), and we test noise-adaptive scaling based on short-term RMS in 1–8 Hz to stabilize performance under varying environmental noise. Looking ahead, we will extend the same workflow to a multi-year, three-component (STS) Syowa dataset, add polarization metrics (rectilinearity, incidence, back-azimuth) and PPSD percentiles, derive detection-probability vs. noise

curves, and translate results into penetrator-ready requirements for bandwidth, sensitivity, and installation stiffness or shallow burial. By packaging inputs, processing, QC metrics, and figures as a reproducible Polar Data Science artifact, we aim to provide a transparent bridge from Syowa-area analog observations to mission-level design decisions for penetrator-mounted single-axis seismometers on icy moons.

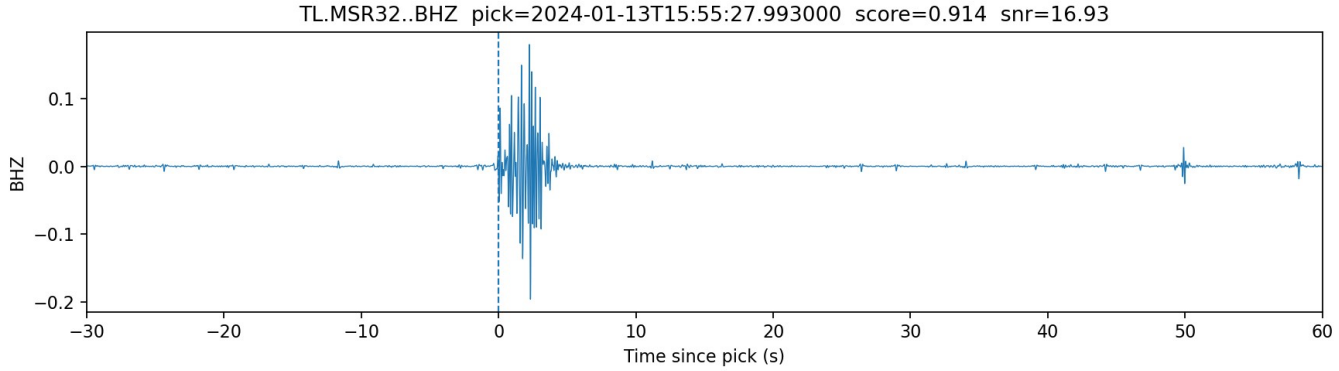


Figure 1. High-score waveform (score = 0.914) detected by PhaseNet on vertical-only single-axis records at the Telen Glacier, shown on a relative-time axis (pick = 0 s).

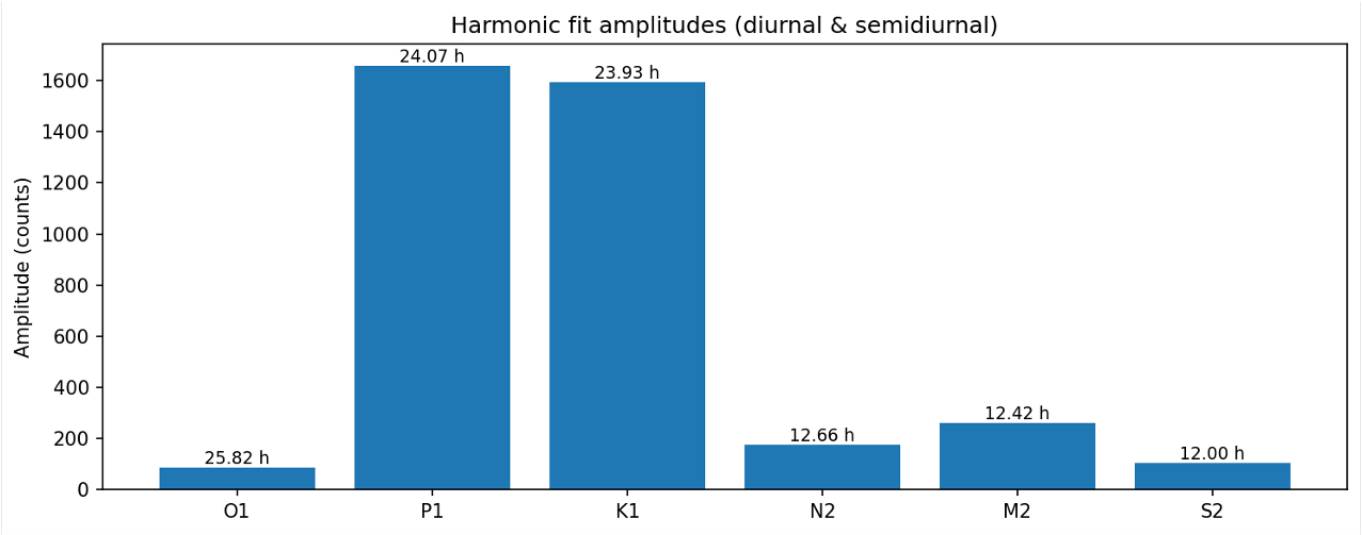


Figure 2. Detection counts with harmonic fit showing diurnal (P1, K1) and semidiurnal (M2, N2, S2) constituents. M2 is the principal lunar semidiurnal constituent (~12.42 h) and is often dominant tidal signal on Earth.

IDo11

Dataset of cosmogenic nuclide Be-7 in Iceland and Bangkok and analysis of latitude effect

Hirohisa Sakurai[1], Fuyuki Tokanai[1], Mirei Takeyama[1], Toru Moriya[1], Emiko Inui[1], Akira Kadokura[2], Masayoshi Kozai[2], Bjornsson Gunnlaugur[3], Fusa Miyake[4], Warit Mitthumsiri[5], David Ruffolo[5]

[1]*Yamagata University*

[2]*ROIS-DS,* [3]*Iceland University,*[4]*ISEE,* [5]*Mahidol University*

Beryllium-7 (Be-7), a cosmic ray-produced nuclide, is generated when cosmic rays interact with nitrogen and oxygen in the atmosphere. Since it descends with aerosols, the concentration of Be-7 in the boundary layer atmosphere (BEC) is influenced by the temporal variations of cosmic rays. Cosmic rays reaching Earth are modulated by solar activity as they pass through the heliosphere, so BEC fluctuations should contain modulation profiles corresponding to solar activity cycles such as the 27-day and 11-year cycles. On the other hand, since cosmic rays are affected by the latitude effect due to geomagnetic shielding, BEC should also be affected by the latitude effect. Furthermore, BEC varies with the global atmospheric circulation that transports aerosols. To investigate the influence of polar latitudes on seasonal variations in air mass transport, not only 27-day cycle fluctuations and the 11-year solar cycle, a daily Be-7 concentration monitoring system was installed at Husafell, Iceland (64°N). BEC has been continuously conducted at this site since September 2003. Furthermore, to investigate the influence of tropical latitudes on the seasonal variation of air mass transport, a similar daily Be-7 concentration monitoring system has been in operation since 2014 in Bangkok, Thailand (13.7°N). Figure 1 shows daily observation datasets of Be-7 concentrations in Iceland and Bangkok. Using these datasets, we describe the characteristics of BEC in polar and tropical regions through data analysis of monthly and annual variations. This information is essential for contributing to open data in the polar science when publishing datasets.

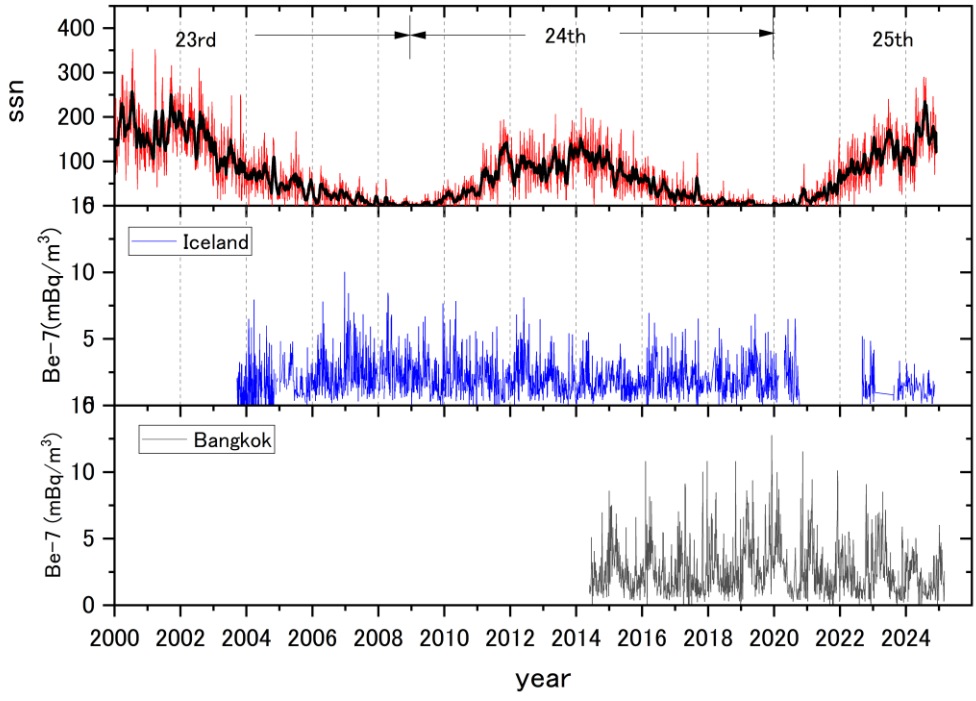


Figure 1. Daily profiles of observed Be-7 concentrations over approximately 23 years from 2003 in Iceland and 11 years in Bangkok with together sunspots number.

Exploring Large Language Models for Use in the Geosciences

Jens Klump[1], John Hille[1], Magda Guglielmo[2], and Brint Gardner[3]
[1]*Mineral Resources, CSIRO, Kensington WA, 6151, Australia*
[2]*Information Management & Technology, CSIRO, Eveleigh NSW, 2015, Australia*
[3]*Information Management & Technology, CSIRO, Clayton VIC, 3188, Australia*

For decades, we have grappled with the challenge of unstructured data, in particular, extracting information from text in scientific publications and reports. Recently, advances in Artificial Intelligence (AI) have become available as a tool to access large text corpora. Of the generative AI systems, Retrieval Augmented Generation (RAG) has attracted attention for its ability to support natural language queries into large text corpora with the help of Large Language Models (LLM). In a pilot project, we explored RAG as a tool for exploring the abstracts of the European Geosciences Union General Assembly (EGU GA) as a text corpus (Klump et al., 2025). Work by other groups has shown the potential for discovery using LLM for data analysis with promising results, but also clear indications for further work needed to make RAG a reliable tool for geoscience research (Zhang et al., 2025).

Working with the abstracts of the EGU GA gave us a large corpus of text with unique characteristics: all texts were in one language, on one broad subject, high quality, and uniform in structure and length. The abstracts were transformed to JSON and imported into a vector storage system. The RAG system integrates with the vector storage to read and use conference abstracts to provide enriched responses to natural language queries.

The RAG system's responses to natural language queries produced promising results. Including links to the source materials allowed us to compare the query response with the information in the source materials. However, evaluating generative AI models is not trivial since one query can produce multiple results. Using a well-understood text corpus and being able to trace the probable origin texts of the results allows us to evaluate the quality of the results and better understand the origin of deficient RAG responses.

The abstracts are grouped into annual volumes and thematically by the scientific divisions of the EGU GA, which allows us to study the responses of a RAG system to incomplete data, e.g., only a single annual volume. With the abstract volumes spanning a quarter-century, we can also study the response of RAG systems to changes in opinion and the evolution of knowledge in the fields covered by the EGU GA.

Based on the lessons learned from our work on the EGU GA abstracts, we plan to expand our work to include further specialised text corpora and to use RAG to generate draft reports, augment sample and drill core descriptions, and support the generation of metadata in data curation workflows.

References

Klump, J., Hille, J., Guglielmo, M., & Gardner, B. (2025). Building a RAG system for querying a large corpus of conference abstracts. In EGU General Assembly 2025 (pp. EGU25-7690). Vienna, Austria: Copernicus Meetings. https://doi.org/10.5194/egusphere-egu25-7690

Zhang, J., Clairmont, C., Que, X., Li, W., Chen, W., Li, C., & Ma, X. (2025). Streamlining geoscience data analysis with an LLM-driven workflow. Applied Computing and Geosciences, 25, 100218. https://doi.org/10.1016/j.acags.2024.100218

Development of a Cross-Disciplinary and Integrated Data Analysis Support System for Polar Research Using Large Language Models

Qi Zhang[1], Masayoshi Kozai[1], Yoshimasa Tanaka[1,2,3], Masaki Kanao[1,2,3] and Akira Kadokura[1,2]
[1] *Joint Support-Center for Data Science Research(ROIS-DS), Research Organization of Information and Systems*
[2]*National Institute of Polar Research*
[3]*The Graduate University for Advanced Studies, SOKENDAI*

Polar science is an essential field for understanding global challenges such as climate change, ecosystem shifts, and resource–environment issues, and its importance has been increasing in recent years. The polar science community in Japan, centered around the National Institute of Polar Research, has been conducting extensive observations and studies in both the Antarctic and Arctic regions, generating diverse datasets across multiple disciplines. However, because these datasets vary widely in format and content, integrated analysis requires considerable expertise and effort. This study aims to develop a "cross-disciplinary data extraction and analysis support system" that leverages large language models (LLMs), a form of generative AI that has recently gained significant attention. With this system, researchers will be able to query multiple databases in natural language and seamlessly extract, integrate, and present information. The system is expected to enhance the efficiency of data utilization in polar science, facilitate the creation of new research outcomes, and strengthen collaboration among researchers both in Japan and abroad.

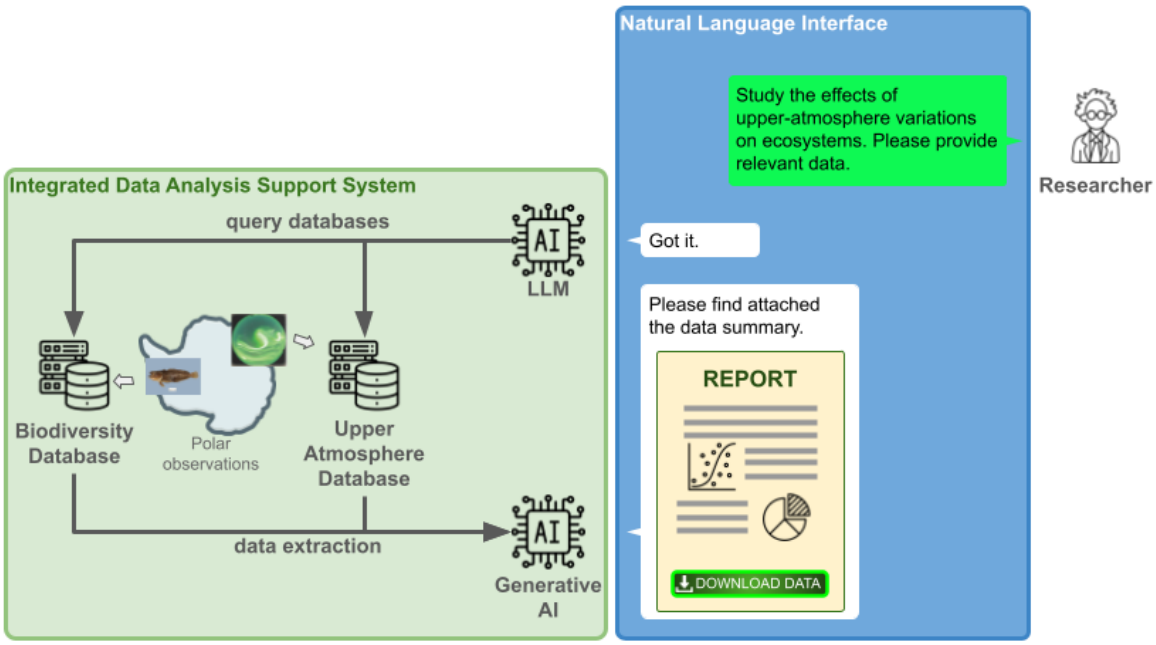


Figure 1. Conceptual overview of the cross-disciplinary data analysis support system, featuring a natural language interface for user queries and a backend module for database access and data extraction.

IDp1

Overview of Antarctic Meteorology and Climate Data: Meteorology, Glaciology, and Satellite Observations

Naohiko Hirasawa[1,2], Keiichiro Hara[3], Fumio Nakazawa[1,2], Masaki Okada[1], and collaborators†

[1]*National Institute of Polar Research*
[2]*Sokendai (The Graduate University for Advanced Studies)*
[3]*Fukuoka University*

This presentation provides an overview of long-term observations and data publications related to meteorology and climate monitoring conducted over several decades at Syowa Station, aboard the *Shirase*, and on the Antarctic ice sheet, with a focus on progress made since last year. Specifically, it will cover: (1) aerosol and cloud observations at Syowa Station and aboard the *Shirase*, (2) precipitation and cloud observations at Syowa Station, (3) Automatic Weather Station (AWS) observations on the Antarctic ice sheet, and (4) satellite observations at Syowa Station.

Some of these observations have been initiated only recently (e.g., enhanced AWS deployments and a Doppler radar), while others have been recently suspended (e.g., aerosol and cloud observations at Syowa Station and aboard the *Shirase*). Certain aerosol and cloud observations will be resumed for a limited period as part of new project activities.

Regarding data publication, quality-controlled datasets are now in preparation. However, some datasets currently provide only raw data (e.g., AWS), and others have not yet established a complete publication framework due to the large data volume (e.g., satellite observations). Most datasets are available via the "Science Database"[R1] and "Arctic and Antarctic Data archiving System (ADS)"[R2] portals operated by the National Institute of Polar Research. The poster will also present information on external public repositories and portals that compile related observational datasets.

References
R1. Science database: https://scidbase.nipr.ac.jp/modules/site/index.php?content_id=14&ml_lang=en
R2. ADS: https://ads.nipr.ac.jp/others/

† Project contributors:
(1) **Aerosol and Cloud Observation:** Masataka Shiobara[1], Hiroshi Kobayashi (University of Yamanashi), Masanori Yabuki (Kyoto University), Masahiko Hayashi (Fukuoka University), Makoto Kuji (Nara Women's University), Takashi Yamanouchi[1]
(2) **Precipitation Observation:** Hiroyuki Konishi (Osaka Kyoiku University), Yasushi Fujiyoshi (Hokkaido University), Katsushi Iwamoto (Monbetsu City)
(3) **Automatic Weather Station (AWS) Observation:** Shun Tsutaki[1 2], Hideaki Motoyama[1], Naoyuki Kurita (Nagoya University), Konosuke Sugiura (University of Toyama)
(4) **Satellite Observation:** Hiroshi Miyaoka[1], Kazue Suzuki (Meiji University)

Study on Automatic Estimation of Cloud Spatial Distribution Using All-Sky Images

Satoshi Ishii[1], Yoshihiro Tomikawa[2, 3, 4] and Yoshimasa Tanaka[2, 3, 4]
[1]*Rikkyo University*
[2]*NIPR*
[3]*SOKENDAI*
[4]*ROIS-DS*

In Japan, meteorological observatories had long recorded cloud amount through human visual observations, but they discontinued this practice at all sites except Tokyo and Osaka in March 2024. Advances in weather radar and geostationary satellite systems motivated this change, but satellite resolution of about 2 km limits detection of small-scale clouds and local weather conditions. Visual estimates every three hours also remain subjective, making them insufficient for detailed monitoring of spatiotemporal cloud distribution. High-frequency ground-based observations can provide valuable information for climate monitoring. Many observatories often operate all-sky cameras and publish images online, but most cloud evaluations still rely on manual inspection. This study explores using all-sky images for objective cloud estimation.

We installed a Sony α6000 mirrorless camera in Shirosato Town, Ibaraki Prefecture (36.5°N, 140.3°E), removed the IR-cut filter, used a fisheye lens, and applied a long-pass filter for near-infrared observations. The camera achieves a horizontal resolution of 0.6–5.5 m/pixel for target altitudes of 1–10 km. We captured images every 15 minutes during the daytime.

We estimated daytime cloud amount by comparing observed images with 2,244 manually selected daytime clear-sky reference images. We selected reference images with similar solar zenith angles and calculated pixel brightness ratios to distinguish clear-sky and cloudy regions (Yabuki et al., 2021; Hirasawa, 2023). The ratio images highlighted cloudy areas, and histograms showed a sharp peak near 1.0 under low cloud amount and broader distributions under higher cloud amount. These results suggest that histogram shapes can be used to quantitatively estimate cloud cover.

Although this study focuses on Japan, we aim to develop a system that automatically derives cloud cover from all-sky cameras worldwide, including polar observatories, enabling automated, objective, and high-frequency monitoring and supporting future interdisciplinary research collaborations.

References

Yabuki et al., 2014, Polar Science, 8, 315-326.
Hirasawa et al., 2023, Pol. Data Jour., 7, 35–49.

IDp3

Study on material exchange between atmosphere and cryosphere in polar regions

Keiihiro HARA[1], Naohiko HIRASAWA[2], Hiroshi KOBAYASHI[3], Masanori YABUKI[4], Seiji KOGA[5], Yayoi INOMATA[6], Kiyoshi MATSUMOTO[3], Norimichi TAKENAKA[7], Hisahiro TAKASHIMA[1], Sumito MATOBA[8], Yoshinori IIZUKA[8], Sakiko ISHINO[6], Kengo SUDO[9], and Saori NISHINO[8]

[1]Fukuoka University, [2]NIPR, [3]Yamanashi University, [4]Kyoto University, [5]AIST,
[6]Kanazawa University, [7]Osaka Meteopolitan University,
[8]ITLS, Hokkaido University, [9]Nagoya University

Atmospheric material cycles play important roles in climate change. In polar regions, natural atmospheric materials such as aerosols and reactive gases are supplied from ocean and cryosphere, and deposited onto snow surface. Recently, material exchange between atmosphere and cryosphere has been concerned from atmospheric chemistry and aerosol processes in polar regions (e.g., Pratt et al., Nature Geoscience, 2013; Raso et al., PNAS, 2017; Hara et al., ACP, 2017; Bognar et al., JGR, 2020; Frey et al., ACP, 2020; Ranjithkumar et al. Elementa, 2025). From viewpoint of deposition of atmospheric materials onto snow surface, better understanding on material exchange between atmosphere and cryosphere relates to interpretation of snow chemistry and ice core records. Our project aims to elucidate material exchange between atmosphere and cryosphere, and its impact on climate in polar regions through analysis of (1) new observation data and samples taken from JARE (AP1009), and (2) reanalysis of archived observation data.

In the presentation, we will show new results such as (1) Antarctic haze, (2) sea-salt aerosol emission from snow surface on sea-ice, (3) radioactive materials, and so on.

IDp4

Anomaly detection in Polarized Micro-Pulse Lider observation data using f-AnoGAN

Kei Nakamitsu[1], Kazue Suzuki[2], Masanori Yabuki[3], Kazuyuki Nakamura[1]
[1]*School of Interdisciplinary Mathmatical Sciences, Meiji University*
[2]*Graduate School of Meiji University*
[3]*RISH, Kyoto University*

This study proposes an f-AnoGAN-based anomaly detection method for vertical distribution images of atmospheric aerosols. The data were obtained from long-term observations using a Polarized Micro-Pulse Lidar (PMPL) at Syowa Station, Antarctica. We applied this method to a PMPL dataset from 2016 (Figure 1) to detect distributions of highly concentrated aerosols, which we define as anomalies in this study. Since the model is trained exclusively on normal data, we first classified the images into normal and anomalous classes for subsequent evaluation. This classification was based on a lag analysis of particle size-resolved time-series data to identify high-concentration events. We then trained a Wasserstein GAN (WGAN) on the normal images, and subsequently, we trained an izif encoder to map these images into a latent space (Table 1). Anomalies were detected based on the reconstruction error between an input image and the image generated from its latent representation (Figure 2). The proposed method achieved the ROC-AUC score of 0.76 and the PR-AUC score of 0.21 (Table 2). The anomaly score distribution showed a slight but noticeable separation between normal and anomalous samples. These results demonstrate the potential of our method for detecting distributions of highly concentrated aerosols. Future work will focus on improving performance by enlarging the training dataset and optimizing hyperparameters.

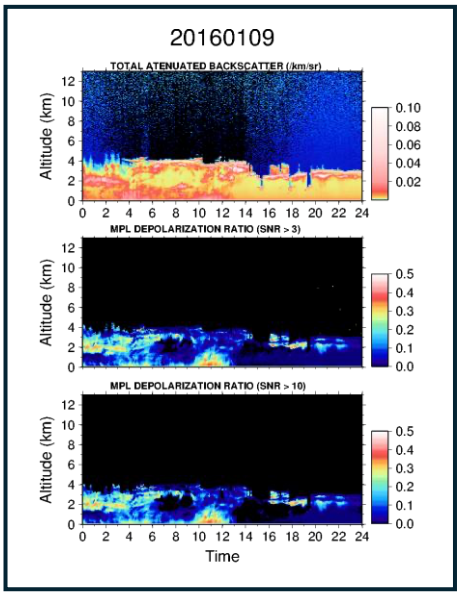


Figure 1. PMPL image from January 9, 2016

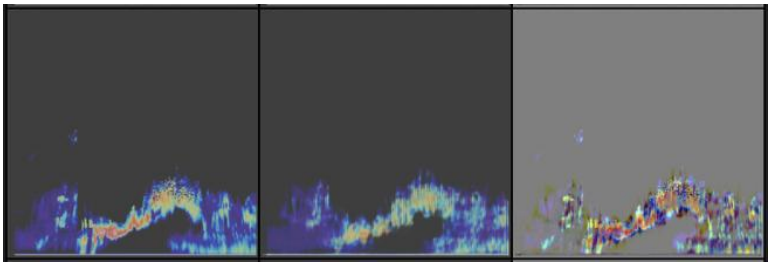


Figure 2. Reconstruction results of normal images

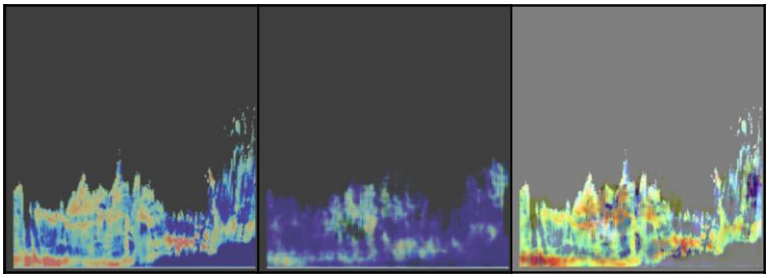


Figure 3. Reconstruction results of anomaly images

Table 1. Parameter List

	n_epochs	batch_size	lr	n_critic	latent_dim	Img_size
1	500	32	0.0004	4	64	256
2	700	32	0.0003	4	64	256
3	1400	32	0.0003	4	100	256

Table 2. Score List

	ROC-AUC	PR-AUC
1	0.703122	0.172171
2	0.705882	0.175525
3	0.743833	0.216792

References

Schlegl T, Seebock P, Waldstein SM, Langs G, Schmidt-Erfurth U.f-AnoGAN: Fast unsupervised anomaly detection with generative adversarial networks. Meidical Image Analysis, 2019.

IDp5

Last Interglacial ice volume changes inferred through statistical analysis of global sea level variations with GIA corrections

Jun'ichi Okuno[1,2,3]
[1]*Research Organization of Information and Systems*
[2]*National Institute of Polar Research*
[3]*The Graduate University of Advanced Studies, SOKENDAI*

The behavior of polar ice sheets during past warm periods provides crucial constraints for understanding their potential response to future climate warming. The Last Interglacial (LIG, ~125 ka) serves as a particularly important analog due to its global mean temperatures being 1-2°C above pre-industrial levels and sea levels 6-9 m higher than present. This study presents an integrated analysis of relative sea level (RSL) observations and numerical modeling to reconstruct ice volume variations during this critical interval.

A fundamental challenge in reconstructing LIG ice volumes from RSL records is to deconvolve spatially heterogeneous solid Earth deformation signals associated with Glacial Isostatic Adjustment (GIA) from the RSL observations. To address this challenge, we developed and implemented a high-resolution numerical model that explicitly incorporates both GIA effects and true polar wander (TPW) during the LIG. Statistical methods are employed to constrain ice volume variations by comparing observed RSL patterns with model predictions across multiple sites worldwide. The recent development of an extensive database of well-dated RSL indicators from the LIG has enabled robust constraints on polar ice sheet volume fluctuations.

Our approach employs a GIA model that incorporates TPW to evaluate spatiotemporal RSL variations during the LIG. Through sensitivity experiments, we demonstrate that ice sheet distributions during the Penultimate Glacial Maximum significantly influence global RSL model predictions through TPW. These findings advance our understanding of how the redistribution of ice masses and subsequent TPW affect sea-level changes, thereby contributing to improving future sea-level projections.

IDp6

Multidisciplinary Research Database AMIDER for Polar Science

Masayoshi Kozai[1], Yoshimasa Tanaka[1], Shuji Abe[2], Yasuyuki Minamiyama[3], Atsuki Shinbori[4], Qi Zhang[1] and Akira Kadokura[1]

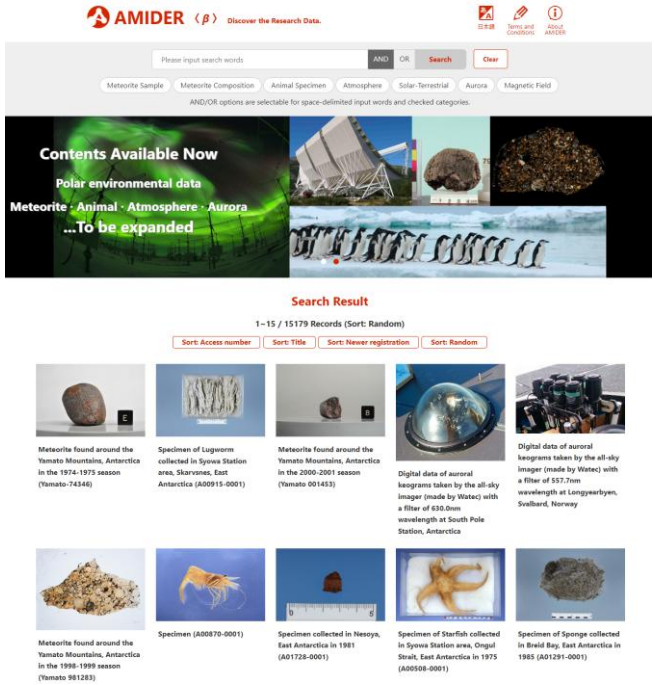
[1]*Polar Environment Data Science Center, Joint Support-Center for Data Science Research, Research Organization of Information and Systems, Tokyo, Japan*

[2]*International Research Center for Space and Planetary Environmental Science, Kyushu University, Fukuoka, Japan*

[3]*Institute of Social Science, The University of Tokyo, Tokyo, Japan*

[4]*Institute for Space-Earth Environmental Research, Nagoya University, Nagoya, Japan*

Sharing research data, such as scientific experiment data, forms a basis for studying complex systems in the polar environment. The AMIDER (Advanced Multidisciplinary Integrated-database for Exploring new Research) website was released at [1] in April 2024 to demonstrate an advanced data-sharing platform targeting non-expertized users interested in connecting to diverse scientific fields[2]. As of July 2025, over 15,000 metadata mainly in polar science, including solar-terrestrial physics data, meteorology data, meteorite samples, and animal specimens, are registered thanks to a collaboration with polar science communities. AMIDER is characterized as an integrated system applicable to different data types, such as physical observation data and specimens, and a user-friendly web application. Users can grasp diverse research data at a glance through AMIDER's catalog view (the figure below), which was inspired by web marketing and consists of thumbnail images and snippets. Multidisciplinary data curation and management are also AMIDER's uniqueness, where multiple metadata schemas are integrated, and standardized and self-describing data formats are adopted. Data curation software, such as metadata handling tools, is developed and released[3,4]. More advanced attempts, such as applying text mining techniques to metadata, are also ongoing. In this presentation, we introduce the concept, design, operation status, and experimental efforts of the AMIDER project.



References

[1] https://amider.rois.ac.jp/

[2] Masayoshi Kozai, Yoshimasa Tanaka, Shuji Abe, Yasuyuki Minamiyama, Atsuki Shinbori, Akira Kadokura, AMIDER: A Multidisciplinary Research Database and Its Application to Promote Open Science, Data Science Journal 24(7), 2025.

[3] https://doi.org/10.5281/zenodo.15590207

[4] https://doi.org/10.5281/zenodo.15621697

IDp7

The Role of GEO-Code in Historical data governance and AI application

Chandra Shekhar Roy[1], Md Biplob Khan[2], Md Maruf Hossen[3], Nayan Sarkar[4]
[1]*Bangladesh Bureau of Statistics,* [2]*Business Automation Ltd,* [3]*Synesis IT Ltd,* [4]*Nano Information Technology*

This paper explores the historical evolution and strategic importance of Geographic Entity Object (GEO) codes in Bangladesh, focusing on their application in census and survey data governance. GEO-codes are hierarchical geographic identifiers that uniquely represent administrative units from Divisions to Villages and Enumeration Areas (EAs). These codes enable precise data collection, integration, service delivery, and support for smart governance. Bangladesh Bureau of Statistics (BBS) has declared these scientific datasets as open-data for national and international use. BBS first introduced a structured GEO-Code system in 1978 in preparation for the year 1981 Population Census. This involved collecting comprehensive Mouza lists and maps in collaboration with the Land Records and Survey Department. Over time, this system has become an essential tool for managing all censuses and socio-economic surveys in the country. Historically, the GEO-Code system evolved from colonial-era Cadastral Surveys and pre-1971 East Pakistan's mouza and thana-based land records. Post-independence, it matured into a structured, digital framework linking Divisions, Districts, Upazilas, Unions, Mouzas, Villages, and EAs. This digital transformation culminated in the 2022 Digital Population and Housing Census, which successfully enumerated 58,846 mouzas and 600 Upazilas. Initially comprising 4 Divisions and 19 Districts in 1961, Bangladesh's administrative landscape has since expanded to 8 Divisions and 64 Districts. The population increased from 50.84 million in 1961 to 169.83 million in 2022. Similarly, the number of villages has risen from approximately 68,038 in 1974 to 90,049 in 2022. The paper concludes that GEO-Coding is fundamental to inclusive development, accurate data governance, and national planning. GEO-codes can support real-time analytics, AI-driven policymaking and integrated public administration across Bangladesh.

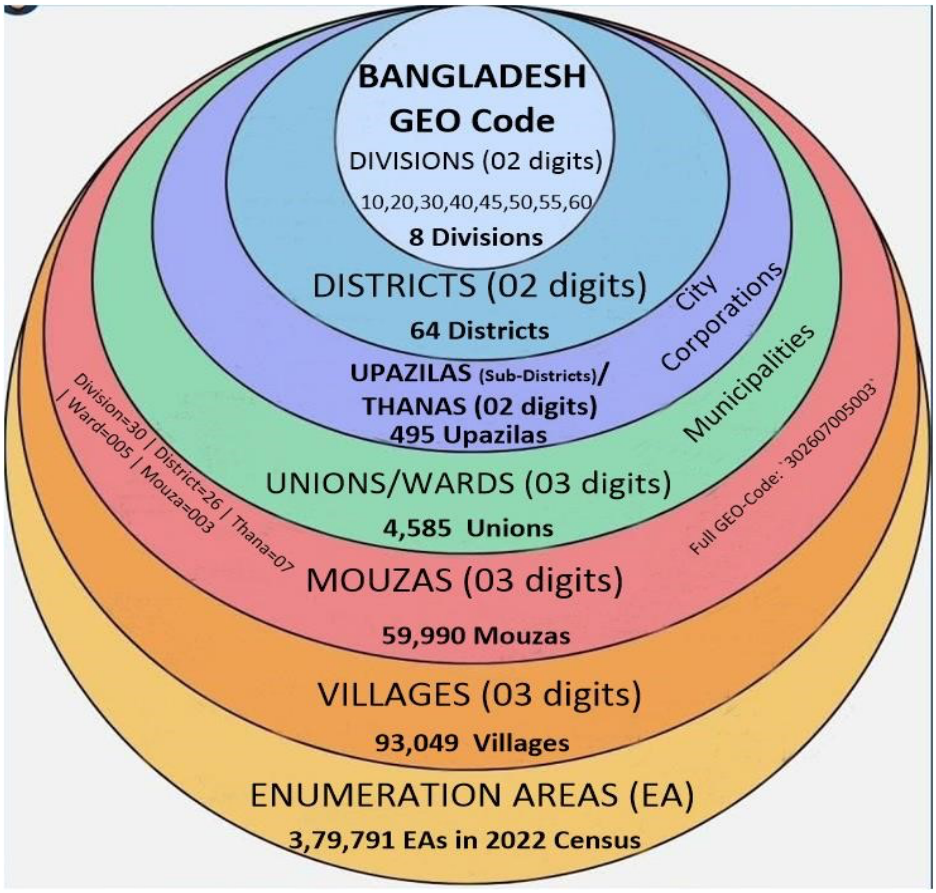


Figure 1. Administrative hierarchy and assign sequential numerical codes.

References

[1] BBS, Population and Housing Census Reports, Various Years.
[2] World Bank, Data Governance in Developing Countries, 2021.
[3] UN-GGIM, Geospatial Information Management in South Asia, 2019.
[4] Goodchild, M., "Geographic Information Science," Int. J. of Geographical Info. Science, 2020.
[5] Li, S., "Spatial Data Integration and Harmonization," Computers, Environment and Urban Systems, 2021.
[6] Zhou et al., "AI for Spatial Analytics," IEEE Transactions on Big Data, 2022.
[7] SDMX Technical Guidelines, 2020.
[8] Planning Commission of India, Geo-Spatial Governance Report, 2021.
[9] ESCAP, "Geospatial Data Governance," 2020.