Changes in the climate, ecology, water, and energy cycles on the three poles (North, South, and the Tibetan Plateau) are loosely connected to other regions through complex, long-range atmospheric and ocean transfer of energy, mass, and momentum. For example, the atmospheric heat change of the Tibetan Plateau is an important driving force for changes in the Asian monsoon and the East Asian climate. The acceleration of Arctic sea ice melting favors extreme climate events in East Asia and North America, and air–sea variations in the Antarctic affect climate anomalies in East Asia.

The changing environments have also deeply affected polar biology, such as a decline in caribou and penguin populations, while sustainable development at the three poles is an enormous challenge as human activity increases. A greater understanding of patterns, trends, and mechanisms is required to adapt to these changes. As a result, polar research is an overarching theme in the interdisciplinary, integrated multisphere study of the Earth system.

Presently, a lack of collaboration and integration in research hinders our progress in solving those problems and is the main motivation for using big data technologies for the three poles. CASEarth Poles—a project within the framework of the Chinese Academy of Sciences–supported Big Earth Data Science Engineering program—has been formally kicked off to address the challenges of polar research integration in the big data era, through data acquisition, curation and stewardship, integration,
sharing, and application, from model management to decision support.

**Big data for three poles research**

During the past decades, a large amount of valuable polar scientific data have been produced, including ground-based observations, remote sensing data, multispherical model simulations, and data assimilation products, as well as socioeconomic data. However, due to three key limitations, there are still obvious data gaps to address.

The first limitation is the lack of an efficient strategy for the acquisition, curation, exchange, and sharing of polar data. Currently, these data exist at different institutions, such as the National Snow and Ice Data Center (NSIDC) and the Chinese National Arctic and Antarctic Data Center. Each processes their polar data independently according to different strategies, standards, and goals. The lack of interoperability hinders the exchange, sharing, and utilization of polar data and further limits integrated scientific research. One positive factor is that most of these data can be accessed from open-source domains.

The second limitation is that current research methods in polar research are not skillful at dealing with multisource, multiscale, high-dimensional, and heterogeneous datasets. A novel approach focused on multifactor analysis is urgently needed. In addition, the big data paradigm has not yet been popularized in polar research and could...
reveal new scientific discoveries and decision support.

The third limitation is that the three poles have not been practically treated as an interconnected entity, partially due to limited data sharing. The result is that the effects and feedback among the three poles, as well as between the poles and global climate system, are still unclear. Scientists are only now realizing the importance of considering the three poles as interconnected.

CASEarth Poles aims to address these limitations, in part, by building a comprehensive polar big data platform for research regarding polar cryosphere, climate, hydrology, ecology, and sustainable development from an integrated and interdisciplinary perspective. The CASEarth Poles platform will support the analysis of key elements of the energy balance, mass balance, and global carbon cycle in the polar regions.

**Paths to integration**

A bilingual (e.g., Chinese and English) data sharing platform on multiple terminals has been established to archive and manage existing polar data according to an international metadata standard. A large number of new datasets will be produced via the reanalysis of ground observations, remote sensing, and model output by assimilating multiple proxies—for example, ice core, tree ring, lake sediment, marine sediment, and speleothem data and historical documents—into advanced paleoclimate models. The data repository now comprises approximately 1,400 datasets from the three poles, available at [http://poles.tpdc.ac.cn/en/](http://poles.tpdc.ac.cn/en/).

Efforts are underway to develop an interoperation with data centers or data portals of domestic and international programs on the three poles, with the ultimate objective of an integrated polar data center. In this way, CASEarth Poles will collaborate with international efforts to enable better data and information services for the three poles in the big data era.

A toolbox is being developed for processing, computing, analyzing, and visualizing multisource and multivariable big data in polar regions.
regions. In addition, a modeling environment is being constructed to contain and manage simulation and assessment models for the polar regions. In particular, an advanced Earth system model, FGOALS-f2.0, has been used in studies of data assimilation, climate change, climate variability, and interactions among climate systems over the three poles.

We plan to produce high-quality and high-resolution remote sensing products of glaciers, frozen soil, periglacial landforms, sea ice, river ice, lake ice, aerosols, and vegetation at the three poles. Innovative methods and algorithms have been developed by combining traditional remote sensing models and big data analytics. Resulting datasets include the freezing and thawing of the three-pole glacier/snow surface, glacier movement speed, crevasses in the ice shelf, snow-cover thickness, permafrost type, sea ice extent, sea ice concentration, vegetation type, vegetation index, aerosol type, and aerosol optical thickness.

With improved integration through CASEarth Poles, we plan to investigate spatio-temporal variation in the vegetation dynamics, sea ice, freeze–thaw state, and ice-shelf stability at the three poles. We will also investigate mechanisms and correlation processes between the glaciers and aerosols in polar regions; and synchronization, asynchronization, and teleconnection between key elements, such as the sea ice change in the Arctic and Antarctica.

CASEarth Poles has also set out to improve the capability of Arctic sea ice monitoring and forecasting to supply real-time information services for Arctic channel planning. The project’s efforts are ultimately expected to support better decision-making for sustainability and development in the polar regions.

---

**METADATA**

**BAMS:** What would you like readers to learn from this article?

**Xin Li (Chinese Academy of Sciences and University of Chinese Academy of Sciences):** We hope that all readers appreciate that CASEarth Poles is an integrated and useful platform that can meet the pressing need for better data acquisition, curation, integration, service, and application to support fundamental scientific research and sustainable development for the three poles.

**BAMS:** How did you become interested in the topic of this article?

**XL:** We became interested in polar research when we realized that unprecedented changes have occurred in the climate and environment over the three poles. Polar research is not only a dominating theme in the integrated multisphere study of the Earth system but is also in line with the interdisciplinary research on the nature–social science of the “Future Earth.” However, at present, the polar data are distributed and used inefficiently, and this prevents us from deeply understanding the three-pole climate and environmental change as an interconnected entity. Therefore, our research team is devoted to studying polar climate and environmental change from the perspective of polar big data.

**BAMS:** What surprised you the most about the work you document in this article?

**XL:** There is so much polar scientific data scattered in different centers around the world with various standards, and the lack of interoperation among different data centers hinders the exchange, sharing, and utilization of the data and further limits the extent and depth of integrated scientific research in the polar regions.

**BAMS:** What was the biggest challenge you encountered while doing this work?

**XL:** The biggest challenge was how to eliminate the barriers of data sharing and integrate the polar data resources scattered in various organizations into this platform. In addition, the integration of polar data in different institutions was subject to uniform standards, which is also a big challenge.

**BAMS:** What’s next?

**XL:** In the near future, we will make efforts to build up a polar information infrastructure, which will be based on the latest Earth observation and big data technology composed of the three-pole observation system and cloud storage, and applied to the interoperability of polar multisource data, online analysis, and simulation and visualization of polar environmental changes, as well as promotion of polar governance and sustainable development.