



OPEN

ArcTiCA: Arctic tidal constituents atlas

DATA DESCRIPTOR

M. G. Hart-Davis¹✉, S. L. Howard², R. D. Ray³, O. B. Andersen⁴, L. Padman⁵, F. Nilsen⁶ & D. Dettmering¹

Tides in the Arctic Ocean affect ocean circulation and mixing, and sea ice dynamics and thermodynamics. However, there is a limited network of available *in situ* tidal coefficient data for understanding tidal variability in the Arctic Ocean; e.g., the global TICON-3 database contains only 111 sites above 60°N and 21 above 70°N. At the same time, the presence of sea ice and latitude limits of satellite altimetry complicate altimetry-based retrievals of Arctic tidal coefficients. This leads to a reliance on ocean tide models whose accuracy depend on having sufficient *in situ* data for validation and assimilation. Here, we present a comprehensive new dataset of tidal constituents in the Arctic region, combining analyses of *in situ* measurements from tide gauges, ocean bottom pressure sensors and GNSS interferometric reflectometry. The new dataset contains 914 measurement sites above 60°N and 399 above 70°N, with each site being quality-assessed and expert guidance provided to help maximise the usage of the dataset. We also compare the dataset to recent tide models.

Background & Summary

Ocean tides play a crucial role in large-scale ocean circulation through their role in mixing at the seabed and in the stratified interior¹. In the Arctic Ocean, tides also influence sea ice distribution, cause the formation of leads within pack ice and play an important role in mixing and stirring deep Arctic Ocean waters, especially at the boundary between shelf and deep-basin water masses^{2–4}. Around Greenland, tides contribute to mass loss from the ice sheet and the few remaining ice shelves⁵. The periodic rise and fall of ocean tides is the principal source of noise in sea-level estimations from satellite altimetry⁶, and the tidal redistribution of mass in the ocean is crucial for a variety of gravity modelling applications⁷. Tidal ranges in the Arctic Ocean are generally below one meter, but with much larger ranges in Baffin Bay, the White Sea, and the Davis Strait^{8,9}.

Our current understanding of the spatial variability of Arctic Ocean tides comes primarily from models and, more recently, from satellite altimetry. However, modelling Arctic tides is complicated by substantial regions of sparse bathymetric data¹⁰, and complex interactions between tidal currents and sea ice (including landfast ice) that dampen propagating tidal energy to modify amplitudes and shift phases of tidal coefficients^{11,12}. The seasonality of the Arctic sea ice¹² including landfast ice¹³ has been shown to impact amplitudes and phase lag estimations for tidal constituents. Additional challenges arise because much of the tidal energy in the eastern Arctic is in the form of diurnal, topographically-trapped vorticity waves¹⁴, which are sensitive to small-scale topographic variability¹⁵ and background stratification and mean flows¹⁶. While assimilation of satellite altimetry can significantly improve tidal solutions¹⁷, these data have several limitations imposed by their orbit characteristics, footprint size (for radar altimeters) and contamination by sea ice and land.

In contrast to satellite altimetry, tide gauges provide fully resolved (unaliased) measurements, usually adjacent to the coast. Tide gauges have been used for centuries for both sea-level monitoring and ocean tide predictions^{18,19} and were essential resources for early global tide modelling efforts²⁰. Even well into the altimetry era, tide gauges remain a crucial data source, particularly in the Arctic Ocean. For models purely constrained by satellite altimetry, tide gauges are a critical, independent data set for validation. Several state-of-the-art tide models also assimilate tidal analyses from *in situ* gauges, with clear benefits for model accuracy [^{9,17,21–23}, for example], especially in regions with limited altimetry coverage. *In situ* tide records also provide insights into less energetic tidal coefficients which cannot be easily addressed with modern satellite altimetry data, e.g. higher

¹Deutsches Geodätisches Forschungsinstitut, Technische Universität München, Munich, Germany. ²Earth and Space Research, Seattle, WA, USA. ³Geodesy & Geophysics Lab., NASA Goddard Space Flight Center, Greenbelt, Maryland, USA. ⁴National Space Institute, Technical University of Denmark, Kongens Lyngby, Denmark. ⁵Earth and Space Research, Corvallis, OR, USA. ⁶The University Centre in Svalbard, Longyearbyen, Norway. ✉e-mail: michael.hart-davis@tum.de

harmonics in shallow water²⁴, as well as tidal dynamics in estuarine systems²⁵. Tide gauges can provide sea-level measurements all year round and, therefore, can be used to determine the fine-structure of the tidal spectrum, including the separation of constituents of nearby frequencies (e.g., S_2 and T_2).

However, for conventional altimetry missions, which have a footprint between 100 and 250 km², this is not always possible due to gaps in the retrievals of sea level anomalies caused by the significant influence of sea ice. Tide gauges are usually limited to coastal regions, underscoring the utility of ocean bottom pressure (OBP) sensors to study ocean tides²⁶ in other regions. These data measure shelf and open ocean, providing crucial insight into tidal dynamics, especially when coupled with the coastal measurements from tide gauges. In the Arctic Ocean, the placement and maintenance of both *in situ* measurement types are often complex, not to mention costly, based on the harsh conditions of sea ice and weather patterns. Satellite altimetry has resulted in significant improvements in the modelled prediction of ocean tides in the open ocean and in shelf and coastal regions^{17,26–28}, for example]. Improved altimetry processing and spatial and temporal coverage continue to drive advances in modelled tides. However, challenges remain in the Arctic regions⁸ due to limited altimetry missions in the Arctic Ocean and the sun-synchronous orbit of many of the missions that do orbit sufficiently far north. The geopolitical sensitivity relating to bathymetry information in the Arctic is another limitation, resulting in a lack of high-resolution bathymetry data for accurate numerical modelling of ocean tides⁹.

Solving all of the issues mentioned above, as well as those not mentioned or not yet identified, to fully understand ocean tides in the Arctic Ocean requires a concerted and collaborative effort to understand the complex tidal dynamics, as well as significant advances in computational resources and improved coverage of the appropriate observational measurements. Although these are challenging prospects, a first step towards these ambitious targets is to provide a harmonised and up-to-date *in situ* dataset that models can rely on for validation or assimilation purposes. The GESLA-3 tide gauge database²⁹, used to determine tidal constituents (TICON-3³⁰), contains only 21 stations above 70°N, with the distribution of gauges being mainly around the North American and Norwegian regions. Additional individual sources of constituents in the Arctic are available, but they provide inconsistent results from one another, and they are often difficult to retrieve and interpret. Furthermore, previous data not exploited for tidal estimations, notably from OBP sensors and GNSS-IR, are also publicly available.

This manuscript describes a common, documented source of tidal constituents from multiple sources and measurement types in the Arctic Ocean. The resultant dataset – the Arctic Tidal Constituents Atlas (ArcTiCA) – considerably increases the number of observations in the Arctic Ocean region compared with previous publicly available datasets. It is provided in two easy-to-use formats, comma-separated variable (CSV) and NetCDF, and can be subsetted based on user requirements. The ArcTiCA dataset and this manuscript complement the work on tidal currents from mooring observations in the Arctic Ocean by Baumann *et al.*³¹ but with a focus on tidal heights. The focus of this dataset is the Arctic Ocean, but data is provided down to 50°N to allow for ease of use of the dataset within Arctic tide model evaluations, whose boundaries often extend that far south^{9,32,33}. ArcTiCa will be a valuable dataset for future scientific research into improved modelling of the tidal dynamics in the Arctic Ocean for applications ranging from altimetric and gravimetric corrections to data assimilation and model boundary forcings. The ArcTiCA dataset will serve as a springboard for the ocean tidal community to assess our current state-of-the-art models and advance future Arctic modelling efforts.

Finally, although we intend for this dataset to be a one-stop-shop for *in situ* tidal data in the Arctic Ocean, there are still significant gaps in the coverage of *in situ* measurements. Therefore, one aspiration of this manuscript is to motivate further concerted efforts to deploy more measurements and make existing measurements public for inclusion in updates to our dataset. As more time series are obtained, we will release updated versions of ArcTiCA aimed at increasing coverage and usability. These datasets will be documented within the README distributed with each new dataset version to allow users to find the data sources used. One additional dataset we intend to include in future iterations is the Alaska Department of Geological & Geophysical Surveys (DGGS) dataset (individual site data from <https://water-level-watch.portal.aos.org/>); based on preliminary studies, this will require improved preprocessing techniques before tidal analysis and subsequent ingestion into ArcTiCA.

Methods

Data acquisition. The starting point for creating ArcTiCA was to acquire all available tidal height data, including sea level and bottom pressure time series, published tidal coefficients, and coefficients obtained by other researchers but unpublished.

As data coverage is scarce, we chose to include all data sources regardless of quality. Where possible, time-series data were preferred as they allowed us to apply a consistent set of quality controls and analysis techniques to estimate tidal constituents. Quality control included the removal of outliers and, where necessary, converting time series from local time to standardised UTC. Where we could not locate or obtain the original time series, we extracted constituents from existing databases and literature sources to expand the overall coverage of observations. Unlike the time-series data we analysed, these can only be quality controlled directly if the respective sources provided sufficient metadata, and only post-processing steps can be applied to these data.

A total of twenty-nine different data sources were identified that provide data in the Arctic Ocean from these three measurement types (Table 1), resulting in a total of 1624 stations (Fig. 1), with 914 of the stations being above 60°N. These data sources are provided in different manners: derived by us from time-series data, provided by previously published datasets, or taken from publications. The latter sources are not ideal for quality control purposes, but we determined that they should be included in ArcTiCA based on the scarcity of data in the region. Three measurement types are included in ArcTiCA and summarised in this manuscript: tide gauges, OBP sensors, and GNSS reflectometry.

Tide gauges measure sea-level changes relative to a vertical datum, which usually varies between gauges. As their name suggests, they are used to study ocean tides, with the fixed position of these measurements and the

Source ID	Source	Source Data Type	Citation
#009	Kowalik_and_Proshutinsky_1994	Constituents	2
#010	Kulikov_etal_2018	Constituents	47
#016	Norwegian_Hydrographic_Service	Constituents and Time Series	https://www.kartverket.no/
#023	Stammer_etal_2014	Constituents	8
#025	Talibi_etal_2020	GNSS-IR Time Series	39
#001	Davis_etal_2014	Literature	48
#002	Dietrich_etal_2007	Literature	49
#004	Gjevik_and_Straume_1989	Literature	50
#005	Gjevik_etal_1992	Literature	51
#006	Greisman_etal_1986	Literature	52
#017	Peralta-Ferriz_etal_2014	Literature	53
#018	Peralta-Ferriz_2012	Literature	54
#021	Richter_etal_2011	Literature	55
#026	Voinov_2006	Literature	56
#028	Russian_Geographical_Society	Literature	https://elib.rgo.ru/
#003	Emily_Shroyer_pers_comms	OBP Time Series	Emily Shroyer, pers. comms.
#007	John_Mortensen_pers_comms	OBP Time Series	John Mortensen, pers. comms.
#008	Janout_etal_2023	OBP Time Series	57
#012	McRaven_and_Pickart_2022	OBP Time Series	58
#013	Morison_etal_2007	OBP Time Series	59
#014	Nilsen_etal_2021	OBP Time Series	60
#015	Frank_Nilsen_pers_comms	OBP Time Series	Frank Nilsen, pers. comms.
#019	Polyakov_2016	OBP Time Series	61
#020	Ray_2013	OBP Time Series	26
#022	Soren_Rysgaard_pers_comms	OBP Time Series	Soren Rysgaard, pers. comms.
#027	WHOI	OBP Time Series	44
#000	DMI	Tide Gauge Time Series	https://ocean.dmi.dk/tides ^{41–43}
#011	MEDS	Tide Gauge Time Series	https://isdm-gdsi.gc.ca/
#024	TICON-3	Tide Gauge Time Series	30

Table 1. List of sources used in the creation of ArcTiCA as well as the type of data provided.

frequent temporal sampling, usually hourly or shorter, allowing for the derivation of tidal constituents. Some older records were based on measurements of sea surface height at the times of high and low tides rather than the ocean height at regular time intervals; however, these data are not incorporated within this dataset. Tide gauges can be susceptible to non-ocean-related influences such as human interference, including instrument repositioning, vertical land motion, and instrument outages. Thousands of tide gauges have been deployed globally, although not particularly well distributed, by various institutions and governments, with several organised efforts being conducted to collect the variety of datasets based on varying applications, for example, GESLA-3²⁹ and PSMSL³⁴. Although these observations are of significant value, they are dominated by measurements at the coast, with only a few island stations globally providing insights into deep ocean tides. In the Arctic, there are no deep ocean island stations, and tides can vary substantially over the broad continental shelves of the Eurasian Arctic; therefore other data types are crucial to understanding Arctic tides.

Measurements of ocean tides by OBP sensors are a crucial part of this new tidal database. As OBP data require efforts to recover the sensor or to maintain an anchored mooring array, they are much less common than tide gauge observations, and their time series are typically shorter. OBP sensors are deployed mainly in the open ocean and shelf sea regions, making them crucial for filling regions with no tide gauge observations. There have been some previous efforts to compile tidal data from OBP measurements, including early work done under the auspices of the International Association for the Physical Sciences of the Ocean (IAPSO)^{35–37}.

The Global Undersea Pressure (GLOUP) database, now housed within the U.K. National Tidal and Sea Level Facility, has also compiled many time series. However, it was last updated in the early 2000s. A compilation of high-quality tidal constants from OBP measurements by Ray²⁶ has been used in several model validation studies^{8,27}. However, that compilation is now ten years old, and several additional data sources have since been deployed or made publicly available in the Arctic Ocean.

Another *in situ* source becoming increasingly important over recent years is GNSS interferometric reflectometry (GNSS-IR). Although not the primary intention of GNSS-IR, ocean tidal constituents have been successfully estimated from sea-level measurements from GNSS-IR stations³⁸. A study by Tabibi *et al.*³⁹ presented tidal constituents from GNSS-IR measurements in both Greenland and Antarctica. These two regions have limited coverage from other *in situ* measurements; therefore, including GNSS measurements here is valuable.

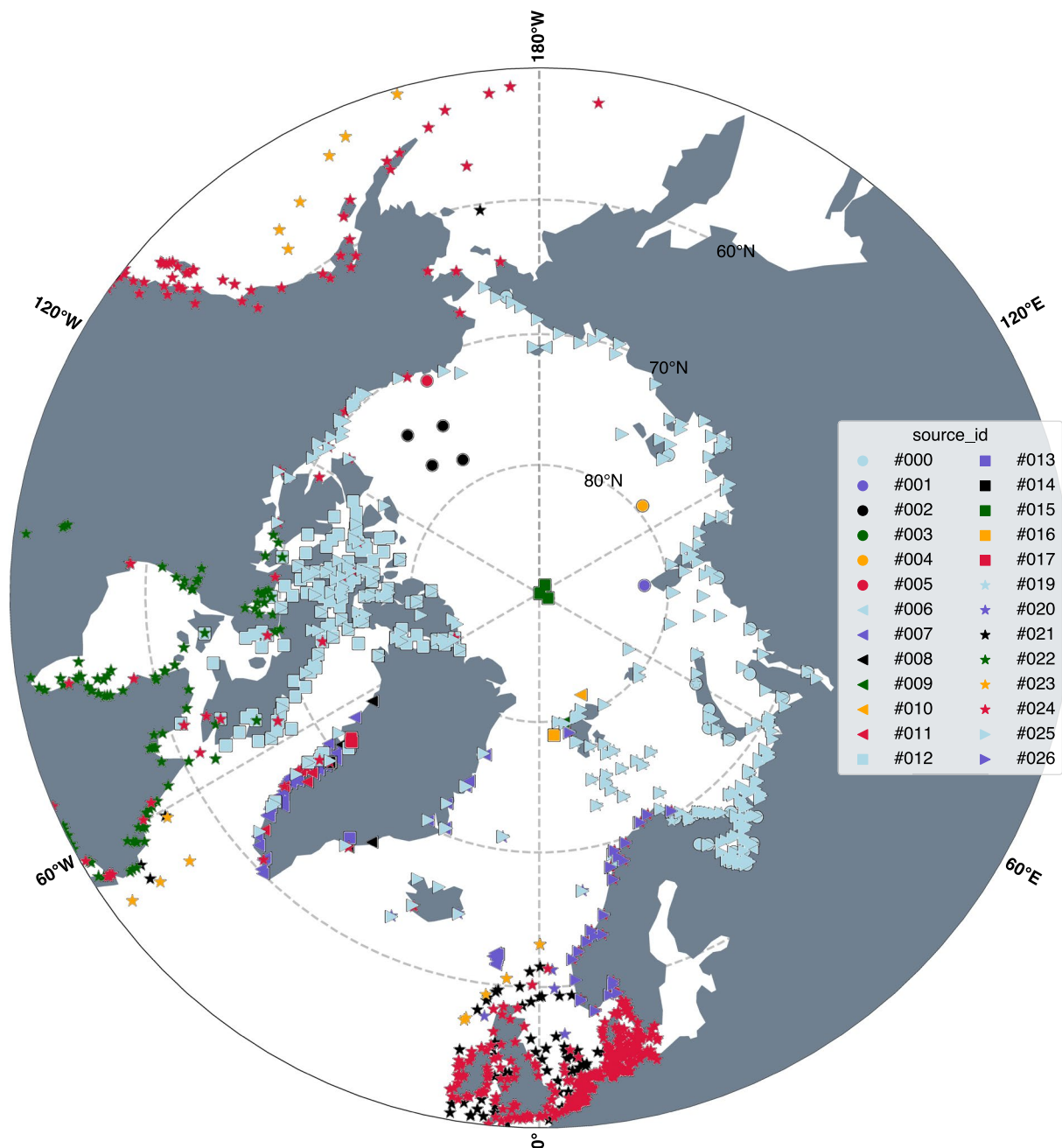


Fig. 1 Distribution of data from each data source, with the Source IDs listed in Table 1.

Data pre-processing. The methodology used in estimating tidal constituents depends on the data source. When the raw sea level or OBP data are available (for Source Data types containing ‘Time Series’), we use a common approach to the estimation of the tidal constituents. The publicly available software package *UTide*⁴⁰ was used for consistency and reliability of the results and to allow for repeatability and well-defined processes for future updating of the dataset. Experiments using our own software (not shown), such as those used in creating TICON-3³⁰, indicate that these packages provide essentially the same results for all the constituents of interest.

For the raw time-series data, outliers were removed if they deviated more than three times the standard deviation from the mean; this reduces the possibility that erroneous data will corrupt the tidal coefficient estimates. All the time series provided in local time were converted to UTC. This step is vital in preventing erroneous phase lag estimations within the dataset. The corrected time series were then used to produce suitable sets of tidal constituents based on the time-series length. As this length varies from gauge to gauge, with some being on the order of a few weeks or months and some being for multiple decades, only a certain number of tidal constituents can be estimated for all gauges. For tidal modelling efforts, the eight major tidal constituents (M_2 , S_2 , N_2 , K_2 , K_1 , O_1 , P_1 and Q_1) are of most interest in the Arctic and, therefore, when we estimate tides within this database from *in situ* time series, these constituents are always provided. The nodal modulation has been accounted for when

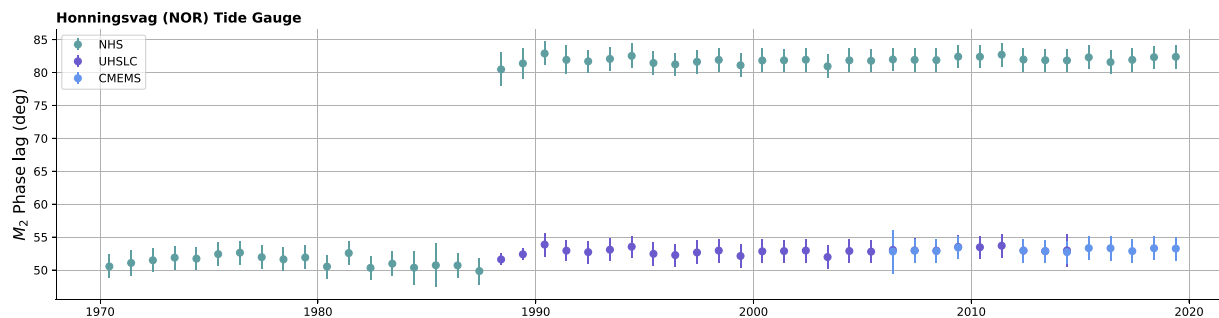


Fig. 2 Yearly M_2 phase estimations from three different GESLA-3 data sources of Honningsvag tide gauge.

deriving the constituents within this dataset using *UTide*. Note that the raw time series were not corrected for the inverse barometer effect.

For each *in situ* time series analysed by the authors, additional uncertainty information is provided based on the confidence interval information of the amplitudes and phases of each constituent, which is directly output by the *UTide* software. The confidence interval information is based on the coloured Monte Carlo method, which is the default selection within *UTide* and is further explained in⁴⁰. The uncertainty or variance information provided by other sources is also included where appropriate, with the type of information being explained with an additional variable. The uncertainty information is provided to aid in the interpretation of the data provided within the dataset. For tide gauges and GNSS-IR instruments, the units for amplitude estimations are in centimetres (cm), while the OBP-derived amplitudes are provided in millibar (mbar).

The Danish Meteorological Institute (DMI) data^{41–43} are obtained from tidal height time series, which are themselves derived from estimated tidal constituents. These constituents are not publicly available, but DMI publishes tidal heights for two years from which tidal constituents can be estimated. Note that the time-series lengths to make the original tidal height time series are not publicly available. All other time series were taken as-is from the source, with the only exception being data from the WHOI source. These data were collected and made available by the Beaufort Gyre Exploration Program based at the Woods Hole Oceanographic Institution (<https://www2.whoi.edu/site/beaufortgyre/>) in collaboration with researchers from Fisheries and Oceans Canada at the Institute of Ocean Sciences⁴⁴. These data have been collected in separate sampling efforts, usually deployed and recovered in one-year periods. To provide more reliable estimations of the tidal constituents, we combined data from these various deployments to create extended time series for the respective moorings.

Data post-processing. Before producing the final dataset, an additional quality control step was performed on the time-series data from source types OBP Time Series, Tide Gauge Time Series and GNSS-IR Time Series. This was done by dividing the time series into yearly blocks and estimating constituents for each of these yearly blocks to check consistency, especially of phase lags. An example for a Norwegian tide gauge, Honningsvag, is provided in Fig. 2, illustrating the value of this procedure. For this location, the GESLA-3²⁹ compilation includes data from three sources: the Norwegian Hydrographic Service (NHS), the University of Hawaii Sea Level Center (UHSLC), and Copernicus Marine Service (CMEMS). It is evident from Fig. 2 that there is a phase shift that occurs within the NHS estimations in 1988, and we determined that this was caused by a change in reference time from UTC pre-1988 to local time post-1988. The NHS data within GESLA-3 has this issue for several gauges (not shown). The cause of this issue is not fully diagnosed, but it occurred somehow during the construction of GESLA-3. At the time of publication, these Norwegian gauges in GESLA-3 have not been updated. Note that the NHS website also provides directly estimated constituents based on internally consistent tide gauge time series, i.e., based on one single time reference. Through a series of tests where we corrected individual gauges and compared them to constituent estimates taken directly from NHS, the use of the NHS-provided constituents was considered safe.

In some cases, tidal estimations were provided on time series in local times, which then required conversion to the UTC reference used throughout ArcTiCA. The standard formulation⁴⁵ was used for the conversion based on the difference in hours between the local time zone and UTC and the frequency of the respective tidal constituent. Where identified, this conversion was done for each station and constituent.

The *in situ* constituents were also compared with global tide models to help further identify potentially incorrect phase lag estimations. Although models have uncertainty, the phase lag estimations can provide an adequate overall picture of the individual tidal constituents. This is illustrated in Fig. 4B, where the modelled DTU16 phase lag usually demonstrates overall consistency, within estimated error ranges, with *in situ* measurements. However, there are some evident exceptions. This step was particularly valuable for data sources that provided the constituents themselves, where we could not confirm the time references from the metadata. This step helped identify that the MEDS (#011) dataset used mainly across North America was all referenced to local time and, therefore, needed to be adjusted to be consistent with UTC throughout the dataset.

Flag determination. ArcTiCA provides two flags: a ‘data’ and an ‘expert opinion’ flag. The idea behind these flags is to allow an ArcTiCA user to determine the usefulness of each data record for their own applications. As several sources provide data from the same sites, these flags also help users decide which source of measurement they prefer. In the cases of both flags, the lower the given flag, the higher the confidence in the usefulness of the

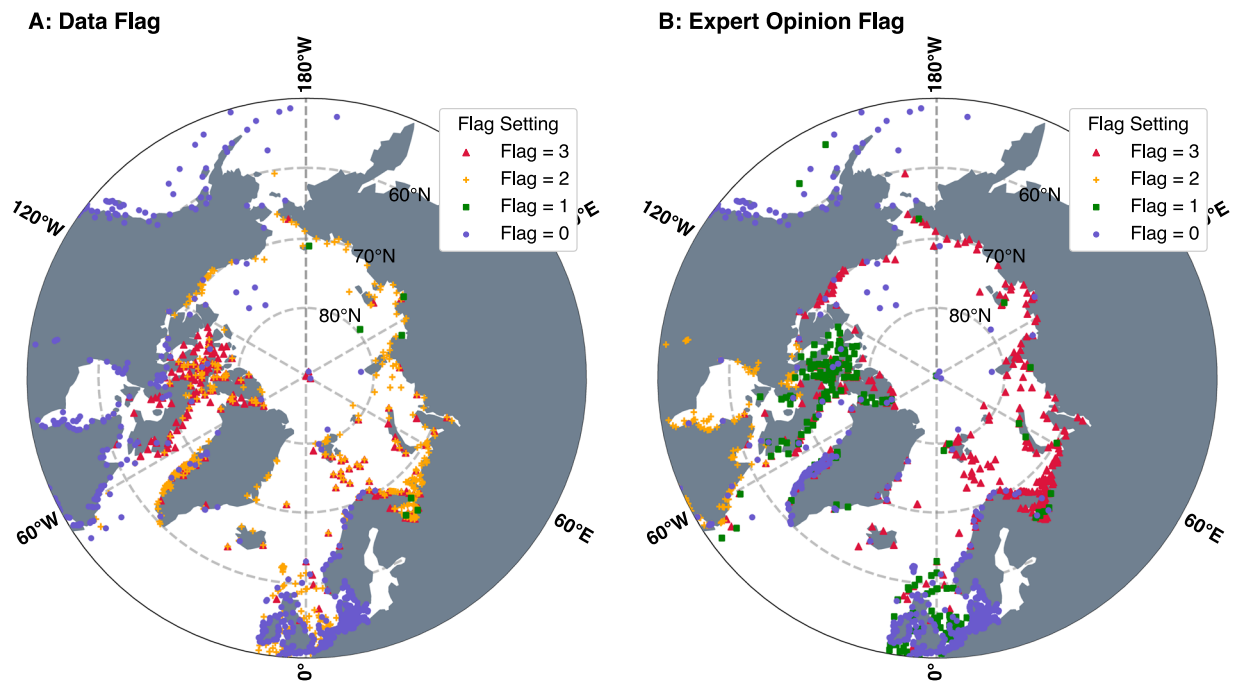


Fig. 3 An illustration of both the (A) data and (B) expert opinion flag for the entire dataset, described in Table 2.

Flag Name	Flag No.	Description
Data Flag	0	Constituents derived from a time series.
	1	Constituents obtained directly from a source.
	2	Constituents taken from published datasets or websites.
	3	Constituents obtained from tables and figures in journal publications or reports.
EO Flag	0	<i>Excellent</i> ; data and metadata is suitable for tidal estimations. Time-series length exceeds 1 year.
	1	<i>Good</i> ; time-series length is shorter than 1 year but exceeds thirty days. Metadata indicates reliable estimations.
	2	<i>Fair</i> ; insufficient meta data for a conclusion to be drawn on the appropriateness for tidal estimations.
	3	<i>Use with caution</i> ; no meta data provided and there is concern on the accuracy of the provided information.

Table 2. Overview of the flags provided in ArcTiCA and appropriate descriptions of flags.

data. As tide data points in the Arctic are scarce, these flags should serve as a guideline for the entire dataset. A map of these two flags is given in Fig. 3, and a summary of the flags and the corresponding values are given in Table 2.

Data flag. The data source for the constituents is used to create the data flag. As the sources of data and techniques vary, this flag is designed as a general guide to a user who may place higher confidence in a particular source type. Where time-series data of either bottom pressure or sea level data are available from a gauge from which we could derive the estimations using consistent approaches, the data flag is 0.

Time series analysis by us is the preferred technique for deriving constituents within this dataset, as it also allows for the removal of outliers and accurate correction of the reference time of the measurement to UTC. A data flag value of 1 is given when the tidal constituents were obtained from a source directly. However, there is no way for us to check the accuracy of these tidal estimations; i.e., the time-series data was not directly available to us. A data flag of 2 is given when tidal coefficients were provided by external sources, meaning they were either taken from websites or published datasets, but where we cannot confirm the methodology of the estimations. These data are helpful, but certain provided measurements could still be imperfect. Finally, a data flag of 3 is given when constituents were extracted from published tables or figures from scientific literature. Again, these data cannot be directly assessed by us, and errors may be present within these data. Note that the data flag is not, by itself, a guarantee that the actual data quality of a specific record is better or worse than another with a different data flag. Several factors such as instrument drift and changing amplitudes and phases as sea ice varies, can influence the quality of individual measurements, regardless of whether the analyses were carried out by us (data flag = 0) or by someone else (data flags 1–3).

Expert opinion flag. The second flag is termed an ‘expert opinion’ flag (EO flag). It provides our assessment of the suitability of the different estimations based on available data or, for data flags 1–3, the detail in the metadata

Dataset variable	Description of variable
<i>source_id</i>	the provided ID number for the respective source
<i>lon</i>	the longitudinal position (in degrees 0 to 360) of the measurement
<i>lat</i>	the latitudinal position (in degrees) of the measurement
<i>cons</i>	the respective tidal constituent
<i>amp</i>	amplitude of the tidal constituent
<i>pha</i>	phase lag of the tidal constituent (from 0 to 360 degrees)
<i>amp_uncert</i>	the standard deviation of the amplitude (cm)
<i>pha_uncert</i>	the standard deviation of the phase lag (in degrees)
<i>start</i>	start time of the measurements
<i>end</i>	end time of the measurements
<i>number_of_obs</i>	the number of observations available within the <i>in situ</i> time series
<i>missing_obs</i>	the number of missing observations within a time series, i.e. gaps in the time series
<i>source</i>	the source of the tidal constants or time series used for tidal constituents estimation
<i>instrument</i>	the type of instrument: tide gauge, OBP or GNSS-IR
<i>site</i>	the site name of the instrument
<i>rec_length</i>	the total record length in days
<i>sampling_rate</i>	the sampling rate of the measurement in minutes
<i>inference</i>	whether inference methods are used in this measurement
<i>data_flag</i>	data flag as described in section <i>Data Flag</i>
<i>expert_flag</i>	expert opinion flag as described in section <i>Expert Opinion Flag</i>
<i>site_record</i>	the observations record number for the particular site
<i>site_total</i>	the total number of sources for the same particular site
<i>amp_units</i>	the units of the amplitude estimations (cm or mbar)
<i>uncert_info</i>	the type of uncertainty information provided, either standard deviations (STD), confidence intervals (CI) or none.
<i>notes</i>	any notes about this particular station

Table 3. The variables of the ArcTiCA dataset and a brief description of the variables.

information. As with the data flag, this flag is not an indictment on the data itself; e.g., in some cases the meta-data is insufficient to judge data quality, but the estimations themselves may still be accurate. It is the dataset user's responsibility to review the original documentation and follow up with original data authors as needed. We encourage feedback to us on further insights into specific dataset elements that may change the EO flag for future ArcTiCA releases.

The EO flag is set to 0 when the data time series is analysed by us, is greater than one year long and the series has less than 10% missing data, allowing for reliable estimation and separation of major tidal constituents. An EO flag of 1 is given to measurements with less than one year's worth of data but more than thirty days. This EO flag value is also given to data containing limited metadata but provided by sources with data flag = 1. This is done to account for estimations by previous studies where those authors did not provide the appropriate meta-data required to determine this flag but are provided by sources that we trust based on their experience with tidal analyses. An EO flag of 2 is given to data where insufficient metadata is available, and, therefore, the reliability of the provided estimations cannot be determined. An EO flag of 3 is given when no metadata is provided, and there is a concern about the accuracy of the provided information, either related to the positional accuracy of the gauges or the tidal coefficients themselves. Additionally, record lengths that are less than thirty days are flagged as 3 as constituents determined from short records can contain large errors, particularly relating to K_1 and S_2 .

ArcTiCA Data Records

The ArcTiCA data are available from the Arctic Data Center at <https://doi.org/10.18739/A2VT1GR6446> in either NETCDF4 or CSV formats. These formats are chosen to allow for ease of use by users and the selection of data by specific criteria. The crucial metadata users need to determine whether the tidal data is appropriate for specific applications, defined in Table 3, are provided within the dataset. Additionally, a README file (*ArcTiCA_README.pdf*) is provided and a history file (*ArcTiCA_Revision_History.txt*) is given for users to keep track of any updates that are released. Updates to ArcTiCA will be released with a new DOI. When new versions are released, the previous DOI will direct users to the newest version of ArcTiCA.

At the time of initial publication, ArcTiCA contains 1924 individual *in situ* stations with at least an estimation for the M_2 constituent and 1020 with at least eight constituents. A total of 29 data sources were used to create the current dataset, where either the tidal constituents were directly estimated by us, or the data were provided from personal communications or literature sources. Where possible, all available metadata from the different sources is provided within the dataset, which is shown in Table 3. The current spatial distribution of 1924 *in situ* stations is presented in Fig. 4 for the amplitude and phase lag of the M_2 tidal constituent and compared to a global tide model, DTU16²⁸.

Although duplicates remain due to obtaining data for the same site from different sources, they are kept within ArcTiCA as they often contain estimations from varying sources or use different time-series lengths. It is

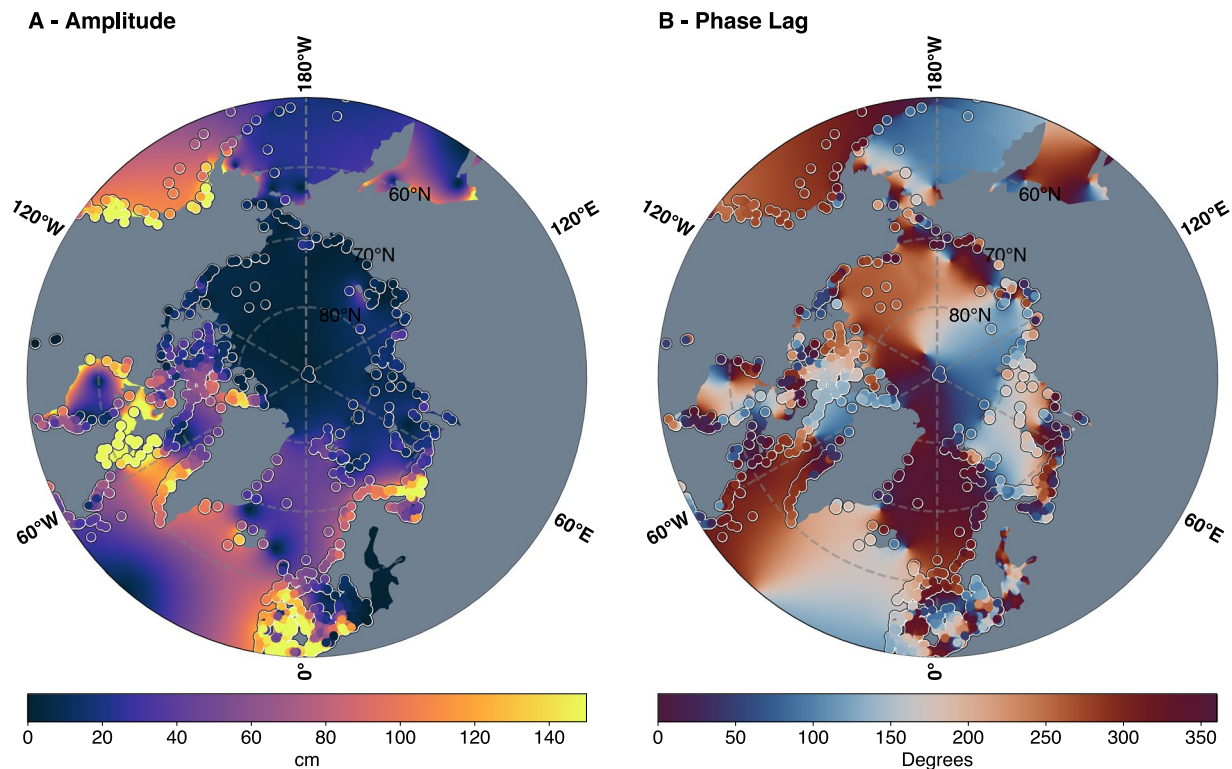


Fig. 4 The distribution of tidal constants from the ArcTiCA dataset overlaid onto an M_2 tidal amplitude and phase estimate from DTU16.

the user's responsibility to decide which source to use for individual sites using the recommendations provided by the data and EO flags. For each site, all the available constituents are given; therefore, sites have differing numbers of constituents. A total of 43 different tidal constituents are available within ArcTiCA. This decision was preferred over limiting the dataset to a certain number of constituents to provide as much data as possible for future modelling efforts, which will continue to expand the number of tidal constituents investigated.

Technical Validation

Evaluation using a model ensemble. We used the M_2 coefficients from all available *in situ* measurements for our initial evaluation of ArcTiCA. Figure 5 demonstrates the overall differences between an ensemble of models for the M_2 constituent in terms of root-mean-square error (RMS), as well as amplitude and phase differences. This comparison demonstrates the value of these flags, as regions with extreme differences between observations in the models are likely those that would be flagged within the dataset.

The ensemble of ocean tide models was estimated by taking the mean amplitude and phase calculated from four tide models; DTU16²⁸, GOT5.1 (26 updated), EOT-Polar-v0.1 (27 updated) and FES2014b¹⁷. The models differ in terms of the data sets that are assimilated; for records in ArcTiCA that have been assimilated in a particular model, we expect reduced errors compared with models with no assimilation of that site. We chose an ensemble approach to account for individual model errors that could influence statistical interpretations and to avoid an inter-model comparison that is beyond the scope of this dataset documentation. Note that, even for models in the lower latitudes with abundant observations and satellite data, the coastal region remains the most challenging region for ocean tide models. Since most observations are tide gauges at the coast, we expect that model and observations would have differences that exceed centimetres^{8,27}. This is due to factors such as limited bathymetry information, complex tidal dynamics and sea-ice interactions. Additionally, many gauges are in river or estuarine environments, including fjords, where the relatively coarse grids of the global models do not provide ocean tide estimates.

The model ensemble's median RMS for M_2 was 5.89 cm for the dataset above 60°N and 4.63 cm above 70°N, which correlates well with previous literature. Cancet *et al.*⁹ and Stammer *et al.*⁸ both independently evaluated global tide models against *in situ* stations in the Arctic, although using only 121 and 20 tide gauge measurements, respectively. Stammer *et al.*⁸ reported RMS errors for the M_2 tide between 3.91 and 5.89 cm, while Cancet *et al.*⁹ found errors between 5.8 and 8.7 cm, both using some of the major global ocean tide models. For the other constituents, these RMS differences align well with previous literature. The regions with the highest RMS were along the Russian coastline and within Hudson Bay. The latter is a well-known region of difficulty for tide models, having a relatively high spread between tidal models in this region⁸, and the RMS in this region is primarily driven by differences in amplitudes between the measurements and the models.

On the other hand, the Russian coastline is a complex tidal region and has troubled modellers due to limited bathymetry information. However, the RMS difference here may additionally be explained by the relatively poor

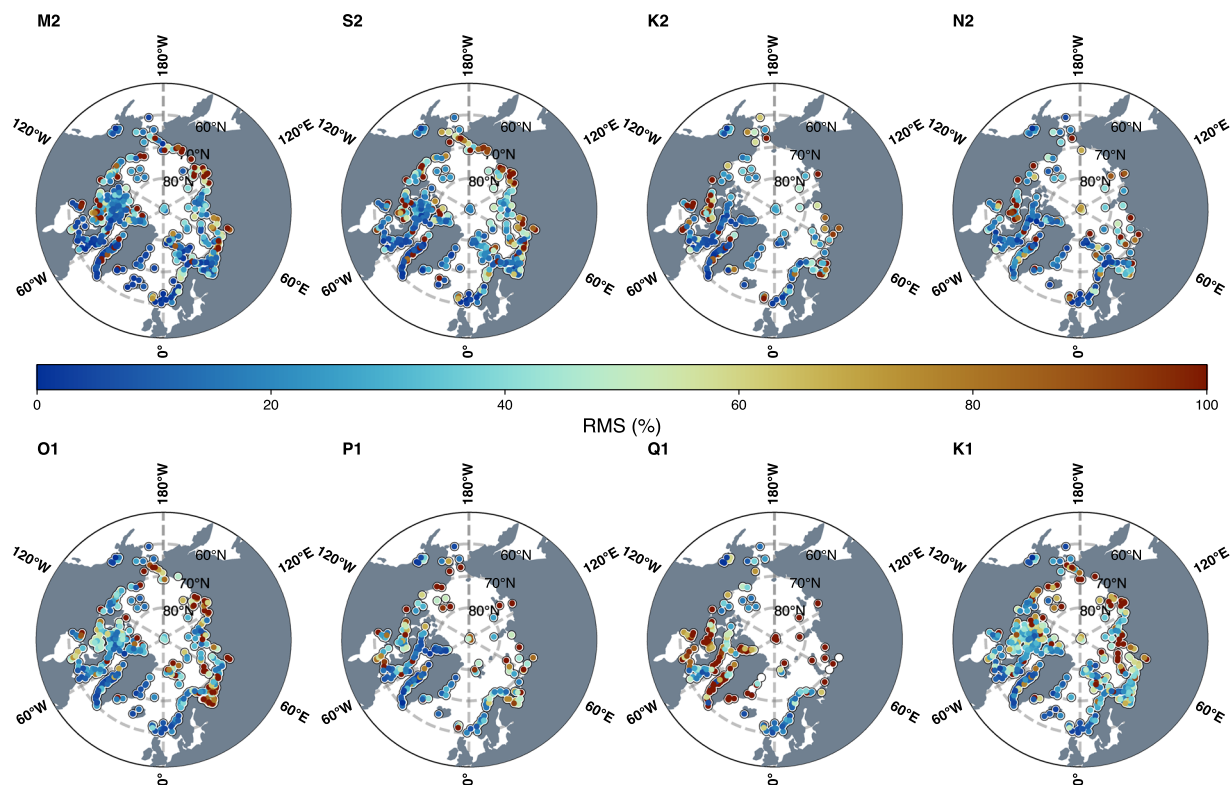


Fig. 5 RMS of an ensemble of models (DTU16, GOT5.1, EOT-Polar v0.1 and FES2014b) relative to the respective amplitudes for the eight major tidal constituents taken from the ArcTiCA dataset.

quality of the situ measurements, which are either extremely old or are provided with very little metadata to confirm their estimations. When making use of the EO flag by selecting only data flagged as 0 and 1, the RMS decreases to 5.20 and 4.31 cm above 60°N and above 70°N, respectively, which remains in line with previous literature. This is primarily due to the removal of several gauges with an EO flag of 2 or 3 (see Fig. 3), which further highlights the importance of both the data and EO flag in the application and interpretation of the dataset.

Usage Notes

ArcTiCA will serve the ocean tide modelling community by providing extensive model validation data and by offering a substantially larger database for assimilation into inverse models. Although we have attempted to make comprehensive flags and remove unreliable or erroneous estimations, users should still proceed with caution when drawing their own conclusions from using these data. In these cases, we strongly recommend that users consider the *data_flags* and *expert_flags* within ArcTiCA when interpreting their results for a specific region. In some cases, users may find that high-quality data are very sparse and that a model that assimilates tidal coefficients obtained from satellite altimetry may provide higher-accuracy coefficients than can be obtained from a poorly-located and essentially undocumented tide gauge record.

ArcTiCA is designed to make using the data as easy to do as possible by using formats that are easily ingested by any popular coding software. Examples of how to use the dataset and select data by specific variable criteria are provided in a public GitHub repository found at <https://github.com/hart-davis/ARCTICA/>.

Code availability

A public GitHub repository has been set up to help users process the data, which will be updated based on user feedback or input as well as when changes occur to the dataset. This repository can be found at <https://github.com/hart-davis/ArcTiCA>.

Received: 3 August 2023; Accepted: 25 January 2024;

Published online: 03 February 2024

References

1. Munk, W. & Wunsch, C. Abyssal recipes II: energetics of tidal and wind mixing. *Deep Sea Research Part I: Oceanographic Research Papers* **45**, 1977–2010, [https://doi.org/10.1016/S0967-0637\(98\)00070-3](https://doi.org/10.1016/S0967-0637(98)00070-3) (1998).
2. Kowalik, Z. & Proshutinsky, A. Y. The Arctic Ocean tides. In O. M. Johannessen, J. E. O., R. D. Muench (ed.) *The Polar Oceans and Their Role in Shaping the Global Environment*, 137–158, <https://doi.org/10.1029/GM085p0137> (American Geophysical Union (AGU), 1994).
3. Holloway, G. & Proshutinsky, A. Role of tides in Arctic ocean/ice climate. *Journal of Geophysical Research: Oceans* **112**, <https://doi.org/10.1029/2006JC003643> (2007).

4. Luneva, M. V., Aksenov, Y., Harle, J. D. & Holt, J. T. The effects of tides on the water mass mixing and sea ice in the Arctic Ocean. *Journal of Geophysical Research: Oceans* **120**, 6669–6699, <https://doi.org/10.1002/2014JC010310> (2015).
5. Padman, L., Siegfried, M. R. & Fricker, H. A. Ocean tide influences on the Antarctic and Greenland ice sheets. *Reviews of Geophysics* **56**, 142–184, <https://doi.org/10.1002/2016RG000546> (2018).
6. Schrama, E. J. O. & Ray, R. D. A preliminary tidal analysis of TOPEX/POSEIDON altimetry. *Journal of Geophysical Research* **99**, 24799, <https://doi.org/10.1029/94jc01432> (1994).
7. Hauk, M. & Pail, R. Treatment of ocean tide aliasing in the context of a next generation gravity field mission. *Geophysical Journal International* **214**, 345–365, <https://doi.org/10.1093/gji/ggy145> (2018).
8. Stammer, D. *et al.* Accuracy assessment of global barotropic ocean tide models. *Reviews of Geophysics* **52**, 243–282 (2014).
9. Cancet, M., Andersen, O., Lyard, F., Cotton, D. & Benveniste, J. Arctide2017, a high-resolution regional tidal model in the Arctic. *Advances in Space Research* **62**, 1324–1343, <https://doi.org/10.1016/j.asr.2018.01.007> (2018).
10. Jakobsson, M. *et al.* The International Bathymetric Chart of the Arctic Ocean Version 4.0. *Scientific Data* **7**, <https://doi.org/10.1038/s41597-020-0520-9> (2020).
11. Kowalik, Z. A study of the M_2 tide in the ice-covered Arctic Ocean. *Modeling, Identification and Control: A Norwegian Research Bulletin* **2**, 201–223, <https://doi.org/10.4173/mic.1981.4.2> (1981).
12. Kagan, B. & Sofina, E. Ice-induced seasonal variability of tidal constants in the Arctic Ocean. *Continental Shelf Research* **30**, 643–647, <https://doi.org/10.1016/j.csr.2009.05.010> (2010).
13. Bij de Vaate, I., Vasulkar, A. N., Slobbe, D. C. & Verlaan, M. The influence of Arctic landfast ice on seasonal modulation of the M_2 tide. *Journal of Geophysical Research: Oceans* **126**, <https://doi.org/10.1029/2020jc016630> (2021).
14. Fer, I., Müller, M. & Peterson, A. K. Tidal forcing, energetics, and mixing near the Yermak Plateau. *Ocean Science* **11**, 287–304, <https://doi.org/10.5194/os-11-287-2015> (2015).
15. Padman, L., Plueddemann, A. J., Muench, R. D. & Pinkel, R. Diurnal tides near the Yermak Plateau. *Journal of Geophysical Research: Oceans* **97**, 12639–12652, <https://doi.org/10.1029/92JC01097> (1992).
16. Skarðhamar, J., Skagseth, Ø. & Albretsen, J. Diurnal tides on the Barents Sea continental slope. *Deep Sea Research Part I: Oceanographic Research Papers* **97**, 40–51, <https://doi.org/10.1016/j.dsr.2014.11.008> (2015).
17. Lyard, F. H., Allain, D. J., Cancet, M., Carrère, L. & Picot, N. FES2014 global ocean tide atlas: design and performance. *Ocean Science* **17**, 615–649, <https://doi.org/10.5194/os-17-615-2021> (2021).
18. Cartwright, D. E. *Tides - A Scientific History* (Cambridge University Press, Cambridge, 1999).
19. Woodworth, P. L. Tide prediction machines at the Liverpool tidal institute. *History of Geo- and Space Sciences* **11**, 15–29, <https://doi.org/10.5194/hgss-11-15-2020> (2020).
20. Schwiderski, E. W. On charting global ocean tides. *Reviews of Geophysics* **18**, 243, <https://doi.org/10.1029/rg018i001p00243> (1980).
21. Padman, L., Erofeeva, S. & Howard, S. AOTIM5: Arctic Ocean Inverse Tide Model, on 5 kilometer grid, developed in 2004, <https://doi.org/10.18739/A2S17SS80> (2020).
22. Padman, L., Fricker, H. A., Coleman, R., Howard, S. & Erofeeva, L. A new tide model for the Antarctic ice shelves and seas. *Annals of Glaciology* **34**, 247–254, <https://doi.org/10.3189/172756402781817752> (2002).
23. Egbert, G. D. & Erofeeva, S. Y. Efficient inverse modeling of barotropic ocean tides. *Journal of Atmospheric and Oceanic Technology* **19**, 183–204 (2002). [10.1175/1520-0426\(2002\)019<0183:eimobo>2.0.co;2](https://doi.org/10.1175/1520-0426(2002)019<0183:eimobo>2.0.co;2)
24. Hart-Davis, M. G. *et al.* Regional evaluation of minor tidal constituents for improved estimation of ocean tides. *Remote Sensing* **13**, <https://doi.org/10.3390/rs13163310> (2021).
25. Talke, S. A. & Jay, D. A. Changing tides: The role of natural and anthropogenic factors. *Annual Review of Marine Science* **12**, 121–151, <https://doi.org/10.1146/annurev-marine-010419-010727> (2020).
26. Ray, R. D. Precise comparisons of bottom-pressure and altimetric ocean tides. *Journal of Geophysical Research: Oceans* **118**, 4570–4584, <https://doi.org/10.1002/jgrc.20336> (2013).
27. Hart-Davis, M. G. *et al.* EOT20: a global ocean tide model from multi-mission satellite altimetry. *Earth System Science Data* **13**, 3869–3884, <https://doi.org/10.5194/essd-13-3869-2021> (2021).
28. Cheng, Y. & Andersen, O. B. Towards further improving DTU global ocean tide model in shallow waters and polar seas. Proceedings of the Ocean Surface Topography Science Team (OSTST) Meeting, Miami, FL, USA, ftp://ftp.space.dtu.dk/pub/DTU16/OCEAN_TIDE/OSTST2017-tide.pdf (2017).
29. Haigh, I. D. *et al.* GESLA version 3: A major update to the global higher-frequency sea-level dataset. *Geoscience Data Journal* **n/a** (2022).
30. Hart-Davis, M. G., Dettmering, D. & Seitz, F. TICON-3: Tidal Constants based on GESLA-3 sea-level records from globally distributed tide gauges including gauge type information (data) PANGAEA <https://doi.org/10.1594/PANGAEA.951610> (2022).
31. Baumann, T. M. *et al.* Arctic tidal current atlas. *Scientific Data* **7**, 275, <https://doi.org/10.1038/s41597-020-00578-z> (2020).
32. Howard, S. L. & Padman, L. Arc2kmTM: Arctic 2 kilometer tide model, 2021, <https://arcticdata.io/catalog/view/doi:10.18739/A2PV6B79W> (2021).
33. European Union-Copernicus Marine Service. Arctic ocean tidal analysis and forecast. *Mercator Ocean International* <https://doi.org/10.48670/MOI-00005> (2020).
34. Holgate, S. J. *et al.* New Data Systems and Products at the Permanent Service for Mean Sea Level. *Journal of Coastal Research* **29**, 493–504, <https://doi.org/10.2112/JCOASTRES-D-12-00175.1> (2013).
35. Cartwright, D. E., Zetler, B. D. & Hamon, B. V. Pelagic tidal constants. Publ. Sci. 30, International Association for the Physical Sciences of the Ocean, Paris (1979).
36. Cartwright, D. E. & Zetler, B. D. Pelagic tidal constants 2. Publ. Sci. 33, International Association for the Physical Sciences of the Ocean, Paris (1985).
37. Smithson, M. J. Pelagic tidal constants 3. Publ. Sci. 35, International Association for the Physical Sciences of the Ocean, Paris (1992).
38. Larson, K. M., Ray, R. D., Nievinski, F. G. & Freymueller, J. T. The accidental tide gauge: A GPS reflection case study from Kachemak Bay, Alaska. *IEEE Geoscience and Remote Sensing Letters* **10**, 1200–1204, <https://doi.org/10.1109/lgrs.2012.2236075> (2013).
39. Tabibi, S., Geremia-Nievinski, F., Francis, O. & van Dam, T. Tidal analysis of GNSS reflectometry applied for coastal sea level sensing in Antarctica and Greenland. *Remote Sensing of Environment* **248**, 111959, <https://doi.org/10.1016/j.rse.2020.111959> (2020).
40. Codiga, D. L. *Unified tidal analysis and prediction using the UTide Matlab functions* <https://www.po.gso.uri.edu/codiga/utide/2011Codiga-UTide-Report.pdf> (2011).
41. Ribergaard, M. Tide tables for Danish waters 2023. *DMI Report* <https://www.dmi.dk/fileadmin/Rapporter/2022/DMIRep22-32.pdf> (2023).
42. Ribergaard, M. Tide tables for Faroe Island waters 2023. *DMI Report* <https://www.dmi.dk/fileadmin/Rapporter/2022/DMIRep22-33.pdf> (2023).
43. Ribergaard, M. Tide tables for Greenlandic waters 2023. *DMI Report* <https://www.dmi.dk/fileadmin/Rapporter/2022/DMIRep22-34.pdf> (2023).
44. Krishfield, R., Toole, J. & Proshutinsky, A. BGFE 2003-2004 MMP EMCTD and ACM data processing procedures. Tech. Rep., Woods Hole Oceanographic Institution. https://www2.whoi.edu/site/beaufortgyre/wp-content/uploads/sites/108/2020/04/BGFE_MMP_Processing_2004-final_5-31-05_85606.pdf (2004).
45. Schureman, P. Manual of harmonic analysis and prediction of tides. Spec. Pub. 98, U.S. Coast & Geodetic Survey, Washington (1940).

46. Hart-Davis, M. *et al.* Arctic Tidal Constituent Atlas (ArcTiCA): A database of tide elevation constituents for the Arctic region from 1800 through present day. *Arctic Data Center* <https://doi.org/10.18739/A2VT1GR64> (2023).
47. Kulikov, M. E., Medvedev, I. P. & Kondrin, A. T. Seasonal variability of tides in the Arctic Seas. *Russian Journal of Earth Sciences* **18**, 1–14, <https://doi.org/10.2205/2018es000633> (2018).
48. Davis, P. E. D., Lique, C. & Johnson, H. L. On the link between Arctic sea ice decline and the freshwater content of the Beaufort Gyre: Insights from a simple process model. *Journal of Climate* **27**, 8170–8184, <https://doi.org/10.1175/JCLI-D-14-00090.1> (2014).
49. Dietrich, R. *et al.* Jakobshavn Isbræ, West Greenland: Flow velocities and tidal interaction of the front area from 2004 field observations. *Journal of Geophysical Research: Earth Surface* **112**, <https://doi.org/10.1029/2006JF000601> (2007).
50. Gjevik, B. & Straume, T. Model simulations of the M2 and the K1 tide in the Nordic Seas and the Arctic Ocean. *Tellus A: Dynamic Meteorology and Oceanography* **41**, 73, <https://doi.org/10.3402/tellusa.v41i1.11822> (1989).
51. Gjevik, B., Nøst, E. & Straume, T. Model simulations of the tides in the Barents Sea. Tech. Rep., Research Report, Matematisk institutt, Department of Mathematics, University of Oslo, Oslo. <https://www.duo.uio.no/handle/10852/49003> (1992).
52. Greisman, P., Grant, S., Blaskovich, A. & Van Hardenburg, B. Tidal propagation measurements in Baffin Bay. *Canadian Contractor Report of Hydrography and Ocean Sciences* (1986).
53. Peralta-Ferriz, C., Morison, J. H., Stalin, S. E. & Meinig, C. Measuring Ocean Bottom Pressure at the North Pole. *Marine Technology Society Journal* **48**, 52–68, <https://doi.org/10.4031/mts.j.48.5.11> (2014).
54. Peralta Ferriz, A. C. *Arctic Ocean circulation patterns revealed by ocean bottom pressure anomalies*. Ph.D. thesis. <https://digital.lib.washington.edu/researchworks/handle/1773/22539?show=full> (2012).
55. Richter, A., Rysgaard, S., Dietrich, R., Mortensen, J. & Petersen, D. Coastal tides in West Greenland derived from tide gauge records. *Ocean Dynamics* **61**, 39–49, <https://doi.org/10.1007/s10236-010-0341-z> (2011).
56. Voinov, G. Harmonic tidal analysis for the nodal 18.61-year period with Amderma settlement taken as an example. *Russian Meteorology and Hydrology* 44–57 (2006).
57. Janout, M. A., Kanzow, T., Hölemann, J. A. & Ivanov, V. Raw data of physical oceanography and current velocity from mooring AK6-1 in the Arctic Ocean in 2015–2018. *PANGAEA* <https://doi.org/10.1594/PANGAEA.949152> (2023).
58. McRaven, L. & Pickart, R. Arctic Observing Network (AON) observations from the 2018–2020 Beaufort shelf-edge mooring array. *NSF Arctic Data Center* <https://doi.org/10.18739/A2HX15S0W> (2022).
59. Morison, J., Wahr, J., Kwok, R. & Peralta-Ferriz, C. Recent trends in Arctic Ocean mass distribution revealed by GRACE. *Geophysical Research Letters* **34**, <https://doi.org/10.1029/2006GL029016> (2007).
60. Nilsen, F., Ersdal, E. A. & Skogseth, R. Wind-Driven Variability in the Spitsbergen Polar Current and the Svalbard Branch across the Yermak Plateau. *Journal of Geophysical Research: Oceans* **126**, e2020JC016734, <https://doi.org/10.1029/2020JC016734> (2021). E2020JC016734 2020JC016734.
61. Polyakov, I. NABOS II - Mooring M1_4a SBE 53 Bottom Pressure Data 2013–2015. *NSF Arctic Data Center* <https://doi.org/10.18739/A2VD6P52T> (2016).

Acknowledgements

This work was funded in part by the German Research Foundation (Deutsche Forschungsgemeinschaft DFG) project TIDUS-2 (DE2174/12-2, Project Number 388296632) to MHD, by the National Science Foundation grant 1708424 to LP and SLH, and FN received support from the Research Council of Norway project 222696 (REOCIRC) and 276730 (Nansen Legacy). We would also like to thank the numerous sources for providing the data used in this dataset. We also would like to thank community members who have provided extensive feedback, particularly Svetlana Erofeeva and the IAPSO Tidal prediction group.

Author contributions

M.G.H.-D., S.L.H., R.D.R., O.B.A., L.P. and D.D. contributed to the project conceptualisation. M.G.H.-D., S.L.H., R.D.R., O.B.A. and L.P. collected and analysed the different sources of data analysed in this manuscript. F.N. provided previously unpublished mooring data used in the dataset. All authors contributed to the manuscript writing as well as reviewed and approved the final version of the manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to M.G.H.-D.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024