

TICON: Tidal CONstants based on GESLA sea-level records from globally located tide gauges

Gaia Piccioni | Denise Dettmering | Wolfgang Bosch | Florian Seitz

Deutsches Geodätisches Forschungsinstitut
der Technische Universität München
(DGFI-TUM), München, Germany

Correspondence

Gaia Piccioni, Deutsches Geodätisches
Forschungsinstitut der Technische
Universität München (DGFI-TUM),
München, Germany.
Email: gaia.piccioni@tum.de

Abstract

The TICON (Tidal CONstants) dataset contains harmonic constants of 40 tidal constituents computed for 1,145 tide gauges distributed globally. The tidal estimations are based on publicly available sea level records of the second version of the Global Extreme Sea Level Analysis (GESLA) project and were derived through a least squares-based harmonic analysis on the single time series. A preliminary screening was performed on all records to exclude doubtful observations. Only the records containing more than 70% of valid measurements were processed, that correspond to 89.7% of the total 1,276 original public GESLA records. The results are stored in a text file, and include additional information on the position of the stations, the starting and ending years of the analysed record, the estimated error of the fit, a code that corresponds to the source of the record and additional information on the single time series. In ocean tide models, data from in situ stations are used for validation purposes, and TICON is a useful and easy-to-handle data set that allows the users to select the records according to different criteria most suitable for their purposes. The data are provided with DOI identification in the PANGAEA repository.

KEYWORDS

GESLA, tide gauge, tides

Data set

The Tidal CONstants (TICON) data set contains information on harmonic constants for 40 tidal constituents, derived from the GESLA-2 tide gauge records (<http://www.gesla.org>) distributed on a quasi-global scale. The data set is freely available to the public from the PANGAEA repository.

Identifier: <https://doi.pangaea.de/10.1594/PANGAEA.896587>

Creators: Piccioni, G., D. Dettmering, W. Bosch, and F. Seitz

Title: TICON: Tidal Constants based on GESLA sea-level records from globally distributed tide gauges (data)

Publisher: PANGAEA

Publication year: 2018

Resource type: dataset

Version: 1

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2019 The Authors. *Geoscience Data Journal* published by Royal Meteorological Society and John Wiley & Sons Ltd.

1 | INTRODUCTION

Tide gauges have been a fundamental element in sea-level analysis since the nineteenth century, and continue to represent a complementary, yet necessary tool to evaluate oceanographic models (e.g. Higginson *et al.*, 2015, satellite data (Passaro *et al.*, 2018; Volkov and Pujol, 2012), and to perform instrumental calibration of satellite missions (Chambers *et al.*, 1998; Wunsch, 1986). Tide gauges are regularly exploited in ocean tide models for data assimilation and performance assessment. In Stammer *et al.* (2014), the tidal harmonic constants were derived from tide gauge time series and were used to estimate the accuracy of the so-called modern tide models in terms of statistical differences. Further examples of such practice can be easily found in the literature, for example, Ray (2013), Lago *et al.* (2017), Cheng and Andersen (2011), Carrère *et al.* (2012). Most of the in situ data used in the aforementioned papers are located in both open oceans and shallow waters. However, the growing concern for our coasts encouraged tide modelling research towards improved solutions in the latter areas Ray *et al.* (2011), and as a consequence, the need for larger data sets of coastal in situ information Cazenave and Nerem (2004) and Restano *et al.* (2018). The Tidal CONstants (TICON) data set was created with this purpose, and specifically to provide the users with a simple tool that helps in tide model validation at the coast. TICON contains information on the harmonic constants of 40 tidal constituents, computed for 1,445 tide gauge stations. The constants are the result of a least squares harmonic analysis performed on time series belonging to the Global Extreme Sea Level Analysis (GESLA) data bank (Woodworth *et al.*, 2017). GESLA was chosen as a basis for TICON not only because of its higher frequency (1 hr or faster) sampling which is suitable for tidal analysis but also because it provides a comprehensive set of sea level records located on a global scale. Also, its public distribution together with the release of harmonized and user-friendly record files, both facilitate the usage of the data. Indeed, even though tide gauge data are often available via a direct download, their temporal resolution can be lower than hours or have a monthly frequency, like in the case of the Permanent Service For Mean Sea Level (PSMSL, Holgate *et al.*, 2013). Higher frequency data are also available to the public, but they generally require a formal request which entails a certain waiting time. TICON is characterized by a simple file format that helps the users to select the records according to their needs. For each tide gauge station, estimates of the tidal constants are provided together with additional information such as the location of the station or the time period of the tidal estimation. The dataset is freely available to public, and it is registered with a digital object identifier on the PANGAEA platform. In the following

pages, the TICON data set will be introduced. In Section 3 the source data set GESLA is illustrated. Section 4 explains the method used to build TICON. The file format is described in Section 5. The work is summarized in Section 6.

2 | INPUT DATA

The latest version of GESLA contains 1,355 harmonized records, collected among 30 different sources such as national authorities, research institutions and international projects. An exhaustive description of the different datasets involved, together with the corresponding source reference, can be found in (Woodworth *et al.*, 2017). A total of 1,276 of these records are publicly available and were used to build the TICON data set. The remaining 77 ‘private’ records were not used, as the intention of the authors was to guarantee a free and direct access to the data. GESLA public stations have a quasi-global extent, with a higher data coverage in the Northern Hemisphere, in particular:

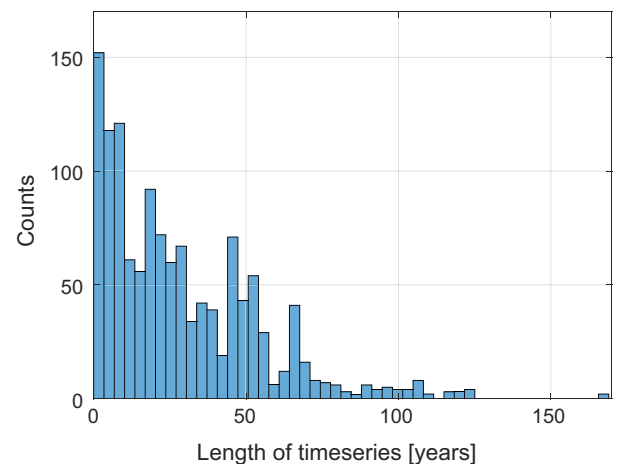


FIGURE 1 Overview of the length (in years) of the GESLA time series

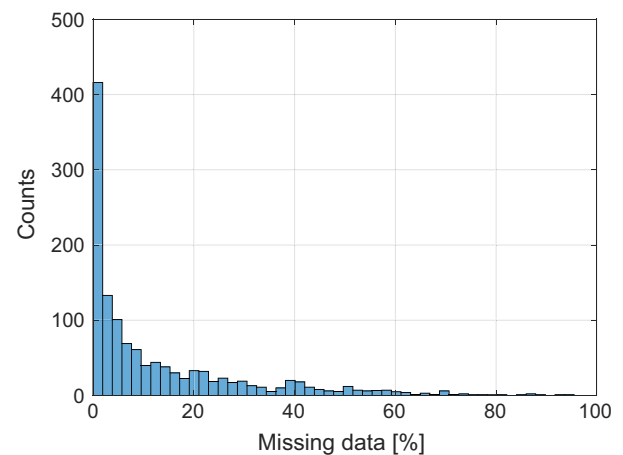


FIGURE 2 Overview of missing data in GESLA records

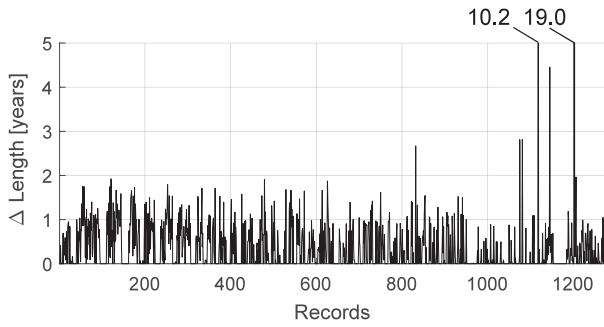


FIGURE 3 Difference in time series length before and after flagging rejection. The record numbers on the x-axis are sorted according to the length of the original timeseries, from the shortest to the longest

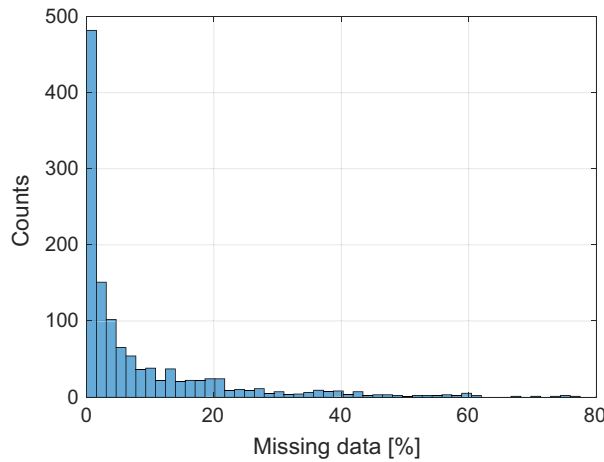


FIGURE 4 Overview of missing data in GESLA records after removing the flagged data at the extremes of the single time series

North America, Europe and Japan. Apart from the sparse number in the open ocean – which is related to the presence of islands – unfortunately almost no station can be found in

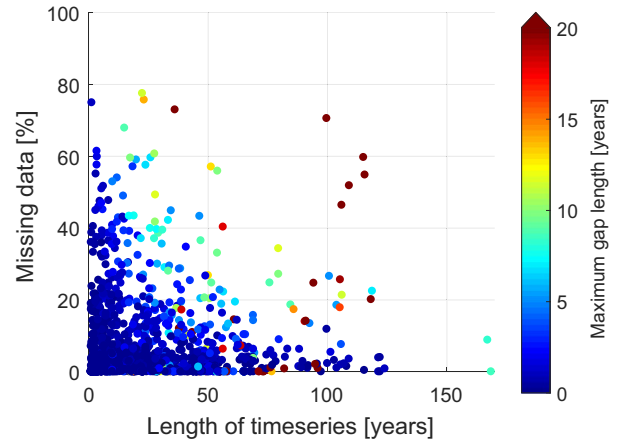


FIGURE 5 Percentage of missing data plotted against the time series length. The maximum gap length is colour coded. The plot shows the records after the removal of flagged data at each end of their time series

regions such as at high latitudes or at the coast of the Indian Subcontinent.

A preliminary screening was performed on all records to analyse the average duration of the measurements and the distribution of the temporal gaps. We have observed that the records span from a minimum of 150 days, to a maximum of 168.6 years, having a median length of 22.2 years and a general distribution shown in Figure 1.

GESLA contains quality information in terms of flagged observations. The flags characterize the quality and the possible usage of the individual measurements. Only measurements assigned as "correct" or "interpolated value" were selected as valid. In addition, data gaps due to missing physical observations can occur. After flagging, 417 records have missing data less than 2% of the total number

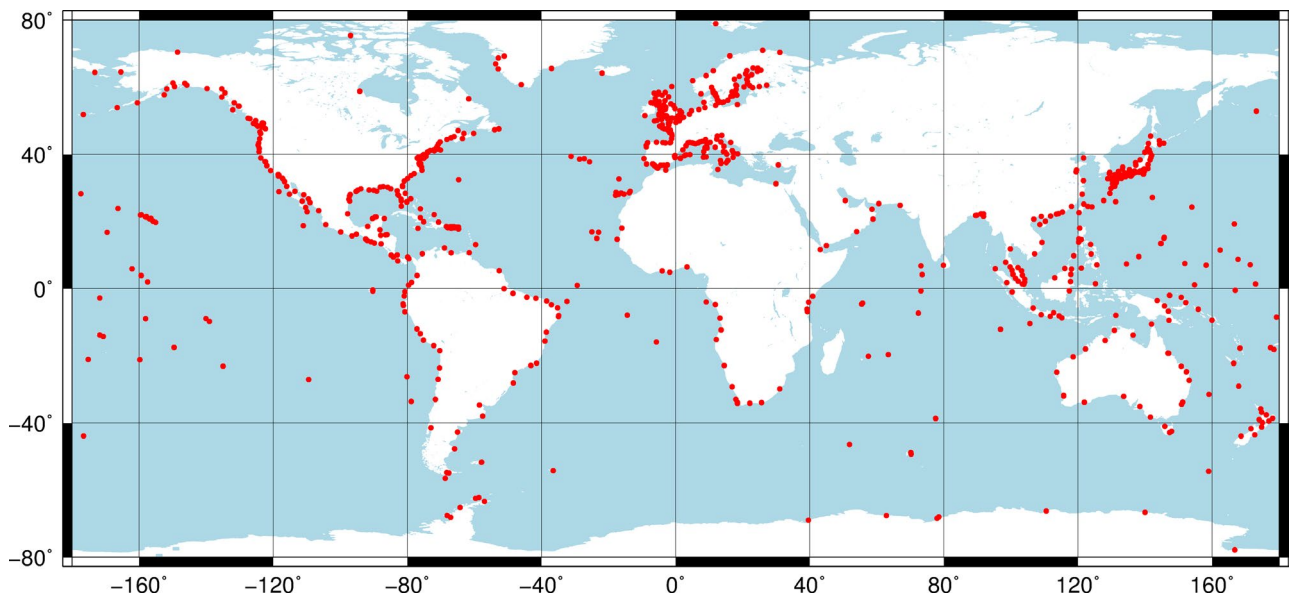


FIGURE 6 Geographical distribution of TICON data

TABLE 1 Part of the TICON dataset for the station of Port Angeles, Washington, USA. EP2, MI2, NI2 and LM2 correspond to constituents ε_2 , μ_2 , ν_2 and λ_2 respectively

Lat [°]	Lon [°]	Constituent symbol	Amplitude [cm]	Phase lag [°]	σ_{am} [cm]	σ_{ph} [°]
48.1250	236.5600	SA	10.612	3.834	0.003	0.001
48.1250	236.5600	SSA	1.896	220.888	0.001	0.004
48.1250	236.5600	MM	2.024	197.866	0.001	0.004
48.1250	236.5600	MSF	0.596	216.913	0.000	0.013
48.1250	236.5600	MF	1.298	164.147	0.000	0.005
48.1250	236.5600	MTM	0.460	183.835	0.000	0.017
48.1250	236.5600	MSQ	0.128	180.800	0.000	0.062
48.1250	236.5600	2Q1	0.804	231.874	0.000	0.009
48.1250	236.5600	Q1	6.649	232.570	0.002	0.001
48.1250	236.5600	O1	38.670	241.412	0.010	0.000
48.1250	236.5600	M1	2.145	347.220	0.000	0.001
48.1250	236.5600	P1	20.848	259.529	0.006	0.000
48.1250	236.5600	S1	2.062	33.524	0.001	0.004
48.1250	236.5600	K1	66.796	261.405	0.018	0.000
48.1250	236.5600	J1	3.402	284.364	0.001	0.002
48.1250	236.5600	OO1	2.499	304.029	0.001	0.002
48.1250	236.5600	EP2	0.665	200.423	0.000	0.012
48.1250	236.5600	2N2	1.462	248.399	0.000	0.005
48.1250	236.5600	MI2	2.754	233.646	0.001	0.003
48.1250	236.5600	N2	11.756	280.099	0.003	0.001
48.1250	236.5600	NI2	2.100	287.739	0.001	0.004
48.1250	236.5600	MA2	1.101	145.594	0.000	0.007
48.1250	236.5600	M2	51.586	307.293	0.014	0.000
48.1250	236.5600	MB2	0.823	57.938	0.000	0.010
48.1250	236.5600	MKS	0.324	169.525	0.000	0.022
48.1250	236.5600	LM2	0.616	54.276	0.000	0.013
48.1250	236.5600	L2	1.124	29.774	0.000	0.007
48.1250	236.5600	T2	0.814	335.278	0.000	0.010
48.1250	236.5600	S2	14.611	326.503	0.004	0.001
48.1250	236.5600	R2	0.275	327.092	0.000	0.029
48.1250	236.5600	K2	2.843	333.255	0.001	0.003
48.1250	236.5600	M3	0.138	341.120	0.000	0.058
48.1250	236.5600	S3	0.043	108.247	0.000	0.185
48.1250	236.5600	N4	0.156	28.526	0.000	0.051
48.1250	236.5600	MN4	0.711	63.974	0.000	0.011
48.1250	236.5600	M4	1.463	96.914	0.000	0.005
48.1250	236.5600	MS4	0.858	112.900	0.000	0.009
48.1250	236.5600	S4	0.197	114.431	0.000	0.040
48.1250	236.5600	M6	1.510	198.201	0.000	0.005
48.1250	236.5600	M8	0.026	153.469	0.000	0.300

of observations, 624 records contain missing data between 2% and 25%, 170 records between 25% and 50%, and only 65 records have gaps above 50% (Figure 2). Because of the different nature of these temporal gaps, their lengths may also vary, ranging between a few samples (in general some hours) and years.

3 | DATA PRODUCTION METHOD

3.1 | Pre-processing

An automatic programme was set up to compute tidal constants for all the records. The script performs a first selection of records suitable for least squares harmonic analysis, that is: the period of the observations of each record must be larger than 1 year. The choice of this minimum duration is based on the Rayleigh criterion for tidal constituent separation (Pugh and Woodworth (2014)). The time series length is measured after rejecting individual observations for unsuitable flags. This is done because temporal gaps due to flagging may occur at time series extremes, shortening their extent. In Figure 3 a comparison of the length of the time series before and after flagging rejection is shown. Five hundred and thirteen files show no difference after flagging, 625 are reduced up to 1 year and 138 records are shorter by more than 1 year. Two extreme cases occur for the Canadian stations of Port Hardy and New Westminster, whose time series, despite a reduction of 10 and 19 years after flagging, are still, respectively, 50.7 and 45.7 years long. Forty-four records – which correspond to 3.45% of the public GESLA dataset – do not reach the minimum time span required and are discarded from the tidal analysis and the final data set. A second selection is made in relation to the amount of missing data within each record. The distribution of missing data for a time series longer than 1 year is shown in Figure 4. After removing flagged values at record extremes, the amount of missing data for more than 500 records is below 3%. In the scatter plot of Figure 5 the percentage of missing data is plotted against the time series length. The maximum length of temporal gaps is represented by the marker colour. In general, records below 50 years duration do not show large gaps, and in the majority of cases missing data are below 30%. Few long records are characterized by extensive temporal gaps that exceed 20 years, which may cause a loss of data larger than 40%. The authors attempt to perform least squares for the longest time series possible, in order to use the full original records. For this reason, a threshold of 70% of valid observations was set, above which the records are processed for their full length. This criterion is used to compute tidal harmonics for the full time series, reducing the risk of processing records with highly scattered observations. A similar method was used by Ruiz-Etcheverry *et*

al. (2015) to sort and compare the annual and semi-annual signal of tide gauge observations against satellite data). In total, 1,145 records were processed with this condition, while 87 were excluded from the data set. The overall number of discarded records (due to short time series or missing observations) is 131, that corresponds to 10.3% of the full GESLA data set. In Figure 6 the locations of the final TICON data are shown.

3.2 | Least squares and error estimation

The 40 tidal constituents are derived using the least squares method. The choice of the number of constituents was mainly based on the purpose of this data set, which is to evaluate ocean tide models. Generally, the modern tide models include no more than 15 tidal constituents; however, there are exceptions such as FES2014 model, that provides 34 constituents Carrère *et al.* (2015). Thus, providing the 40 most important constituents should be sufficient for an adequate model evaluation.

The least squares approach is often preferred over the spectral analysis because it allows to perform the tidal estimation on incomplete time series (Ponchaut *et al.*, 2001). A matrix system was set up following equation 4 in Piccioni *et al.* (2018), with which the in-phase and quadrature coefficients (here called $H = A_k \cos P_k$ and $G = A_k \sin P_k$) – and consequently the amplitude and phase lag – are computed:

$$SLA_i = Z_0 + a \cdot t_i + \sum_{k=1}^n (A_k \cos P_k \cdot f_k \cos(\theta_k + u_k) + A_k \sin P_k \cdot f_k \sin(\theta_k + u_k)) \quad (1)$$

with Z_0 the mean sea level, a the trend of the time series, t_i the time at observation i , n the number of tidal constituents which are defined by the amplitude A_k , the phase lag P_k , θ_k the astronomical arguments, and f_k and u_k the nodal corrections for the amplitude and phase lag respectively. The values of the nodal corrections and the Doodson extended numbers needed to compute the astronomical arguments are taken from International Hydrographic Organization (2006), while the expressions for the astronomical arguments are from Tamura (1987). The amplitude and phase lag are assigned a statistical error based on the standard error of the regression (Heij *et al.*, 2004) and the error propagation principles (Taylor, 1997). In detail, the errors for the amplitude and phase lag are described by the formulae:

$$\sigma_{am} = \frac{\sqrt{H^2 \sigma_H^2 + G^2 \sigma_G^2}}{\sqrt{H^2 + G^2}} \quad (2)$$

$$\sigma_{\text{ph}} = \frac{\sqrt{G^2\sigma_H^2 + H^2\sigma_G^2}}{H^2 + G^2} \quad (3)$$

where σ_H and σ_G are the standard errors of H and G , and are related to the number of observations analysed (because of the df in the standard error of the regression), and therefore for longer time series smaller errors may be computed.

Finally, the results are merged and saved in a user-friendly text file, together with supplementary information relevant to the tide gauge station and the time series (Data S1).

4 | DATA SET LOCATION AND FORMAT

TICON is stored in the PANGAEA public repository as a text/ASCII format, and it is freely available for any research purpose. The data set is a single file that contains the harmonic constants of 40 tidal constituents and their related errors. An example of part of the TICON file is shown in Table 1 for the station of Port Angeles, Washington, USA. The file has a tab-separated column structure and the columns display information on: left to right you have latitude and longitude (with domain 0–360) of the station's location, constituent's name, amplitude of the tidal constituent in cm, phase lag (Greenwich lag) of the tidal constituent in degrees, percentage of missing data within the time series analysed, number of observations used for the least squares estimation, length in days of the largest gap found in the record, date of the first and the last observation, and a code that corresponds to the source of the record. The constituents are sorted in ascending order, according to their frequency. The user manual contains also a validation session, in which TICON constants are compared to the Finite Element Solution 2014 (FES2014) global tide model (Carrère *et al.*, 2015). Additional comparisons are shown between two or more 'duplicate' records, that is, records at the same location coming from different GESLA-2 sources. For major details about the file structure and usage, the authors suggest to read the user manual provided in PANGAEA.

5 | SUMMARY

This article describes the TICON data set, which contains information on tidal harmonic constants of 40 constituents for 1,145 tide gauge stations located worldwide. The constants were computed from the time series provided in the GESLA project, and were selected according to the lengths of the individual records and the percentage of missing data. The final results are stored in a single text file enabling a simple record selection according to the user's needs. TICON will find applications in the sea level and ocean tide community,

as it represents a directly accessible validation dataset of easy to handle. With this paper the authors want to highlight the importance of a freely available, harmonized data set and express the wish that more and more data centres will make records available for a unified dataset in a near future.

ACKNOWLEDGEMENTS

We acknowledge the providers of the GESLA data set, which is freely available on the web page: <https://www.gesla.org>. A special thank goes to Philip Woodworth, who kindly supported this idea and clarified details of the GESLA records. The authors are also grateful to Martin Saraceno and Richard Ray for sharing insights on tide gauge time series analysis and related issues. Finally, a big thank you goes to Nitin Ravinder who contributed to refine the TICON final product.

OPEN PRACTICES

This article has earned an Open Data badge for making publicly available the digitally-shareable data necessary to reproduce the reported results. The data is available at https://doi.panga.ea.de/10.1594/PANGA_EA.896587 Learn more about the Open Practices badges from the Center for Open Science: <https://osf.io/tvyxz/wiki>.

REFERENCES

- Carrère, L., Lyard, F., Cancet, M., Guillot, A. and Roblou, L. (2012) .FES 2012: a new global tidal model taking advantage of nearly 20 years of altimetry. In: Conference proceeding for the European Space Agency meeting, 20 Years of Progress in Radar Altimetry, Venice, p. 6.
- Carrère, L., Lyard, F., Cancet, M. and Guillot, A. (2015) FES 2014, a new tidal model on the global ocean with enhanced accuracy in shallow seas and in the Arctic region. In: Conference contribution for the European General Assembly meeting, EGU General Assembly Conference Abstracts, Vienna, Austria, Vol. 17, p. 5481.
- Cazenave, A. and Nerem, R.S. (2004) Present-day sea level change: observations and causes. *Reviews of Geophysics*, 42(3), RG3001. ISSN 8755–1209. Available at: <https://doi.org/10.1029/2003RG000139>.
- Chambers, D.P., Ries, J.C., Shum, C.K. and Tapley, B.D. (1998) On the use of tide gauges to determine altimeter drift. *Journal of Geophysical Research: Oceans*, 103(C6), 12885–12890. ISSN 01480227. Available at: <https://doi.org/10.1029/98JC01197>.
- Cheng, Y. and Andersen, O.B. (2011) Multimission empirical ocean tide modeling for shallow waters and polar seas. *Journal of Geophysical Research*, 116(C11), C11001. ISSN 0148–0227. Available at: <https://doi.org/10.1029/2011JC007172>.
- International Hydrographic Organization. (2006). Harmonic constants product specification. Version 1.0. Tides Committee of the International Hydrographic Organization. Dated November 2006. 29 pp.
- Heij, C., De Boer, P., Franses, P.H., Kloek, T. and van Dijk, H.K. (2004) Least squares in matrix form. In: *Econometric Methods with Applications in Business and Economics*, Chapter 3. Oxford: Oxford University Press, pp. 118–133. ISBN 9780199268016.

- Higginson, S., Thompson, K.R., Woodworth, P.L. and Hughes, C. (2015) The tilt of mean sea level along the east coast of North America. *Geophysical Research Letters*, 42. Available at: <https://doi.org/10.1002/2015GL063186>.
- Holgate, S.J., Matthews, A., Woodworth, P.L., Rickards, L.J., Tamisiea, M.E., Bradshaw, E *et al.* (2013) New data systems and products at the permanent service for mean sea level. *Journal of Coastal Research*, 288, 493–504. ISSN 0749–0208. Available at: <https://doi.org/10.2112/JCOASTRES-D-12-00175.1>.
- Lago, L.S., Saraceno, M., Ruiz-Etcheverry, L.A., Passaro, M., Oreiro, F.A., D'Onofrio, E.E *et al.* (2017) Improved sea surface height from satellite altimetry in coastal zones: a case study in Southern Patagonia. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(8), 3493–3503. ISSN 1939–1404. Available at: <https://doi.org/10.1109/JSTARS.2017.2694325>.
- Passaro, M., Kildegaard Rose, S., Andersen, O.B., Boergens, E., Calafat, F.M., Dettmering, D *et al.* (2018) ALES+: adapting a homogenous ocean retracker for satellite altimetry to sea ice leads, coastal and inland waters. *Remote Sensing of Environment*, 211, 456–471. ISSN 00344257. Available at: <https://doi.org/10.1016/j.rse.2018.02.074>.
- Piccioni, G., Dettmering, D., Passaro, M., Schwatke, C., Bosch, W. and Seitz, F. (2018) Coastal improvements for tide models: the impact of ALES retracker. *Remote Sensing*, 10(5). ISSN 20724292. Available at: <https://doi.org/10.3390/rs10050700>.
- Ponchaut, F., Lyard, F. and Le Provost, C. (2001) An analysis of the tidal signal in the WOCE sea level dataset. *Journal of Atmospheric and Oceanic Technology*, 18(1), 77–91. ISSN 07390572. Available at: [https://doi.org/10.1175/1520-0426\(2001\)018<0077:AAOTT S>2.0.CO;2](https://doi.org/10.1175/1520-0426(2001)018<0077:AAOTT S>2.0.CO;2).
- Pugh, D. and Woodworth, P.L. 2014. *Sea-Level Science*. Cambridge: Cambridge University Press, pp. 395. ISBN 9781139235778. Available at: <https://doi.org/10.1017/CBO9781139235778.007>.
- Ray, R.D. (2013) Precise comparisons of bottom-pressure and altimetric ocean tides. *Journal of Geophysical Research: Oceans*, 118(9), 4570–4584. Available at: [10.1002/jgrc.20336](https://doi.org/10.1002/jgrc.20336).
- Ray, R.D., Egbert, G.D. and Erofeeva, S.Y. (2011) Tide predictions in shelf and coastal waters: status and prospects, In: Vignudelli, S., Kostianoy, A.G., Cipollini, P. and Vignudelli, J.B. (Eds.) *Coastal Altimetry*. Ch. 8. Berlin: Springer-Verlag, pp. 191–216. Available at: https://doi.org/10.1007/978-3-642-12796-0_8.
- Restano, M., Cipollini, P., Passaro, M., Vignudelli, S., Vandemark, D. and Benveniste, J. (2018) Coastal Altimetry Workshop (CAW11) Final report. ESA Publication. Available at: https://doi.org/10.5270/esa.caw11_2018.final_report.
- Ruiz-Etcheverry, L.A., Saraceno, M., Piola, A.R., Valladeau, G. and Moeller, O.O. (2015) A comparison of the annual cycle of sea level in coastal areas from gridded satellite altimetry and tide gauges. *Continental Shelf Research*, 92, 87–97. ISSN 18736955. Available at: <https://doi.org/10.1016/j.csr.2014.10.006>.
- Stammer, D., Ray, R.D., Andersen, O.B., Arbic, B.K., Bosch, W., Carrère, L *et al.* (2014) Accuracy assessment of global barotropic ocean tide models. *Reviews of Geophysics*, 52(3), 243–282. Available at: <https://doi.org/10.1002/2014RG000450>.
- Tamura, Y. (1987) A harmonic development of the tide-generating potential. *Bulletin d'Informations Marines Terrestres*, 99(99), 6813–6855.
- Taylor, J.R. (1997) *An Introduction to Error Analysis*, 2nd edition. Sausalito, CA: University Science Books. ISBN 978-0-935702-75-0.
- Volkov, D.L. and Pujol, M.I. (2012) Quality assessment of a satellite altimetry data product in the Nordic, Barents, and Kara seas. *Journal of Geophysical Research: Oceans*, 117(C3), C03025. ISSN 01480227. Available at: <https://doi.org/10.1029/2011JC007557>.
- Woodworth, P.L., Hunter, J.R., Marcos, M., Caldwell, P., Menéndez, M., Haigh, I. (2017) Towards a global higher-frequency sea level dataset. *Geoscience Data Journal*, 3(2), 50–59. ISSN 20496060. Available at: <https://doi.org/10.1002/gdj3.42>.
- Wunsch, C. (1986) Calibrating an altimeter: how many tide gauges is enough? *Journal of Atmospheric and Oceanic Technology*, 3, 746–754. Available at: [https://doi.org/10.1175/1520-0426\(1986\)003<0746:CAHMT>2.0.CO;2](https://doi.org/10.1175/1520-0426(1986)003<0746:CAHMT>2.0.CO;2).

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Piccioni G, Dettmering D, Bosch W, Seitz F. TICON: Tidal CONstants based on GESLA sea-level records from globally located tide gauges. *Geosci. Data J.* 2019;6:97–104. <https://doi.org/10.1002/gdj3.72>