About the GEOTRACES Intermediate Data Product 2021

The *GEOTRACES Intermediate Data Product 2021* (*IDP2021*) is the result of a community effort and would not exist without the intensive collaboration of the marine geochemical community within the GEOTRACES programme and the willingness of the more than 350 data contributors to release their published or unpublished data as part of this product. In a field where sampling, calibration and measurement of the various tracers require very special skills found only in a small number of laboratories world-wide, this degree of collaboration, openness and sharing of data is exceptional.

GEOTRACES decided in 2012 to create and release a series of intermediate data products during the lifetime of the programme, before eventually producing the final, complete data product after the programme has ended. After the release of *IDP2014* and IDP2017, this is now the third such intermediate data product. The main motivation for producing intermediate data products is to not wait until the end of the programme, but instead create and release datasets at times when the programme is still very active and expanding, both, in terms of observational activities as well as the scientific analysis and synthesis of the data produced so far. By releasing and sharing the data early GEOTRACES aims at strengthening and intensifying the collaboration within the geochemical community itself, but also attracting colleagues from other communities, such as physical and biological oceanography as well as modelling, to join and devote their unique knowledge and skills to marine geochemical problems. GEOTRACES also hopes to ensure future GEOTRACES funding by showing that the programme is productive and successful.

Like the previous data products, the *IDP2021* consists of two parts: (1) the digital data and (2) the *eGEOTRACES Electronic Atlas* providing section plots and animated 3D scenes of the data.

1. Digital Data

The *IDP2021* digital data consist of five datasets: (a) *seawater discrete sample data*, (b) *seawater CTD sensor data*, (c) *aerosol data*, (d) *precipitation data* and (e) *data for snow and ice*. Each of these datasets is made available in three formats: (1) plain ASCII suitable for usage with text readers and spreadsheet software, (2) netCDF commonly used by modelers, and (3) Ocean Data View collection for use with the popular *Ocean Data View* software (https://odv.awi.de).

The digital data are available as large, complete packages at http://www.bodc.ac.uk/geotraces/data/idp2021/. As an alternative, users can retrieve customized subsets of the GEOTRACES IDP2021 data for smaller domains, individual cruises or specific parameter sets via an online data extraction service at https://geotraces.webodv.awi.de/. This site also lets users explore, analyze and visualize the data without the need to download the data or install software. In addition, the IDP2021 datasets are also made available for analysis and visualization as part of the online *webODV*

Explore site at https://explore.webodv.awi.de/ that provides easy access to many popular environmental datasets for various compartments of the earth system (navigate to *Ocean > tracer > geotraces > idp2021*). The online exploration services offer "ODV-like" interface and functionality in the web browser. All access routes require users to agree to the *IDP2021 Fair Data Use Statement*. Registration or login is no longer required.

Great efforts have been made to incorporate a rich set of metadata, describing sampling, analytical procedures and provenance of the data. Detailed information about data originator, measurement methods and associated publications are linked to the data on the individual sample level. This easily accessible extensive data context information is important, and will facilitate proper future data usage, even after the GEOTRACES programme has ended. In ODV, for instance, it only requires two mouse clicks to learn who produced a particular data value and with what methods. Another click shows the references of the associated original publications. The publication retrieval service is dynamic, including papers written and published after the release date of the IDP2021. Proper linkage and prominent display of data originator information is used throughout the *IDP2021* to provide well-deserved credit to data contributors.

The IDP2021 represents a major new data release. Compared to IDP2017, IDP2021 improves data coverage significantly in all ocean basins, especially in the Arctic, the Indian Ocean and the Pacific. Overall, the new data product contains more than twice the number of cruises, stations and actual data values. As an example, the number of dissolved Fe values has increased to more than 14,600 from about 7,500 in the IDP2017.

2. eGEOTRACES Electronic Atlas

The *eGEOTRACES Electronic Atlas* is based on the digital data described above and provides section plots and animated 3D scenes for many of the parameters. Users select tracers, cruise tracks and ocean basins using list-boxes and interactive maps. *eGEOTRACES* then presents tracer distributions along the selected sections, or animated 3D scenes showing tracer distributions along all available sections in the selected basin. Section plots and 3D animations show the names of the scientists who produced or are responsible for the data as well as links to the original publications associated with the given tracer and cruises.

eGEOTRACES provides overviews of the occurrence of geochemically relevant tracers quickly. The 3D scenes provide geographical context crucial for correctly assessing extent and origin of tracer plumes as well as for inferring processes acting on the tracers and shaping their distribution. The links to other tracers, sections and basins found at the bottom of section and 3D animation pages allow quick switching between tracers and domains and facilitate comparisons and the identification of changes. In addition to scientific usage, the *eGEOTRACES* visual material is suitable for teaching and outreach activities and can help conveying societal-relevant scientific results to interested non-scientists or decision makers.

eGEOTRACES is available online at www.egeotraces.org.

Please see the *Acknowledgements* document for the list of individuals and groups who contributed to the creation of this product. See the *Contributors* document for the list of scientists who have contributed data.

IDP2021 is the third in a series of intermediate data products. Future products will include data from new, more recent GEOTRACES cruises, but also data from older cruises for parameters that take longer to measure and complete.

Reiner Schlitzer
October 2021