Spatiotemporal Variability in Global Storm Surge and Tidal Water Levels from Satellite Radar Altimetry

Bij de Vaate, I.

DOI
10.4233/uuid:c9aa7e4e-5425-49cc-93ff-e53382402549

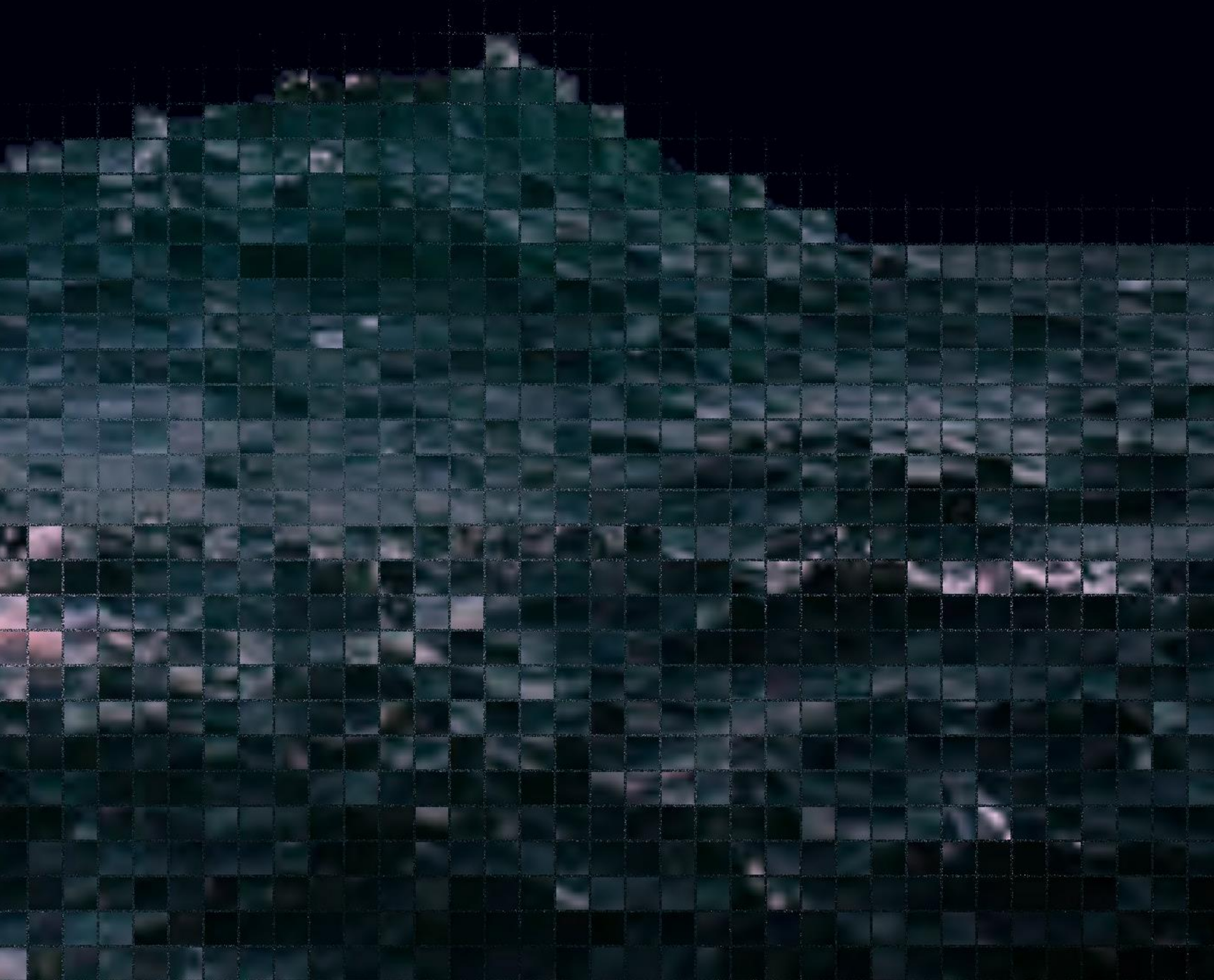
Publication date
2023

Document Version
Final published version

Citation (APA)
Bij de Vaate, I. (2023). *Spatiotemporal Variability in Global Storm Surge and Tidal Water Levels from Satellite Radar Altimetry*. [Dissertation (TU Delft), Delft University of Technology].
https://doi.org/10.4233/uuid:c9aa7e4e-5425-49cc-93ff-e53382402549

Spatiotemporal Variability in Global Storm Surge and Tidal Water Levels from Satellite Radar Altimetry

Inger Bij de Vaate



SPATIOTEMPORAL VARIABILITY IN GLOBAL STORM SURGE AND TIDAL WATER LEVELS FROM SATELLITE RADAR ALTIMETRY

SPATIOTEMPORAL VARIABILITY IN GLOBAL STORM SURGE AND TIDAL WATER LEVELS FROM SATELLITE RADAR ALTIMETRY

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus Prof. dr. ir. T.H.J.J. van der Hagen,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op vrijdag 21 april 2023 om 10:00 uur

door

Inger BIJ DE VAATE

Master of Science in Earth Sciences, Universiteit Utrecht te Utrecht,
geboren te Hengelo.

Dit proefschrift is goedgekeurd door de

promotor: Prof. Dr.-Ing. habil. R. Klees
promotor: Prof. dr. ir. M. Verlaan
copromotor: dr. ir. D.C. Slobbe

Samenstelling promotiecommissie:

Rector Magnificus,	voorzitter
Prof. Dr.-Ing. habil. R. Klees	Technische Universiteit Delft, promotor
Prof. dr. ir. M. Verlaan	Technische Universiteit Delft, promotor
dr. ir. D.C. Slobbe	Technische Universiteit Delft, copromotor

Onafhankelijke leden:

Prof. dr. I. Haigh	University of Southampton
Prof. Dr.techn. M. Schindelegger	Universität Bonn
Prof. dr. C.K. Shum	Ohio State University
Prof. dr. J.D. Pietrzak	Technische Universiteit Delft
Prof. dr. L.L.A. Vermeersen	Technische Universiteit Delft, reservelid

Keywords: tides, storm surges, global change, satellite radar altimetry, arctic sea ice

Printed by: ProefschriftMaken || www.proefschriftmaken.nl

Front & Back: Gridded visualization of the ocean and moon, representing the attempt to digitize natural processes.

ISBN 978-94-6469-294-5

An electronic version of this dissertation is available at
http://repository.tudelft.nl/.

Waar je voelt dat er niets is gemaakt om te blijven.
Waar je omkijkt en ziet hoe je sporen verdwijnen.
Want de golven die komen en de golven die gaan.
En de zee die beweegt op de maat van de maan.

Stef Bos: Witsand

CONTENTS

SUMMARY

Extreme (still) sea levels and the possibly associated coastal floods, are generally linked to (high) tides and storm surges. The risk of coastal floods will likely intensify in the future. This is because, on the one hand, the population of coastal zones is expected to continue to grow, and, on the other hand, climate change may lead to an increase in the frequency and magnitude of extreme sea levels. Although observations suggest that on the global scale, sea level rise is the primary driver behind the increase in extreme sea levels, locally the increase in extreme sea levels may be amplified or even dominated by changes in storm surges and tidal dynamics.

Previous studies that have demonstrated that tides and storm surges are subject to temporal changes, predominantly relied on in-situ data from tide gauges. The latter are scattered across the global coastlines, with the majority being located on the North-American and European continents. Some areas hardly comprise any tide gauges, the Arctic being a paramount example. While we know that this region is highly affected by climate change, knowledge about the Arctic tidal and surge water levels is sparse. As a consequence of the data gap over the global oceans and the Arctic in particular, the large-scale (open ocean) temporal variability in storm surges and tidal water levels remains uncertain. Obtaining the full-scale global picture would enable us to better understand the drivers behind the observed change which may ultimately contribute to better projections and the identification of the consequences for coastal environments.

In an attempt to fill the identified data gap, this thesis exploits the wealth of satellite radar altimetry data. While radar altimeter data have been occasionally used to study temporal variations in tides and surge water levels, they have not been used to study the seasonal or secular variability at a global scale. Although the relatively low resolution of the data may be problematic, the current length of the satellite altimeter records (> 25 years) allows to obtain estimates of the temporal changes in tides and surges. On the contrary, conventional low resolution radar altimetry (known as LRM altimetry) cannot be used to retrieve water levels in the largely ice-covered Arctic Ocean. In this light, the higher resolution Synthetic Aperture Radar altimetry has been proposed as a solution. However, as there is currently no cohesive data set of (reliable) instantaneous Arctic water level available, an appropriate processing scheme must be implemented. In this thesis the findings of four studies are presented, which all rely on satellite radar altimetry to increase our understanding of the temporal variability in global tides and surge.

In Chapter 2, satellite radar altimetry from the TOPEX/Poseidon & Jason series has been used to study worldwide linear trends in tidal harmonic constants of four major tides (M_2, S_2, O_1, and K_1). This study demonstrates both the potential and challenges of using satellite data for the quantification of such long-term changes. Two alternative methods were implemented. In the first method, tidal harmonic constants were estimated for consecutive four-year periods, through which a linear trend was then fitted. In the second method, the estimation of linear trends in the tidal constants was integrated in the harmonic analysis. The results show that the tides were subject to significant changes, with both decreases and increases in tidal amplitude of the order of several centimeters and phase changes of several degrees over the past decades. The level of agreement between altimeter-derived trends and estimates from tide gauge data has been found to differ per region and per tide.

In the third chapter, satellite altimetry data from multiple missions have been used to study the time-averaged, seasonal, and interannual variability in global storm surge water levels. For this purpose, a time-dependent extreme value distribution was fitted to the monthly maximum detided water levels. To account for the limited temporal resolution of the satellite data, the data were first stacked on a 5° × 5° grid and additional scaling was applied to the extreme value analysis. For the latter, the scaling factors were determined by means of a synthetic experiment using model data. Significant seasonal and secular variability in storm surge water levels was inferred across the global oceans. Where possible, the results were compared to similar analysis using in-situ data. Except for secular changes, the satellite-derived results were found to be comparable to the information derived from tide gauges, although the latter show more local variability. Where limited correlation was observed for the secular change, it was suggested that the satellites may not be able to fully capture the temporal variability in the short-lived, tropical storms, as opposed to extra-tropical storms.

The fourth chapter concerns the seasonal variability in the Arctic tide. Using water levels derived from Synthetic Aperture Radar altimetry (CryoSat-2 and Sentinel-3), the seasonal modulation of the M_2 tide was derived. Results were compared to numerical simulations that model the effect of two limiting cases of seasonal landfast ice cover on the M_2 tide. The largest seasonal modulation (up to 0.25 m) was observed along coastlines and in bays. Locally, the presence of landfast ice was found to decrease the tidal amplitude, while in some cases, the opposite effect was observed further afield. In most of the Arctic, winter months experienced a retarded arrival of the tide. While most of the altimeter-derived seasonal modulation could be explained by the modeled impact of landfast ice, discrepancies in several regions suggest that other seasonal processes are also important. Finally, the results suggested that the consequences of variations in Arctic landfast ice are not restricted to the Arctic but may affect tidal water levels on a global scale.

Finally, the fifth chapter describes a thorough assessment of various surface type classification methods that may be used for lead detection in the Arctic Ocean. With the ultimate goal to improve water level retrieval, this study compared the more traditional thresholding method to nine supervised-, and two unsupervised machine learning methods. In analysis of the Sentinel-3 SAR altimeter data, the study relied on the simultaneously sensed images from the Ocean and Land Color Instrument, onboard Sentinel-3, for training and validation of the classifiers. Applied to data from winter months, the supervised Adaptive Boosting, Artificial Neural Network, Naïve-Bayes, and

Linear Discriminant classifiers showed robust results. On the other hand, the unsupervised Kmedoids classifier produced excellent results and emerged as an attractive classifier when ground truth data are limited. All classifiers performed poorly on summer data, rendering surface classifications that are solely based on altimetry data from summer months unsuitable. Finally, the Adaptive Boosting, Artificial Neural Network, and Bootstrap Aggregation classifiers were found to perform best when the altimeter data also include measurements from the open ocean.

The work presented in this thesis is anticipated to contribute to a better understanding and ultimately, improved forecasting of future extreme sea levels. The findings aid in the assessment of the impact of climate change on tides and surges and can be used to improve the models that are used for forecasting extreme sea levels. In addition, progress has been made in understanding the role of sea ice in the Arctic sea level variability and tools have been provided for use in future studies on the Arctic.

SAMENVATTING

Extreme waterstanden en de mogelijke kustoverstromingen die daarmee gepaard gaan, worden doorgaans in verband gebracht met hoogtij en stormvloeden. Het risico op overstromingen langs de kust zal in de toekomst waarschijnlijk toenemen, enerzijds door de verwachtte bevolkingsgroei in de kustgebieden en anderzijds door een mogelijke toenamen van de frequentie en omvang van extreme waterstanden als gevolg van klimaatverandering. Hoewel waarnemingen suggereren dat, op wereldschaal, de zeespiegelstijging de belangrijkste oorzaak is van de stijging van de extreme zeespiegels, kan de toename in extreme waterstanden lokaal worden versterkt of zelfs gedomineerd door veranderingen in stormvloeden en getijdendynamiek.

Eerdere studies die hebben aangetoond dat getijden en stormvloeden onderhevig zijn aan veranderingen door de tijd, zijn voornamelijk gebaseerd op in-situ gegevens van getijmeters. Deze laatste bevinden zich langs de wereldwijde kustlijnen, maar met name op het Noord-Amerikaanse en Europese continent. In sommige gebieden zijn er nauwelijks getijmeters, waarvan de Arctische regio een belangrijk voorbeeld is. Hoewel we weten dat deze regio sterk wordt getroffen door klimaatverandering, is er weinig kennis over het Arctische getij en de hoge waterstanden. Als gevolg van de datakloof over de mondiale oceanen en het Noordpoolgebied in het bijzonder, is de grootschalige tijdsgebonden variabiliteit in stormvloeden en getijdenwaterstanden (op open oceaan) onzeker. Het verkrijgen van het volledige mondiale beeld zou ons beter in staat stellen om de drijvende krachten achter de waargenomen verandering beter te begrijpen, wat uiteindelijk kan bijdragen aan betere projecties en de identificatie van de gevolgen voor kustomgevingen.

In een poging de geïdentificeerde datakloven op te vullen, maakt dit proefschrift gebruik van de rijkdom aan satellietradarhoogtemetingen. Hoewel deze laatste af en toe zijn gebruikt om tijdsgebonden veranderingen in getijden en hoge waterstanden te bestuderen, zijn ze niet gebruikt om de (mondiale) seizoensgebonden of seculiere variabiliteit te bestuderen. Hoewel de relatief lage resolutie van de gegevens de mogelijkheden beperkt, zou de huidige lengte van de satelliethoogtemeterrecords (> 25 jaar) voldoende moeten zijn om schattingen te krijgen van de veranderingen in de getijden en stormvloeden over de tijd. De conventionele radarhoogtemeting kunnen echter niet gebruikt worden om de waterstanden in de grotendeels met ijs bedekte Noordelijke IJszee te achterhalen. In dit licht is verwacht dat de hogere resolutie Synthetic Aperture Radar altimetrie de oplossing kan bieden. Aangezien er momenteel echter geen samenhangende dataset van (betrouwbare) momentane Arctische waterstanden beschikbaar is, vereist dit de implementatie van een geschikt verwerkingsmethodiek. In dit proefschrift worden de bevindingen van vier afzonderlijke onderzoeken gepresenteerd, die allen, met behulp van satellietradar-altimetrie, proberen ons begrip van temporele variabiliteit in mondiale getijden en stormvloeden te vergroten.

In hoofdstuk 2 worden satellietradarhoogtemetingen van TOPEX/Poseidon & Jason-serie gebruikt om wereldwijde lineaire trends in de amplitudes en fases van vier grote getijdecomponenten te bestuderen (M_2, S_2 , O_1 en K_1). Deze studie toont zowel het potentieel als de uitdagingen aan van het gebruik van satellietgegevens voor de kwantificering van dergelijke langetermijnveranderingen. Er werden twee alternatieve methoden geïmplementeerd. Bij de eerste methode werden harmonische getijdenconstanten geschat voor opeenvolgende perioden van vier jaar, op basis waarvan de lineaire verandering werd geschat. In de tweede methode werd de schatting van lineaire trends in de getijconstanten van de vier getijden geïntegreerd in de harmonische analyse. Eerst werden beide methoden beoordeeld door toepassing op synthetische data op basis van getijdemetergegevens. Daarna werden de methoden toegepast op de echte satellietgegevens. De resultaten tonen aan dat de veranderingen in getijden significant zijn, met zowel afnames als toenames van de getijamplitude in de orde van enkele centimeters en faseveranderingen van enkele graden in de afgelopen decennia. De mate van overeenstemming tussen van hoogtemeters afgeleide trends en schattingen uit getijmetergegevens blijkt per regio en per getij te verschillen.

In het derde hoofdstuk zijn satellietaltimetriegegevens van meerdere missies gecombineerd om zowel de tijdsgemiddelde als de seizoensgebonden en interjaarlijkse variabiliteit in wereldwijde stormvloedwaterstanden te bestuderen. Dit is gedaan door middel van een tijdsafhankelijke extremewaardeanalyse die is toegepast op de maandelijkse maximale waterstanden. Om rekening te houden met de beperkte temporele resolutie van de satellietgegevens, werden de gegevens eerst gestapeld op een $5° \times 5°$ raster. Bovendien is extra schaling toegepast op de extremewaardeanalyse waarvoor de schalingsfactoren zijn bepaald met behulp van model data. Waar mogelijk werden de resultaten vergeleken met vergelijkbare analyses van in-situ data. Met uitzondering van seculiere veranderingen, zijn de satellietresultaten vergelijkbaar met de informatie die is afgeleid van getijdenmeters (correlatie $> 0,5$), hoewel de getijdenmeters meer lokale variabiliteit laten zien. Waar een beperkte correlatie werd waargenomen voor de seculaire verandering, werd gesuggereerd dat de satellieten mogelijk niet in staat zijn om de temporele variabiliteit in de kortdurende, tropische stormen volledig vast te leggen, in tegenstelling tot extra-tropische stormen.

Seizoensmodulatie van het M_2 getij is gekwantificeerd voor de gehele Noordelijke IJszee, met behulp van getijdenharmonische analyse van waterstanden afgeleid van Synthetic Aperture Radar-altimetrie (CryoSat-2 en Sentinel-3). De resultaten worden vergeleken met numerieke simulaties die het effect van twee uiterste gevallen van seizoensgebonden landvaste ijsbedekking op het M_2-getij modelleren. De grootste seizoensmodulatie (tot 0,25 m) is waargenomen langs kusten en in baaien. Lokaal vermindert de aanwezigheid van landijs de amplitude van het getij, maar in sommige gevallen werd verderop het tegenovergestelde effect waargenomen. In het grootste deel van het noordpoolgebied, behalve in de Hudsonbaai, is er in de wintermaanden sprake van een vertraagde aankomst van het getij. De waargenomen seizoensmodulatie kan grotendeels worden verklaard door de gemodelleerde impact van landvast ijs. Met name in de Hudsonbaai suggereren verschillen tussen beide producten, dat andere seizoensprocessen een rol spelen. Ten slotte laten de resultaten zien dat variaties in landvast Noordpoolzee-ijs niet beperkt zijn tot het Noordpoolgebied, maar van invloed zijn op de wereldwijde getijdewaterstanden.

Het afleiden van waterstanden in de Noordelijke IJszee vereist nauwkeurige detectie van scheuren in het zee-ijs. Het vijfde hoofdstuk beschrijft een uitgebreide beoordeling van verschillende classificatiemethoden voor oppervlaktetypes, waaronder een klassieke methode die gebruik maakt van drempelwaarden, negen gecontroleerde en twee ongecontroleerde machinale leermethoden, toegepast op Sentinel-3 Synthetic Aperture Radar Altimeter-data. Voor het eerst werden de gelijktijdig waargenomen beelden van het Ocean en Land Color Instrument, aan boord van Sentinel-3, gebruikt voor training en validatie van de classificatiemethoden. Wanneer toegepast op data van de wintermaanden, lieten de gecontroleerde classificaties Adaptive Boosting, Artificial Neural Network, Naïve-Bayes en Linear Discriminant robuuste resultaten zien. De ongecontroleerde Kmedoids-classificator produceerde ook uitstekende resultaten en is een aantrekkelijke classifier wanneer de controle data beperkt zijn. Waar alle classificaties slecht presteren op hoogtegegevens van zomermaanden, wordt aanbevolen om aanvullend gebruik te maken van metingen door andere sensoren. Ten slotte verkrijgen de classificaties Adaptive Boosting, Artificial Neural Network en Bootstrap Aggregation de hoogste nauwkeurigheid wanneer de hoogtemetingen ook metingen van de open oceaan omvatten.

Het werk gepresenteerd in dit proefschrift zal naar verwachting bijdragen aan een beter begrip en uiteindelijk betere voorspelling van toekomstige extreme waterstanden. De bevindingen helpen bij de beoordeling van de impact van klimaatverandering op het getijde en stormvloeden en kunnen worden gebruikt om de modellen te verbeteren die worden gebruikt voor het voorspellen van extreme waterstanden. Daarnaast is er vooruitgang geboekt in het verkrijgen van begrip over de rol van zee-ijs in de variabiliteit van de Arctische zeespiegel en zijn er instrumenten aangereikt voor gebruik in toekomstige studies.

1

INTRODUCTION

It storms outside. The wind blows with at least wind force 10 from the north-northwest. The North Sea crashes on the coasts of Zeeland. It's early in the morning when I wake up from loud banging on the door and the neighbor who yells: "The dike has broken!" Outside, land has turned to water. Father decides that we must flee. It is about to become high tide so the water will rise even more. Meanwhile, the sea is already flooding the country in high waves, destroying houses and killing people and animals.

Personal account of the North Sea flood of 1953 by Bart Bij de Vaate,
based on recordings by Inger Bij de Vaate (2003).

1.1. EXTREME SEA LEVELS AND COASTAL FLOODING

The North Sea flood of 1953 is an example of coastal flooding that truly wreaked havoc. The flood caused the loss of over 2,000 lives and its great destruction of houses and land forced over 100,000 people to evacuate (Hall, 2013). However, coastal floods have affected mankind for centuries and still do. Recent examples are the flooding of New Orleans (2005: about 1,800 fatalities), of Myanmar (2008: over 138,000 fatalities), and of the central Philippines (2013: over 6,000 fatalities) (Adeola, 2009; Kron, 2014). The large death tolls underline the severity of the threat that coastal floods pose to densely populated coastal areas. Already in 2007, over 600 million people lived in low lying coastal zones that are affected by the risk of coastal flooding (McGranahan et al., 2007), a number that has been monotonically increasing over the course of time.

Like the North Sea flood of 1953, coastal floods are typically associated with storm surges in combination with high tides (Muis et al., 2016). Tides - the rise and fall of the sea level principally caused by the gravitational force from the Moon and Sun (Wahr, 1995) - are a continuous process and an important contributor to the dynamics of the coastal zone. They affect the coastal morphology and marine life in coastal ecosystems and are often used to our advantage in, for example, fisheries or marine navigation.

Storm surges on the other hand, are an anomalous rise of the sea level due to low atmospheric pressure and strong winds, that are associated with (extra) tropical storms (Resio and Westering, 2008). The combined effect of high tide and a storm surge can result in extreme sea levels (ESLs) that may have aforementioned devastating effects.

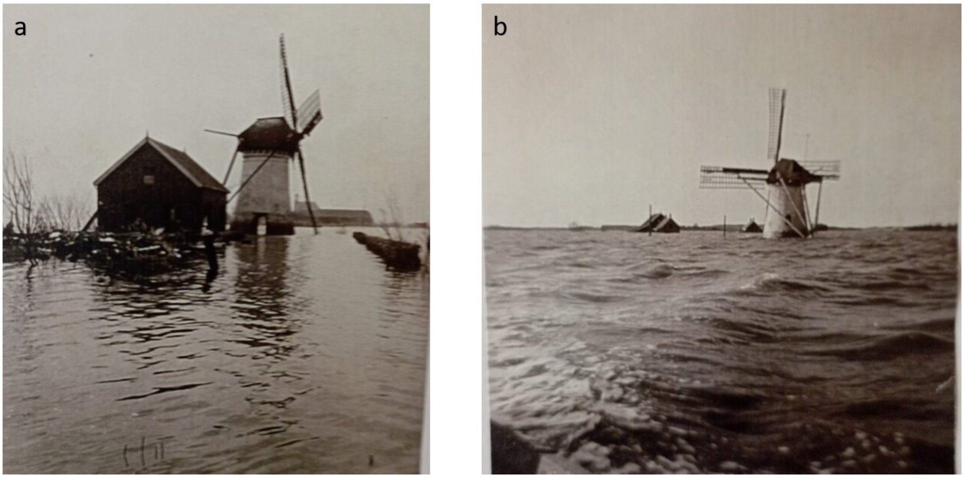


Figure 1.1: Windmill De Lelie in Ellemeet, Zeeland after the North Sea flood of 1953 at low (a) and high tide (b). Picture courtesy of Bart Bij de Vaate.

The risk of coastal flooding is expected to increase in the future for two main reasons. On the one hand, the world population has grown enormously over the past century and although the growth has slowed down, the population is expected to increase at least till the end of the century (Roser et al., 2013). Coastal areas experience a faster population growth than the hinterland as their rich resources and logistical advantages make them particularly attractive for habitation. In fact, most of the world's largest cities (e.g., Tokyo, Mumbai, New York City, Shanghai) are located in the coastal zone. The population of coastal zones is expected to increase in the future. Although the estimates of population growth vary depending on the input data and adopted methodologies (McMichael et al., 2020), up to 1.4 billion people could inhabit the low-elevated coastal zones by 2060 (McGranahan et al., 2007; Neumann et al., 2015). On the other hand, the risk of coastal flooding is expected to intensify under the influence of climate change. The rising sea level in combination with atmospheric circulation changes under global warming will most definitely increase the likelihood of ESLs (IPCC, 2021).

1.2. CHANGES IN EXTREME SEA LEVELS, TIDES AND SURGES

Observations show that ESL events have increased both in frequency and magnitude over the past decades and projections indicate this trend will continue in the future (e.g., Oppenheimer et al., 2019; Wahl et al., 2017). While observations suggest that on the global scale, sea level rise is the primary driver behind the increase in ESLs (e.g., Oppenheimer et al., 2019; Wahl et al., 2017), locally the increase in ESLs may be amplified or even dominated by changes in tidal dynamics (e.g., Devlin et al., 2017) or changes in storm surges (e.g., J. Feng et al., 2019; Haigh et al., 2014; Rashid and Wahl, 2020). Note that in addition to the effect of the mean sea level variability, tides and surges, waves also contribute to ESLs (IPCC, 2021). However, in this thesis, ESLs refers to the extreme still water levels, thus excluding the very local effect of waves and wave set-up (Caires, 2011).

An increasing number of studies indicate that both tidal and storm surge water levels are subject to considerable changes on seasonal to long time scales (e.g., tide: Mawdsley et al., 2015; Müller et al., 2011; Müller et al., 2014; R. D. Ray, 2016; Woodworth, 2010 and surge: Bromirski et al., 2003; X. Feng et al., 2018; Grinsted et al., 2012; Wahl and Chambers, 2015; Zhang et al., 2000). However, in many cases, the identification of the drivers behind the observed changes remains challenging (e.g., Haigh et al., 2020; Woodworth, 2010). For instance, changing tidal dynamics may be affected by local processes, such as modifications of the shoreline or altered river flows, but they have also been linked to climatic processes (Haigh et al., 2020). The latter include changes in ocean stratification in response to ocean warming (e.g., Müller, 2012; Müller et al., 2014; Schindelegger et al., 2022), in the radiational forcing (e.g., R. D. Ray and Egbert, 2004; Schindelegger et al., 2016), and in the mean sea level (Devlin et al., 2017; Pickering et al., 2017; Schindelegger et al., 2018). Changes in storm surges have been linked to dominant modes of climate variability such as ENSO and NAO (e.g., Mawdsley and Haigh, 2016; Wahl and Chambers, 2016). In addition, several studies suggest that climate change affects the frequency and severity of (extra)tropical storms, which may exacerbate the build up of storm surges in the future (e.g., Calafat et al., 2022; Gori et al., 2022). However, in the latest IPCC report (IPCC, 2021), long-term changes in storminess are considered of low confidence, which is predominantly related to the rare and short-lived nature of the storm events and the large local variability (IPCC, 2021).

Another important climatic process that affects the temporal variability in both tides and surges, and hence ESLs, is fluctuations in the extent of sea ice cover. Where sea ice exerts a frictional stress on the ocean surface, it influences the tidal variability. Model studies showed that the presence of sea ice can cause tidal dampening, leading to tidal amplitude decay and phase delay (e.g., Godin, 1986; Kowalik and Untersteiner, 1978; Murty, 1985). At the same time, the presence of a sea ice cover shields the ocean surface from the wind. Hence, a reduction in sea ice cover would result in a greater fetch and as such allows higher storm surges to reach the shore (Kim et al., 2021; Lintern et al., 2013; Vermaire et al., 2013). The impact of sea ice on temporal variability in water levels is particularly important in the Arctic Ocean. While this region is largely covered by sea ice during winter months (up to 15 million km^2, average 2011-2020 (National Snow and Ice Data Center, 2020)), it undergoes a significant loss of coverage during summer (5 million km^2). In addition, the Arctic Ocean is highly affected by global warming, experiencing warming rates that are up to fourfold that of the global average (Kwok, 2018). As a result, the (perennial) Arctic sea ice cover has been subject to a dramatic decrease since the 1970s, while this does not apply to the Antarctic sea ice (e.g., IPCC, 2021). So far, the effect of variations in sea ice cover on tidal and surge water levels has mainly been observed on seasonal timescales (e.g., Müller et al., 2014; St-Laurent et al., 2008) and the long-term effect of the Arctic sea ice decline is unsure. However, any changes in Arctic tides and surges may not only pose higher risks to fragile Arctic ecosystems in low-lying areas (e.g., Kokelj et al., 2012), but may also affect global water levels. Namely, it has been indicated that local changes in tidal dynamics may have a remote effect on connected basins (e.g., Arbic et al., 2007; Arbic et al., 2010; Harker et al., 2019) and storm surges that develop in the Arctic could potentially travel to lower latitudes (e.g., Budikova, 2009; Zhao et al., 2019).

1.2.1. OBSERVATIONAL EVIDENCE

Until now, the observational evidence that indicates that ESLs, tides and storm surges are subject to temporal changes, predominantly consists of in-situ data from tide gauges (e.g., Bromirski et al., 2003; X. Feng et al., 2018; Grinsted et al., 2012; Mawdsley et al., 2015; Müller et al., 2011; R. D. Ray, 2016; Wahl and Chambers, 2015; Woodworth, 2010; Zhang et al., 2000). Historically, these tide gauges were installed in ports, where their measurements were of direct use for maritime navigation (Woodworth et al., 2011). Today, the majority of the tide gauges is still located at the coast or in shelf waters, with uneven spatial distribution across the globe. Where water level variations recorded by tide gauges are influenced by local processes, the large scale patterns in storm surge and tidal water levels remain uncertain.

The Arctic Ocean is a key example of an area where in-situ recordings of the water level are limited. The Arctic is a difficult environment for data acquisition. The region is isolated, sparsely populated and subject to severe weather, while the ocean itself is greatly affected by the dynamic and seasonally variable sea ice cover. For example, there are only twelve tide gauges located along the coastlines of the Arctic Ocean (over 45,000 km; Atland, 2013) of which recent data are publicly available (Figure 1.2). This is a significant contrast to the 350 km of Dutch coastlines (Roeland and Piet, 1995) that host 56 tide gauges. Consequently, there is only limited observational support for the expected impact of climate change on Arctic tidal and surge water levels.

Obtaining the full-scale global picture of the temporal variability in tides and storm surges would enable us to better understand the drivers behind the observed changes. This will ultimately improve projections and the identification and prediction of any consequences for coastal environments.

1.3. SPATIOTEMPORAL VARIABILITY IN TIDES AND SURGES FROM SATELLITE RADAR ALTIMETRY

In addition to tide gauges, satellite radar altimeters have been measuring the quasi-global water level since 1992. In the basic measurement principle of satellite altimetry, a satellite orbits the Earth and continuously sends radar signals to the surface. Knowing the altitude of the satellite with respect to a specified reference ellipsoid, the timing of the reflection of the radar signal can be used to compute the instantaneous height of the sea surface (with respect to the same reference surface). TOPEX/Poseidon (in or-

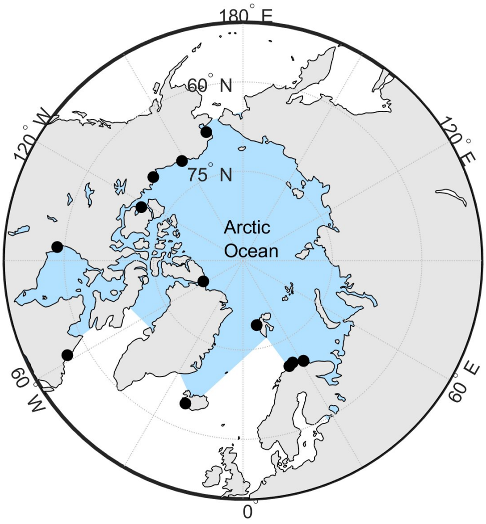


Figure 1.2: Map of Arctic Ocean according to the limits defined by International Hydrographic Organization (1953), showing all tide gauges (black scatters) that have recent data (after 2000) publicly available.

bit from 1992-2002) was the first in a series of missions providing continuous measurements of sea levels, followed by Jason-1 (2002-2008), Jason-2 (2008-2017), and Jason-3 (2016-present). All satellites of this series (further referred to as TPJ) orbit the Earth with a 10-day revisit period, an along track resolution of about 6 km and global coverage up to 66° N/S. The data from the TPJ satellites have often been used for tidal analyses (e.g., Cherniawsky et al., 2010; Müller et al., 2014; R. D. Ray, 2016). In addition to TPJ, the ERS-1 (1991-1995), ERS-2 (1995-2003), Envisat (2002-2008) and SARAL (2013-present) missions are also equipped with a radar altimeter with the primary purpose to map the ocean, land, and sea ice topography (van't Klooster, 2011). These missions follow a higher-inclined orbit around the Earth, resulting in spatially denser data that also include high latitudes (up to 82° N/S) yet at a lower temporal resolution (35-days revisit period). Although the low temporal resolution makes the data less suitable for tidal analysis, they have been successfully used in combination with the TPJ-data to increase spatial data density (e.g., Lyard et al., 2021).

Although satellite altimetry is widely used for tidal studies and estimation of stationary tidal properties, there have only been few efforts to derive temporal variability in tides. The latter includes studies of the nodal cycle (e.g., Cherniawsky et al., 2010) and local studies on the seasonal modulation of the major tide (e.g., Müller et al., 2014). Up to now, radar altimetry has never been used to study long-term trends in tidal constants. As the relatively low resolution of the data causes aliasing of high-frequency signals (higher than half the sampling frequency) on lower frequencies, several years of data are required to be able to distinguish the major tides (Chen and Lin, 2000). Consequently, studying seasonal or secular variability in the tidal signal cannot be done by means of respectively monthly or yearly tidal analysis and subsequent curve fitting, as is commonly done with data from tide gauges. However, given the current length of the satellite altimeter records (> 25 years), from a theoretical point of view it should be possible to obtain estimates of the secular changes in tides from these data.

Regarding storm surges, the situation is not much different. Up to now, there have been few studies that focused on single storm surge events captured by satellite altimetry (Han et al., 2012; Lin et al., 2013; Scharroo et al., 2005). In addition, only a number of studies have attempted to use satellite altimetry to map storm surges over an extended period (e.g., Andersen et al., 2015; Antony et al., 2014; Ji et al., 2019). The latter studies showed that with adequate data processing techniques it is possible to derive storm surge properties (e.g., magnitude and spatial features) from the low resolution satellite data comparable to those from high-frequency tide gauge data. However, as the aforementioned studies considered only specific geographic regions and did not cover temporal changes in storm surges, the large scale spatiotemporal variability remains uncertain. Nevertheless, it should be possible to derive this information from the wealth of multi-mission altimeter data that are currently available.

1.3.1. SAR ALTIMETRY IN THE ARCTIC REGION

In the Arctic Ocean, the applicability of radar altimetry is hampered by the presence of the temporally dynamic (perennial) sea ice cover. The footprint of so-called low-resolution-mode (LRM) altimeters (i.e., all aforementioned missions) is too big to distinguish between measurements from the sea ice pack and alternating sections of open water (leads). However, in this respect, Synthetic Aperture Radar (SAR) altimetry provides a solution. SAR altimeters apply coherent processing of groups of transmitted pulses, using the Delay-Doppler effect (Raney, 1998). This results in a higher along-track resolution compared to conventional radar altimeters (Donlon et al., 2012), which allows measuring the water level from leads. CryoSat-2 was the first satellite equipped with a SAR altimeter. Since the launch of this satellite (2010-present), followed by Sentinel-3A (2016-present) and Sentinel-3B (2018-present) - the SAR data availability in the Arctic has strongly increased (Quartly et al., 2018). However, to use these data for derivation of water levels from leads, careful discrimination between measurements from different surface types is required.

For this purpose, various classification methods have been implemented that rely on the fact that different surfaces result in characteristic radar return signals or so-called waveforms (Figure 1.3). Empirical methods, that use certain thresholds for specific features of the waveforms, have been widely used to classify radar returns (e.g., Laxon, 1994; Peacock and Laxon, 2004; Poisson et al., 2018; Zakharova et al., 2015). More recently, machine learning-based classification methods have gained popularity (e.g., Dettmering et al., 2018; Lee et al., 2016; Muller et al., 2017; Poisson et al., 2018). Studies suggest that machine learning-based methods can produce higher accuracies, although performances obtained by different studies cannot be directly compared due to e.g., differences in study area, satellite mission, or validation data. In addition, the validation in previous studies was often limited because the ground truth data was based on manual interpretation of imagery data (Quartly et al., 2018). Moreover, some of the aforementioned studies make use of validation data that are not acquired simultaneously with the SAR data, which may introduce additional uncertainties (Quartly et al., 2018). In this respect, the use of measurements from the Sentinel-3 satellites is particularly interesting as these missions are also equipped with an optical imaging instrument. So far, the optical imagery from Sentinel-3 has not been applied in lead detection.

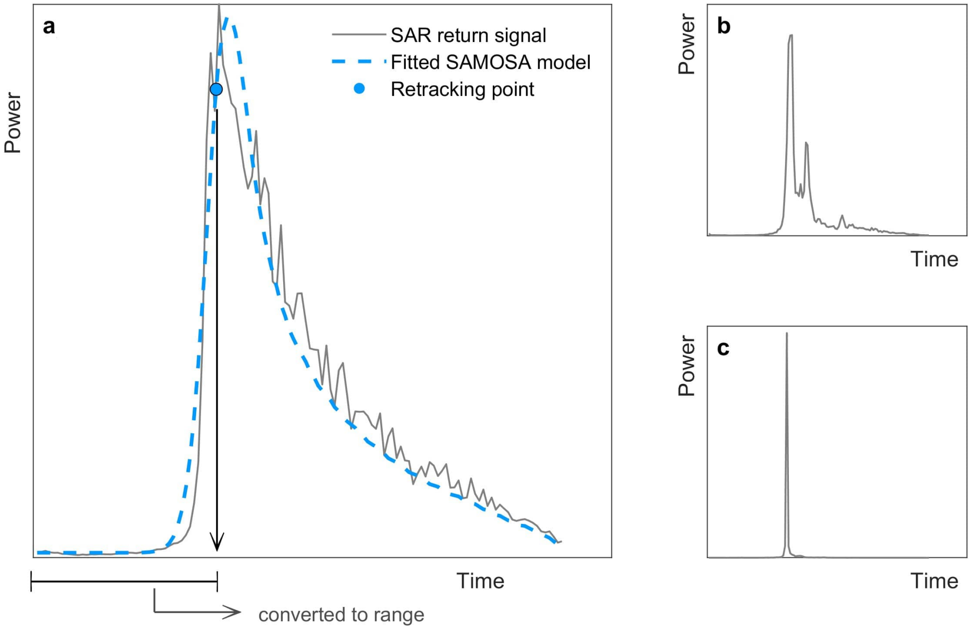


Figure 1.3: Typical SAR waveforms of radar return signals from open ocean (a), sea ice (b) and a lead (c). Note that over sea-ice regions various complex waveform shapes can be observed that relate to different types/forms of sea ice or signal contamination. The markings on (a) indicate how the waveform can be used to compute a water level.

After classification, the radar signals are converted to water levels by finding the so-called retracking point of the waveform and then use the timing of that point to compute the instantaneous distance between the satellite and the surface (see Figure 1.3a for an example). For open ocean returns there are many widely implemented methods to obtain the retracking point. However, the specular waveform of reflections from leads (Figure 1.3c) requires the implementation of another method. Recent studies have provided suitable retrackers for leads such as: ALES+ (Passaro et al., 2018) or the physical retracker developed by Poisson et al. (2018). Both retrackers rely on an adapted version of the Brown-Hayne model (Brown, 1977; Hayne, 1980) that allows specular curve fitting. These retrackers have so far only been applied to LRM altimeter data (ENVISAT and ERS-2) and can not be directly used for SAR data (Kurtz et al., 2014). On the other hand, several studies have focused on retracking of lead waveforms from CryoSat-2. For instance, Jain et al. (2014) obtained better performance over sea ice-affected regions by adapting the traditional empirical retrackers, that are typically applied to ocean data, to solely focus on the primary peak of the waveform. In addition, Kurtz et al. (2014) developed a physical retracker specifically for CryoSat-2 data that also showed improved performance compared to the more subjective empirical retrackers in use. However, their CS2WfF-retracker has been predominantly validated for sea ice freeboard measurements rather than leads and had a bias in the order of centimeters. An alternative was introduced by Dinardo et al. (2018), whose retracker makes use of the SAMOSA-model (C. Ray et al., 2015) that is commonly used for retracking of ocean returns, but fixes the significant wave height (one of the model parameters to be fitted) to zero. This retracker resulted in slightly more precise results than the primary peak retrackers.

Despite the efforts on retrieving Arctic water levels, general consensus on the best classifier/retracker is lacking and there is as of yet no cohesive dataset of (reliable) instantaneous Arctic water levels available. Although earlier work has used SAR-derived water levels (CryoSat-2) to study the Arctic tide (Cancet et al., 2018; Cancet et al., 2015), the data have not been used to study the temporal variability in the Arctic tide, nor (the temporal variability of) surge characteristics.

1.4. RESEARCH OBJECTIVES

The main aim of this thesis is to assess the spatiotemporal variability of global tide and storm surge characteristics over the last three decades, using satellite radar altimetry. Hereby the thesis attempts to fill the gaps of earlier studies on the topic that predominantly relied on in-situ data. In addition, special attention is paid to the Arctic region. In the previous discussion it has become apparent that the Arctic tides and surges are expected to be greatly affected by climate-related processes, while still much progress ought to be made in actually observing the Arctic tides and surge water levels. The objective of this thesis can be further detailed by means of the following research questions:

1. **What is the radar altimetry-derived secular variability in global tides, and the related uncertainty, and how does this compare to the secular variability derived from tide gauge data?** To answer this question, 28 years of water level measurements by the TPJ-satellites (1993-2020) were used to study secular (linear) trends in the tidal amplitudes and phases of four major tides (M_2, S_2, K_1, and O_1). In addition, the uncertainty was assessed by means of repeated subsampling of reanalysis data. The obtained results were compared to the secular variability derived from a global set of tide gauge data (GESLA3; Haigh et al. (2022)).

2. **What is the radar altimetry-derived spatiotemporal variability in global storm surge characteristics, including their time-averaged magnitude and seasonal and secular variability, and how does this compare to the analysis of tide gauge data?** Here, monthly maximum detided water levels, derived from the combination of eight satellite radar altimeters (TPJ-satellites, ERS-1, ERS-2, Envisat and SARAL; 1993-2021), entered an extreme value analysis. In order to assess the seasonal and secular variability in storm surge water levels, we estimated a time-dependent generalized extreme value distribution. The results were compared to results obtained for a global set of tide gauge data (GESLA3; Haigh et al. (2022)).

3. **What is the SAR altimetry-derived seasonal variability in the Arctic tide and to what extent can this be linked to changes in the Arctic sea ice cover?** In light of this question, a threshold classification scheme and a physical retracker were implemented to retrieve Arctic water levels from the SAR data provided by CryoSat-2 and Sentinel-3 (2010-2019). Subsequently, the seasonal modulation of the M_2 tide was derived and compared to numerical simulations that model the effect of two limiting cases of seasonal landfast ice cover on the tides.

4. **Can machine learning classification methods contribute to a better lead detection from SAR altimetry in the Arctic Ocean?** Whilst in response to the previous question a traditional threshold classifier was implemented, the purpose of this question is to assess the added value of machine learning-based classification methods. Therefore, the performance of nine supervised- and two unsupervised classifiers in lead detection from Sentinel-3 SAR altimetry, were compared to that of the threshold classifier. Optical imagery obtained by the same satellite was used to train and validate the classification and the classification performance was assessed using pan-Arctic and full-year data.

1.4.1. OUTLINE OF THE THESIS

The content of this thesis is based on three published journal articles (Chapter 2, 4, and 5) and one submitted article (Chapter 3). Chapter 2 and 3 address respectively the first and second research question. They present global temporal changes in the tide (Chapter 2) and surge water levels (Chapter 3), based on LRM satellite altimetry-derived water levels. Chapter 4 and 5 zoom in on the Arctic, where Chapter 4 considers the seasonal modulation of the major Arctic tide under the influence of seasonal variations in the Arctic sea ice extent. Chapter 5 addresses the fourth research question by providing a thorough assessment of different classification techniques to distinguish between SAR return signals from respectively leads, open ocean and sea ice. An overarching summary with conclusions, recommendations and a reflection on the societal and scientific value of this work, is given in Chapter 6.

BIBLIOGRAPHY

Adeola, F. O. (2009). Mental health & psychosocial distress sequelae of Katrina: an empirical study of survivors. *Human Ecology Review*, 195–210.

Andersen, O. B., Cheng, Y., Deng, X., Steward, M., & Gharineiat, Z. (2015). Using satellite altimetry and tide gauges for storm surge warning. *Proceedings of the International Association of Hydrological Sciences*, *365*, 28–34.

Antony, C., Testut, L., & Unnikrishnan, A. S. (2014). Observing storm surges in the Bay of Bengal from satellite altimetry. *Estuarine, Coastal and Shelf Science*, *151*, 131–140.

Arbic, B. K., St-Laurent, P., Sutherland, G., & Garrett, C. (2007). On the resonance and influence of the tides in Ungava Bay and Hudson Strait. *Geophysical Research Letters*, *34*(17).

Arbic, B. K., Wallcraft, A. J., & Metzger, E. J. (2010). Concurrent simulation of the eddying general circulation and tides in a global ocean model. *Ocean Modelling*, (3–4), 175–187.

Atland, K. (2013). The security implications of climate change in the Arctic Ocean. Dordrecht: Springer.

Bromirski, P. D., Flick, R. E., & Cayan, D. R. (2003). Storminess variability along the California coast: 1858–2000. *Journal of Climate*, *16*(6), 982–993.

Brown, G. (1977). The average impulse response of a rough surface and its applications. *IEEE transactions on antennas and propagation*, *25*(1), 67–74.

Budikova, D. (2009). Role of Arctic sea ice in global atmospheric circulation: A review. *Global and Planetary Change*, *68*(3), 149–163.

Caires, S. (2011). Extreme value analysis: Still water level. *Geneva, Switzerland: World Meteorological Organization, Intergovernmental Oceanographic Commission (of UNESCO), JCOMM Technical Report*, (58).

Calafat, F. M., Wahl, T., Tadesse, M. G., & Sparrow, S. N. (2022). Trends in Europe storm surge extremes match the rate of sea-level rise. *Nature*, *603*(7903), 841–845.

Cancet, M., Andersen, O., Lyard, F., Cotton, D., & Benveniste, J. (2018). Arctide2017, a high-resolution regional tidal model in the Arctic Ocean. *Advances in space research*, *62*(6), 1324–1343.

Cancet, M., Andersen, O. B., Cotton, D., Lyard, F., & Benveniste, J. (2015). High resolution tidal modeling in the Arctic Ocean : needs and upcoming developments. *17*, 7231.

Chen, G., & Lin, H. (2000). The effect of temporal aliasing in satellite altimetry. *Photogrammetric engineering and remote sensing*, *66*(5), 639–644.

Cherniawsky, J. Y., Foreman, M. G., Kang, S. K., Scharroo, R., & Eert, A. J. (2010). 18.6-year lunar nodal tides from altimeter data. *Continental Shelf Research*, *30*(6), 575–587.

Dettmering, D., Wynne, A., Müller, F. L., Passaro, M., & F., S. (2018). Lead detection in polar oceans-a comparison of different classification methods for CryoSat-2 SAR data. *Remote Sensing, 10*. https://doi.org/doi:10.3390/rs10081190

Devlin, A. T., Jay, D. A., Talke, S. A., Zaron, E. D., Pan, J., & Lin, H. (2017). Coupling of sea level and tidal range changes, with implications for future water levels. *Scientific Reports*, *7*(1), 1–12.

Dinardo, S., Fenoglio-Marc, L., Buchaupt, C., Becker, M., Scharroo, R., Fernandes, M. J., & Benveniste, J. (2018). Coastal SAR and PLRM altimetry in German Bight and West Baltic Sea. *Advances in Space Research*, *62*(6), 1371–1404.

Donlon, C., Berruti, B., Buongiorno, A., Ferreira, M. H., Féménias, P., Frerick, J., ..., & Nieke, J. (2012). The global monitoring for environment and security (GMES) sentinel-3 mission. *Remote Sensing of Environment*, *120*, 37–57.

Feng, J., Li, D., Wang, T., Liu, Q., Deng, L., & Zhao, L. (2019). Acceleration of the Extreme Sea Level Rise Along the Chinese Coast. *Earth and Space Science*, *6*(10), 1942–1956. https://doi.org/doi:10.1029/2019ea000653

Feng, X., Li, M., Yin, B., Yang, D., & Yang, H. (2018). Study of storm surge trends in typhoon-prone coastal areas based on observations and surge-wave coupled simulations. *International Journal of Applied Earth Observation and Geoinformation*, *68*, 272–278.

Godin, G. (1986). Modification by an ice cover of the tide in James Bay and Hudson Bay. *Arctic*, 65–67.

Gori, A., Lin, N., Xi, D., & Emanuel, K. (2022). Tropical cyclone climatology change greatly exacerbates us extreme rainfall–surge hazard. *Nature Climate Change*, *12*(2), 171–178.

Grinsted, A., Moore, J. C., & Jevrejeva, S. (2012). Homogeneous record of Atlantic hurricane surge threat since 1923. *Proceedings of the National Academy of Sciences*, *109*(48), 19601–19605.

Haigh, I. D., MacPherson, L. R., Mason, M. S., Wijeratne, E. M. S., Pattiaratchi, C. B., Crompton, R. P., & George, S. (2014). Estimating present day extreme water level exceedance probabilities around the coastline of Australia: tropical cyclone-induced storm surges. *Climate Dynamics*, *42*(1), 139–157.

Haigh, I. D., Marcos, M., Talke, S. A., Woodworth, P. L., Hunter, J. R., Haugh, B. S., ..., & Thompson, P. (2022). GESLA Version 3: A major update to the global higher-frequency sea-level dataset. [Dataset]. *Geoscience Data Journal*. https://doi.org/https://doi.org/10.1002/gdj3.174

Haigh, I. D., Pickering, M. D., Green, J. M., Arbic, B. K., Arns, A., Dangendorf, S., ..., & Woodworth, P. L. (2020). The tides they are a-Changin': A comprehensive review of past and future nonastronomical changes in tides, their driving mechanisms, and future implications. *Reviews of Geophysics*, *58*(1), e2018RG000636.

Hall, A. (2013). The North Sea flood of 1953. *Arcadia*, *5*, 148.

Han, G., Ma, Z., Chen, D., Deyoung, B., & Chen, N. (2012). Observing storm surges from space: Hurricane Igor off Newfoundland. *Scientific reports*, *2*(1), 1–9.

Harker, A., Green, J., Schindelegger, M., & Wilmes, S.-B. (2019). The impact of sea-level rise on tidal characteristics around Australia. *Ocean Science*, *15*(1), 147–159.

Hayne, G. S. (1980). Radar altimeter mean return wave forms from near-normal-incidence ocean surface scattering.

International Hydrographic Organization. (1953). *Limits of oceans and seas.*

IPCC. (2021). *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]* Cambridge University Press.

Jain, M., Martin-Puig, C., Andersen, O. B., Stenseng, L., & Dall, J. (2014). Evaluation of SAMOSA3 adapted retracker using CryoSat-2 SAR altimetry data over the Arctic ocean. *IEEE Geoscience and Remote Sensing Symposium*, 5115–5118.

Ji, T., Li, G., & Zhang, Y. (2019). Observing storm surges in China's coastal areas by integrating multi-source satellite altimeters. *Estuarine, Coastal and Shelf Science, 225*.

Kim, J., Murphy, E., Nistor, I., Ferguson, S., & Provan, M. (2021). Numerical analysis of storm surges on Canada's western Arctic coastline. *Journal of Marine Science and Engineering, 9*(326).

Kokelj, S. V., Lantz, Solomon, S., Pisaric, D., M. F. J. andKeith, Morse, P., Thienpont, J. R., Smol, J. P., & D., E. (2012). Utilizing multiple sources of knowledge to investigate northern environmental change: Regional ecological impacts of a storm surge in the outer Mackenzie Delta. *Arctic*, *65*, 257–272.

Kowalik, Z., & Untersteiner, N. (1978). A study of the M2 tide in the Arctic Ocean. *Deutsche bHydrografische Zeitschrift*, *31*(6), 216–229.

Kron, W. (2014). Flood Risk–a Global Problem. *In ICHE 2014. Proceedings of the 11th International Conference on Hydroscience & Engineering*, 9–18.

Kurtz, N. T., Galin, N., & Studinger, M. (2014). An improved CryoSat-2 sea ice freeboard retrieval algorithm through the use of waveform fitting. *The Cryosphere*, *8*(4), 456–471.

Kwok, R. (2018). Arctic sea ice thickness, volume, and multiyear ice coverage: losses and coupled variability (1958–2018). *Environmental Research Letters*, *13*(10), 105005.

Laxon, S. (1994). Sea ice altimeter processing scheme at the EODC. *International Journal of Remote Sensing*, *15*(4), 915–924.

Lee, S., Im, J., Kim, J., Kim, M., M. Shin, Kim, K., & L., Q. (2016). Arctic Sea Ice Thickness Estimation from CryoSat-2 Satellite Data Using Machine Learning-Based Lead Detection. *Remote Sensing*, *8*. https://doi.org/doi:10.3390/rs8090698

Lin, M., Lillibridge, J., & Shum, C. (2013). Hurricane Sandy storm surge measured by satellite altimetry. *Oceanography*.

Lintern, D. G., Macdonald, R. W., Solomon, S. M., & Jakes, H. (2013). Beaufort Sea storm and resuspension modeling. *Journal of Marine Systems*, *127*, 14–25.

Lyard, F. H., Allain, D. J., Cancet, M., Carrère, L., & Picot, N. (2021). FES2014 global ocean tide atlas: design and performance. *Ocean Science*, *17*(3), 615–649.

Mawdsley, R. J., & Haigh, I. D. (2016). Spatial and temporal variability and long-term trends in skew surges globally. *Frontiers in Marine Science*, *3*, 29.

Mawdsley, R. J., Haigh, I. D., & Wells, N. C. (2015). Global secular changes in different tidal high water, low water and range levels. *Earth's Future*, *3*(2), 66–81.

McGranahan, G., Balk, D., & Anderson, B. (2007). The rising tide: assessing the risks of climate change and human settlements in low elevation coastal zones. *Environment and urbanization*, *19*(1), 17–37.

McMichael, C., Dasgupta, S., Ayeb-Karlsson, S., & Kelman, I. (2020). A review of estimating population exposure to sea-level rise and the relevance for migration. *Environmental Research Letters*, *15*(12), 123005.

Muis, S., Verlaan, M., Winsemius, H. C., Aerts, J. C., & Ward, P. J. (2016). A global reanalysis of storm surges and extreme sea levels. *Nature Communications*, *7*(1), 1–12.

Muller, F. L., Dettmering, D., Bosch, W., & Seitz, F. (2017). Monitoring the arctic seas: How satellite altimetry can be used to detect open water in sea-ice regions. *Remote Sensing*, *9*, 1–20. https://doi.org/doi:10.3390/rs9060551

Müller, M. (2012). The influence of changing stratification conditions on barotropic tidal transport and its implications for seasonal and secular changes of tides. *Continental Shelf Research*, *47*, 107–118. https://doi.org/https://doi.org/10.1016/j.csr.2012.07.003

Müller, M., Arbic, B. K., & Mitrovica, J. X. (2011). Secular trends in ocean tides: Observations and model results. *Journal of Geophysical Research: Oceans*, *116*(C5).

Müller, M., Cherniawsky, J. Y., Foreman, M. G. G., & von Storch, J. S. (2014). Seasonal variation of the M 2 tide. *Ocean Dynamics*, *64*(2), 159–177. https://doi.org/https://doi.org/10.1007/s10236-013-0679-0

Murty, T. S. (1985). Modification of hydrographic characteristics, tides, and normial modes by ice cover. *Marine Geodesy*, *9*(4), 451–468.

National Snow and Ice Data Center. (2020). Charctic Interactive Sea Ice Graph [Last checked on 06/06/2019].

Neumann, B., Vafeidis, A. T., Zimmermann, J., & Nicholls, R. J. (2015). Future coastal population growth and exposure to sea-level rise and coastal flooding-a global assessment. *PloS one*, *10*(3), e0118571.

Oppenheimer, M., Glavovic, B., Hinkel, J., van de Wal, R., Magnan, A. K., Abd-Elgawad, A., ..., & Sebesvari, Z. (2019). Sea level rise and implications for low lying islands, coasts and communities. *In: IPCC Special Report on the Ocean and Cryosphere in a Changing Climate [Pörtner, H.-O., D.C. Roberts, V. Masson- Delmotte, P. Zhai, M. Tignor, E. Poloczanska, K. Mintenbeck, M. Nicolai, A. Okem, J. Petzold, B. Rama, and N. Weyer (eds.)]*, 321–445.

Passaro, M., Rose, S. K., Andersen, O. B., Boergens, E., Calafat, F. M., Dettmering, D., & Benveniste, J. (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.

Peacock, N., & Laxon, S. (2004). Sea surface height determination in the Arctic Ocean from ERS altimetry. *Journal of Geophysical Research: Oceans*, *109*, 1–14. https://doi.org/doi:10.1029/2001JC001026

Pickering, M. D., Horsburgh, K. J., Blundell, J. R., Hirschi, J. M., Nicholls, R. J., Verlaan, M., & Wells, N. C. (2017). The impact of future sea-level rise on global tides. *Continental Shelf Research*, *142*, 50–68.

Poisson, J. C., Quartly, G. D., Kurekin, A. A., Thibaut, P., Hoang, D., & Nencioli, F. (2018). Development of an ENVISAT altimetry processor providing sea level continuity between open ocean and Arctic leads. *IEEE Transactions on Geoscience and Remote Sensing*, *56*(9), 5299–5319.

Quartly, G. D., Rinne, E., Passaro, M., Andersen, O. B., Dinardo, S., Fleury, S., & Tsamados, M. (2018). Review of radar altimetry techniques over the Arctic ocean: Recent progress and future opportunities for sea level and sea ice research. *The Cryosphere Discussions*, 1–51.

Raney, R. K. (1998). The delay/doppler radar altimeter. *IEEE Transactions on Geoscience and Remote Sensing*, *36*(5), 1578–1588.

Rashid, M., & Wahl, T. (2020). Predictability of Extreme Sea Level Variations Along the U.S. Coastline. *Journal of Geophysical Research: Oceans*, *125*(9). https://doi.org/doi:10.1029/2020jc016295

Ray, C., Martin-Puig, C., Clarizia, M. P., Ruffini, G., Dinardo, S., Gommenginger, C., & Benveniste, J. (2015). SAR altimeter backscattered waveform model. *IEEE Transactions on Geoscience and Remote Sensing*, *53*(2), 911–919.

Ray, R. D. (2016). On measurements of the tide at Churchill, Hudson Bay. *Atmosphere-Ocean*, *54*(2), 108–116.

Ray, R. D., & Egbert, G. D. (2004). The global S_1 tide. *Journal of Physical Oceanography*, *34*(8), 1922–1935.

Resio, D. T., & Westering, J. J. (2008). Modeling the physics of storm surges. *Physics Today*, *61*, 33–38.

Roeland, H., & Piet, R. (1995). Dynamic preservation of the coastline in the Netherlands. *Journal of Coastal Conservation*, *1*(1), 17–28.

Roser, M., Ritchie, H., Ortiz-Ospina, E., & Rodés-Guirao, L. (2013). World population growth. *Our world in data*.

Scharroo, R., Smith, W. H., & Lillibridge, J. L. (2005). Satellite altimetry and the intensification of hurricane katrina.

Schindelegger, M., Green, J. A. M., Wilmes, S. B., & Haigh, I. D. (2018). Can we model the effect of observed sea level rise on tides? *Journal of Geophysical Research: Oceans*, *123*, 4593–4609.

Schindelegger, M., Einšpigel, D., Salstein, D., & Böhm, J. (2016). The Global S_1 Tide in Earth's Nutation. *Surveys in Geophysics*, *37*(3), 643–680.

Schindelegger, M., Kotzian, D. P., Ray, R. D., Green, J. M., & Stolzenberger, S. (2022). Interannual changes in tidal conversion modulate M2 amplitudes in the Gulf of Maine. *Geophysical Research Letters*, *49*(24), e2022GL101671.

St-Laurent, P., Saucier, F. J., & Dumais, J. F. (2008). On the modification of tides in a seasonally ice-covered sea. *Journal of Geophysical Research: Oceans*, *113*(11), 1–11. https://doi.org/https://doi.org/10.1029/2007JC004614

van't Klooster, K. (2011). Ers-1, european remote-sensing satellite was launched 20 years ago. *2011 21st International Crimean Conference" Microwave & Telecommunication Technology"*, 117–118.

Vermaire, J. C., Pisaric, M. F., Thienpont, J. R., Courtney Mustaphi, C. J., Kokelj, S. V., & Smol, J. P. (2013). Arctic climate warming and sea ice declines lead to increased storm surge activity. *Geophysical Research Letters*, *40*(7), 1386–1390.

Wahl, T., Haigh, I. D., Nicholls, R. J., Arns, A., Dangendorf, S., Hinkel, J., & Slangen, A. B. (2017). Understanding extreme sea levels for broad-scale coastal impact and adaptation analysis. *Nature Communications*, *8*(1), 1–12.

Wahl, T., & Chambers, D. P. (2015). Evidence for multidecadal variability in us extreme sea level records. *Journal of Geophysical Research: Oceans*, *120*(3), 1527–1544.

Wahl, T., & Chambers, D. P. (2016). Climate controls multidecadal variability in us extreme sea level records. *Journal of Geophysical Research: Oceans*, *121*(2), 1274–1290.

Wahr, J. (1995). Earth Tides. *Global Earth physics, AGU Reference Shelf*, *1*, 40–45.

Woodworth, P. L. (2010). A survey of recent changes in the main components of the ocean tide. *Continental shelf research*, *30*(15), 1680–1691.

Woodworth, P. L., Gehrels, W. R., & Nerem, R. S. (2011). Nineteenth and twentieth century changes in sea level. *Oceanography*, *24*(2), 80–93.

Zakharova, E. A., Fleury, S., Guerreiro, K., Willmes, S., Rémy, F., Kouraev, A. V., & Heinemans, G. (2015). Sea ice leads detection using SARAL/AltiKa altimeter. *Marine Geodesy*, *38*, 522–533. https://doi.org/doi:10.1029/2001JC001026

Zhang, K., Douglas, B. C., & Leatherman, S. P. (2000). Twentieth-century storm activity along the us east coast. *Journal of Climate*, *13*(10), 1748–1761.

Zhao, J., Zhong, W., Diao, Y., & Cao, Y. (2019). The rapidly changing arctic and its impact on global climate. *Journal of Ocean University of China*, *18*(3), 537–541.

2

SECULAR TRENDS IN GLOBAL TIDES

Previous studies have demonstrated that tides are subject to considerable changes on secular time scales. However, these studies rely on sea level observations from tide gauges that are predominantly located in coastal and shelf regions and therefore, the large-scale patterns remain uncertain. Now, for the first time, satellite radar altimetry (TOPEX/Poseidon & Jason series) has been used to study worldwide linear trends in tidal harmonic constants of four major tides (M_2, S_2, O_1, and K_1). This study demonstrates both the potential and challenges of using satellite data for the quantification of such long-term changes. Two alternative methods were implemented. In the first method, tidal harmonic constants were estimated for consecutive four-year periods, from which the linear change was then estimated. In the second method, the estimation of linear trends in the tidal constants of the four tides was integrated in the harmonic analysis. First, both methods were assessed by application to tide gauge data that were sub-sampled to the sampling scheme of the satellites. Thereafter the methods were applied to the real satellite data. Results show both statistically significant decreases and increases in amplitude up to 1 mm/year and significant phase changes up to $\sim 0.1\,^{\circ}$*/year. The level of agreement between altimeter-derived trends and estimates from tide gauge data differs per region and per tide.*

This chapter has been published as Bij de Vaate, I., Slobbe, D. C., & Verlaan, M. (2022). Secular trends in global tides derived from satellite radar altimetry. *Journal of Geophysical Research: Oceans*, *127*(10), e2022JC018845. https://doi.org/https://doi.org/10.1029/2022JC018845

2.1. INTRODUCTION

Knowledge of tides is important for many practical (e.g., marine navigation, fishery, coastal engineering) and scientific purposes. Although tide predictions often treat tidal harmonic constants as stationary over time, considerable changes in tides have been observed on seasonal (e.g, Bij de Vaate et al., 2021; Müller et al., 2014) to long-term timescales (e.g., Müller et al., 2011; Ray, 2016). On the one hand, modifications of the tides can be the result of local processes, such as changes in coastal morphology or altered river flow (Haigh et al., 2020). On the other hand, observed variations in tides have been linked to regional climatic conditions, e.g., the extent of sea ice coverage (e.g., Bij de Vaate et al., 2021; Müller et al., 2014; St-Laurent et al., 2008), ocean stratification (e.g., Müller, 2012; Müller et al., 2014), and sea level rise (e.g., Devlin et al., 2017; Ross et al., 2017). Modelling studies suggest that climate change will continue to affect tides for centuries (Pickering et al., 2017; Schindelegger et al., 2018). Nevertheless, Haigh et al. (2020) indicated the need for better understanding of individual contributions of small-scale and large-scale processes.

An increasing number of studies are devoted to mapping and understanding secular changes in the tides. However, most of these studies rely on sea level observations from tide gauges that are mainly restricted to coastal and shelf regions. Hence, observed changes in tides could be dominated by local processes and the large-scale patterns remain unclear. Obtaining the global picture of long-term changes in tides would contribute to a better understanding of the drivers behind secular changes in tides. Understanding secular changes in tides may result in better identification and prediction of any consequences for coastal environments such as flooding (Li et al., 2021), salt intrusion (Hinton, 2000), or altered estuarine dynamics (Khojasteh et al., 2021).

To gain more insight in the large-scale secular changes in tides, we supplemented the clustered and sparsely distributed tide gauge dataset with data from satellite radar altimeters. Altimeter-derived water levels are being widely used to estimate tidal constants, and have recently been used to study seasonal changes in tides (Bij de Vaate et al., 2021; Müller et al., 2014). However, up to now, only Ray (2016) used altimeter data from successive missions to compare the amplitude of the main semi-diurnal tide (M_2) near Churchill, Hudson Bay (Canada). Given the length of the current satellite altimeter records (>25 years), from a theoretical point of view it should be possible to obtain estimates of the secular changes in tides from these data. For that reason, we have exploited the provided opportunities and used data from TOPEX/Poseidon and the Jason satellites to obtain a global estimate of the linear secular trends in the major tides. In this paper we first describe the data, including satellite radar altimetry, high-frequency tide gauge records and reanalysis data used for validation of the results. Then an outline is given of two approaches to study secular changes in tides and an experiment to test these methods. Finally, the results are introduced and compared to observations at tide gauges and documented changes in tides.

2.2. Data

2.2.1. Satellite Radar Altimetry

Data from the TOPEX/Poseidon and Jason satellite altimeters were combined (further referred to as TPJ) resulting in 28 years of sea level data (1993-2020). Data from interleaved orbits were not considered. The TPJ satellites have a ground coverage up to 66° N/S and an along-track resolution of about 5.8 km. Altimeter data were obtained through the Radar Altimeter Data System (RADS, Naeije, 2022). The following geophysical and range corrections were applied (Scharroo et al., 2016): ionosphere (NIC09 for TOPEX/Poseidon, GIM for Jason), dry troposphere (ECMWF), wet troposphere (if available: radiometer, otherwise: ECMWF), solid tide (Cartwright and Edden, 1973; Cartwright and Taylor, 1971), pole tide (Wahr, 1985), load tide (FES2014), mean sea surface (DTU18-MSS), sea state bias (CLS), and dynamic atmosphere (DAC) (MOG2d (ERA Interim forcing)). The center-of-gravity (CG) correction that RADS by default applies to TOPEX/Poseidon ranges was removed to reduce intermission biases in the solar S_2 tide (Beckley et al., 2021; Zawadzki et al., 2018). In addition, to minimize aliasing of non-tidal sea level variability on tidal frequencies, an additional correction was applied. Following Ray and Zaron (2016), the multi-mission, gridded sea level anomalies (SLA) from the Data Unification and Altimeter Combination System (DUACS) (Taburet et al., 2019) were subtracted from the TPJ-water levels. This removes seasonal and interannual variability from the obtained water levels and specifically reduces the noise in regions with high mesoscale activity. In the remainder of the paper, this correction will be referred to as the 'mesoscale correction'. Finally, outliers in the time series were detected and removed based on threefold the median absolute deviation.

In this paper, results are presented on global maps, supplemented by a zoom in on the North West European Shelf. For the global analysis, data are treated as follows. First, the locations where two tracks intersect (crossovers) were identified. For all of those locations, the data of the two crossing tracks within a radius of 30 km were assigned to the location of the respective crossover. 30 km equals half the distance between the closest neighbouring crossovers. Note that the along-track distance between crossovers depends on latitude: from $\sim$ 460 km at the equator to $\sim$ 60 km at 66° N/S. By stacking the data at crossover locations, the temporal resolution is increased and tidal analysis is deemed more reliable. For the zoom in on the North West European Shelf, data were processed on a track-by-track basis. Data from different cycles were collocated following Cherniawsky et al. (2010). The along-track analysis allows for a higher spatial resolution and to get closer to the tide gauge locations, at the price of an increase in uncertainty levels.

2.2.2. Tide Gauges

Alongside the altimeter data, data from a selection of tide gauges were processed to allow for a comparison of the derived trends. For this purpose, only tide gauge data from the TPJ-period were considered (1993-2020). Data from the GESLA-3 dataset (Haigh et al., 2021) were complemented with quality controlled water level records from tide gauges on the North West European Shelf, provided by nine European organizations (see Acknowledgments). The latter comprise data from 1997 onward, and are manually inspected to exclude possible outliers. Records that span less than 19 years were excluded. The temporal resolution of the tide gauge data varies from one minute to one hour, mainly depending on the country where the stations are located and the time of data acquisition. Tide gauge records were corrected for atmospheric loading using the same product as was used for altimetry (DAC).

2.2.3. REANALYSIS DATA

Finally, reanalysis data were used to obtain uncertainty estimates of the estimated linear change in tidal constants. For this, the Global Tides and Surge Model (GTSM, Wang et al., 2021) was used, forced by ERA5 reanalysis data. GTSM is a barotropic (2D) model that makes use of an unstructured grid with a resolution that increases from 25 km at open ocean to 2.5 km at the coast. Time series with a sampling rate of 10 minutes were reconstructed for the full TPJ-period. This was done for over 600 locations covering the global oceans and about 300 locations on the western North West European Shelf. Subsequently, the time series were corrected for atmospheric loading (using the DAC), temporarily detided and then subjected to a high-pass filter to remove any non-tidal signal with periods larger than 2 days. This was done to mimic the 'mesoscale correction' that was applied to the TPJ-data. Although the GTSM does not resolve ocean circulation and associated mesoscale sea level variability, atmospheric forcing may induce seasonal/interannual sea level variability (e.g., Dangendorf et al., 2014), which is to some extent also contained in the 'mesoscale correction'.

2.3. METHODS

Earlier studies on secular changes in tides typically relied on year-by-year harmonic analyses of high-frequency data, followed by the fitting of a linear trend through the yearly tidal harmonic constants (e.g., Müller et al., 2011; Ray, 2009; Zaron and Jay, 2014). In this paper, a similar procedure was adopted to process the tide gauge data. However, for satellite data, such a procedure is not possible due to the relatively low sampling rate and consequent aliasing of high-frequency tidal signals onto lower frequencies. That is, for the major tides, the TPJ-sampling interval of 9.9156 days results in alias periods of 62.1 (M_2), 58.7 (S_2), 173.2 (K_1), and 45.7 days (O_1) (Cherniawsky et al., 2010; Schrama and Ray, 1994). By applying the Rayleigh criterion to these alias frequencies, we can find the minimum record length that is required to separate the tides of interest from other signals (Savcenko and Bosch, 2007). For M_2, S_2, and O_1, records of three (2.97) years are sufficient to separate them from other considered constituents, while at least 9.19 years are required to separate K_1 from S_{sa} (semi-annual tide). Hence, a year-by-year harmonic analysis of TPJ-data is not possible. In this paper, two different methods were implemented.

Both approaches make use of UTide (Codiga, 2020). This software executes a harmonic analysis for a given set of frequencies similar as in TTide (Pawlowicz et al., 2002), yet it is able to deal with irregular temporal sampling. The latter is a requirement for processing stacked altimeter-derived water levels. For the analysis of tide gauge data, a large set of constituents (including shallow water constituents) was considered following from the automated constituent selection method in UTide (Codiga, 2020; Foreman, 2004). For satellite data, a fixed set of constituents was considered, as explained below.

2.3.1. SEGMENTED HARMONIC ANALYSIS (SEGHA)

The first approach, referred to as the ‘segmented harmonic analysis’ (SegHA) approach (inspired by Jin et al. (2018)), is a two-step procedure that is very similar to the conventional analysis of secular changes using tide gauge data. This approach could be carried out with standard tidal analysis tools, but comes at the price of a slight simplification in error propagation.

STEP 1: ESTIMATION OF TIDAL HARMONIC CONSTANTS

Instead of processing the data year-by-year, time series were split in seven consecutive periods of four years. Thereafter, tidal harmonic constants were calculated and referred to the center date of the respective four-year period. The time span of four years was chosen primarily because this allows the separation of M_2, S_2, and O_1 from other signals (this requires at least 3 years). On the other hand, there is in some instances (mainly coastal) a discrepancy between the actual nodal modulation of lunar tides (18.6 year cycle) and the theoretical value (Hagen et al., 2021). Hence, although amplitude/phase estimates are corrected for the nodal modulation during tidal analysis, there may be a residual modulation left. To separate the trend in tidal amplitude from this possible residual nodal modulation, the difference between the respective center data of the first and last period was required to be at least 18.6 years. This can be achieved by processing segments of up to five years (segments are not allowed to overlap). Hence it is anyway not possible to study the secular trend in K_1 harmonic constants from the available data using the SegHA approach. Given the minimum of three years and the maximum of five years, a time span of four years was chosen since this allows making full use of the available data (28 years).

For each four-year period, tidal amplitudes and phases were estimated for 20 tidal constituents, including: three long-period tides (S_a, M_m, and M_{sf}), five diurnal tides (Q_1, O_1, P_1, S_1, and K_1), eight semi-diurnal tides ($2N_2$, μ_2, N_2, ν_2, M_2, L_2, T_2, S_2, and $2SM_2$), and four shorter period tides (M_3, MN_4, M_4, and MS_4). This selection of constituents eliminates possible conflicts between constituents pairs that cannot be separated from four years of data (e.g., K_1 and S_{sa}). In addition, from each four-year period the mean sea level (Z_0) and a possible trend in mean sea level were estimated to account for any remaining interannual sea level variability.

95% confidence intervals for the estimated harmonic constants were computed with UTide. This measure is derived from linearized error propagation of the total residual power (using the detided signal) within the frequency band surrounding the frequency in question (M_2/S_2 ±0.2 cycles/day and O_1 ±0.1 cycles/day), obtained using the Lomb-Scargle periodogram (Codiga, 2011; Pawlowicz et al., 2002). However, it is stated by Codiga (2011) that certain assumptions underlying this procedure are strictly valid only for uniformly sampled data. The resulting confidence intervals “should be considered potentially reasonable and approximate first estimates, but should be compared against the results for uniform times whenever possible, and used with a measure of caution." (Codiga, 2011, p. 21). Indeed, it was found that both the frequency and timing of the sea level measurements by the satellite influenced the accuracy of the resulting tidal estimates (Guarneri et al., 2022). Moreover, UTide averages the spectral density distribution of the residuals over nine frequency bands resulting in similar confidence intervals for all diurnal tides, all semi-diurnal tides, and so on. In line with the advice from Codiga (2011), but due to the lack of tide gauge data in the vicinity of the altimeter points, we have therefore obtained an additional (alternative) uncertainty estimate using the reanalysis data that were introduced in Section 2.2.3. These time series were reduced to a four-year period (2015-2018) and interpolated to the TPJ-sampling interval of which the start time was iteratively shifted by about 4.75 hours (TPJ-sampling period divided by 50), resulting in 50 time-shifted time series. The median absolute deviation (MAD) of the tidal harmonic constants estimated from these 50 time series, was scaled by 1.4826 to obtain the standard error of the estimate (Rousseeuw and Croux, 1993). The final values are location- and tide specific, but assumed to be independent of the four-year period.

Step 2: Linear Trend Estimation

The linear secular trends in harmonic constants were estimated by fitting the following equations through the series of seven values, using weighted least squares. Here the error propagation was simplified by ignoring the correlations between amplitudes and phases estimates. For amplitudes follows:

$$\tilde{A}_k(t_i) = \underbrace{a^A_{N_k}\cos(2\pi\frac{t_i - t_c}{18.6} + N_c)}_{\text{residual nodal modulation}} + \underbrace{b^A_k(t_i - t_c)}_{\text{trend}}. \tag{2.1}$$

where, $\tilde{A}_k(t_i)$ is the residual amplitude for the ith four-year period of the tidal constituent in question (k) (obtained by subtracting the time averaged amplitude), b^A_k the linear change in amplitude, t_i the center time of the ith four-year periods, and t_c the center time of the full TPJ-period. In addition, the nodal modulation was included in the problem formulation (see Section 2.3.1). N_c represents the nodal phase at the center date. Both the magnitude of the residual nodal modulation ($a^A_{N_k}$), and the linear amplitude change (b^A_k) were estimated, resulting in a redundancy of five. For phases the following equation was used:

$$\tilde{\phi}_k(t_i) = \underbrace{a^{\phi}_{N_k}\cos(2\pi\frac{t_i - t_c}{18.6} + N_c)}_{\text{residual nodal modulation}} + \underbrace{b^{\phi}_k(t_i - t_c)}_{\text{trend}}. \tag{2.2}$$

Where $\tilde{\phi}_k(t_i)$ is the residual phase for the ith four-year period, $a^{\phi}_{N_k}$ the magnitude of the residual nodal phase modulation, and b^{ϕ}_k the linear coefficient describing the change in phase.

Both the standard errors of the harmonic constants derived from UTide and from GTSM (Section 2.3.1) were used to assess the significance of the fitted trends. For the S_2

tide, the choice of ionospheric correction applied to the data may affect the estimated tidal harmonic constants (Jee et al., 2010; Ray, 2020; Zawadzki et al., 2018). Therefore, an additional error estimate was obtained (see Appendix A) and added to the estimates obtained by UTide and GTSM respectively. Given the standard errors of the tidal harmonic constants, the standard error of the trend was derived through error propagation. Finally, confidence intervals were obtained by multiplying the standard error with the appropriate z-score.

2.3.2. TREND-INTEGRATED HARMONIC ANALYSIS (TINTHA)

In the second approach, the linear trends in the four tides of interest (M_2, S_2, O_1, and K_1) were estimated jointly with the average tidal harmonic constants. This required an extension of the available tidal analysis software but allowed for a full error propagation (that is, including the co-variances between amplitude and phase estimates). Since we are now using the full 28 years of data, this approach allows the analysis of changes in the K_1 tide. Moreover, the set of constituents included in the analysis was extended by S_{SA}, K_2 and T_2. In the SegHA approach, these had to be excluded due to aliasing issues.

The TintHA approach uses a different formulation of the tides. Within UTide, the complex formulation is used in which the tidal water level for constituent k, i.e. $\hat{h}_k(t)$, is written as the product of three terms:

$$\hat{h}_k(t) = \left(A_k e^{i\phi_k}\right)\left(f_k(t) e^{iu_k(t)}\right) e^{iv_k(t)}, \tag{2.3}$$

where the term $e^{iv_k(t)}$ is the phase of the equilibrium tide, $\left(f_k(t)e^{iu_k(t)}\right)$ is the nodal correction, and the term $\left(A_k e^{i\phi_k}\right)$ is the complex amplitude-phase pair that needs to be estimated. To keep the equations linear, we consider the complex amplitude-phase pair $\hat{A}_k = A_k e^{i\phi_k}$:

$$\hat{A}_k(t) = \hat{F}_k + \hat{G}_k \frac{t - t_0}{T}, \tag{2.4}$$

where the time period considered starts at t_0 and ends at $t_0 + T$, so that $\hat{A}_k(t_0) = \hat{F}_k$ and $\hat{A}_k(t_0 + T) = \hat{F}_k + \hat{G}_k$. The relative change over this time period is $\hat{\Delta} = (\hat{F}_k + \hat{G}_k)/\hat{F}_k$. The angle and absolute value of this complex number give the phase change and relative amplitude change. A disadvantage of this linear model is that the rate of change of the amplitude and phase is not constant over the time interval. For small changes, however, the approximation error will be small. Note that in this method no empirically estimated correction for any residual of the nodal modulation is determined as this, in combination with the trend estimation, would result in a non-linear estimation problem.

Similar to the first approach (SegHA), alternative error estimates were obtained by means of the GTSM reanalysis data. For the latter, the full 28 year time series were interpolated to TPJ-sampling intervals while iteratively shifting the start time 50 times. From these time series the linear change in tidal harmonic constants was computed and the MAD of these values was again scaled by 1.4826 to obtain the standard error of the trend estimates. For the S_2 tide, the error estimates were supplemented by the possibly error due to the ionospheric correction (as described in Appendix A). Finally, the confidence intervals were obtained by multiplying the error estimate by the appropriate z-score and interpolating the GTSM-derived product to the TPJ-tracks.

2.3.3. VALIDATION

COMPARISON OF SEGHA AND TINTHA USING TIDE GAUGE DATA

Both methods (SegHA and TintHA) were assessed by application to tide gauge data that were sub-sampled to TPJ-sampling intervals. This was done in a similar manner as the GTSM reanalysis data were used to compute confidence intervals. Data were subsampled both to an along-track sampling of 9.9156 days and a crossover sampling which was determined based on the TPJ crossover sampling at the latitude of the respective tide gauge. The start time of the subsampled time series was iteratively shifted, resulting in 50 time series for each tide gauge. For the assessment, only tide gauges were considered that have full data coverage during the entire TPJ-period. In addition to DAC, the ‘mesoscale correction’ was applied to the data to resemble the processing of altimeter data. As this altimetry-derived product is not available everywhere across the globe, only the data from 109 tide gauges could be used. The secular change in tidal harmonic constants derived from both methods was compared to the ‘true’ change that was obtained by processing the original high-frequency data on a year-by-year basis. Assessment of the different methods was done by comparing the median absolute error (MedAE) for respectively each tide gauge, tidal constituent and sampling scheme.

COMPARISON OF CONFIDENCE INTERVALS USING TIDE GAUGE DATA

As discussed in Section 2.3.1, two alternative confidence intervals were obtained for the trend estimates: one following from UTide, the other from processing of GTSM reanalysis data. To validate both alternatives, an additional experiment was performed using the results from the experiment with tide gauge data. This time, the scaled MAD of the trend estimates from the TPJ-sampled tide gauge data was used to compute 95% confidence intervals. These were then compared to the 95% confidence intervals obtained from UTide (based on the TPJ-sampled tide gauge data) and those derived using GTSM reanalysis data from the exact location as the tide gauges. This was done for both the regular along-track, as well as the latitude dependent crossover sampling. Results were analysed based on the correlation, the root-mean-square error (RMSE), and the median underestimation of respectively the GTSM/UTide product.

VALIDATION OF ESTIMATED SECULAR CHANGE USING TIDE GAUGE DATA

In order to validate the secular changes derived from TPJ-data by means of another product, ideally the data needs to be from the exact same location. Since these data is not available, we need to consider the impact of the spatial separation on the consistency of estimated secular change. As can be seen in Figure 2.1a, the distance between the tide gauges and the nearest TPJ-crossover varies from approximately 50 km to over 400 km. In addition, Figure 2.1b shows that the spatial error (MedAE) in estimated amplitude change (comparing every individual tide gauge to other tide gauges within a certain radius) increases with distance. Based on this figure, it was decided to only use tide gauges that are closer than 75 km to one or more TPJ-crossovers (indicated by the thicker outlines in Figures 2.7 and 2.8) for the assessment of agreement between tide gauges and

TPJ-data. From the 176 remaining tide gauge-crossover combinations, the absolute differences in estimated trend were computed. The differences were then classified as being insignificant for combinations where the confidence interval of the trend estimate at the crossover (as derived from GTSM, see Figure 2.6b) exceeded the difference.

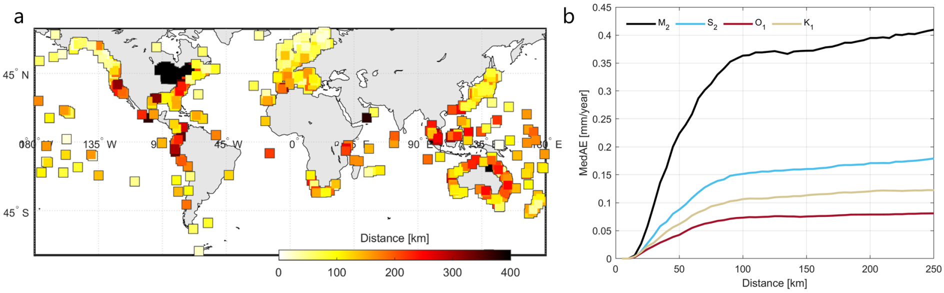


Figure 2.1: Distance between each tide gauge and the nearest TPJ-crossover (a). Spatial error (MedAE) in estimated secular trends in tidal amplitude as a function of distance, derived from the tide gauge data (b).

2.3.4. POST-PROCESSING

Estimated trends were omitted for locations where at least one of the following criteria was not met. If not mentioned otherwise, these criteria were applied in the analysis of the crossovers, individual tracks, and tide gauges:

- The root-mean-square (RMS) of the residual signal should be below 0.15 m. Globally, this removes ~ 8% of the data.
- There should be consistent data coverage throughout the year. A location was not considered when there were more than ten sequential day numbers without data. Globally, this removes ~ 20% of the data.
- The estimated linear coefficient should be larger than its confidence interval. Which confidence intervals were used is mentioned in figure captions.
- Only applied in along-track analysis: crossovers where there is no overlap between the estimated linear trends of the two crossing tracks (interpolated to the location of the crossover) ± the local confidence interval, were flagged. In such a case, all derived trends of the two crossing tracks within half the distance between neighbouring crossovers were omitted.

2.4. RESULTS

2.4.1. VALIDATION

COMPARISON OF SEGHA AND TINTHA USING TIDE GAUGE DATA

Comparison of both methods applied to tide gauge data shows little difference between the SegHA and TintHA methods (Figure 2.2). Regardless of the method that was used,

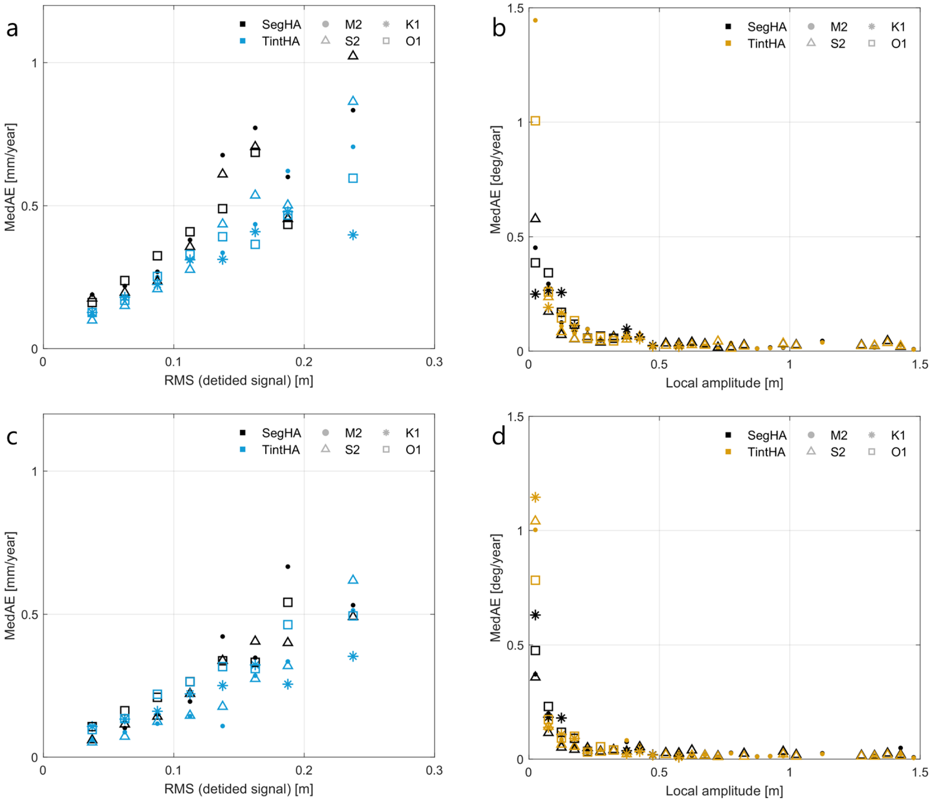


Figure 2.2: Median absolute error (MedAE) between the ‘true’ linear change in amplitudes (a, c) and phases (b, d) derived from high-frequency tide gauge records and the product derived from the SegHA and TintHA approaches applied to the data sub-sampled at TPJ along-track sampling intervals (a, b) and TPJ crossover sampling (c, d). Colours indicate which method was used and the marker style depicts the different tidal constituents that were studied. For visualisation purposes the errors are averaged for intervals of 0.025 m and 0.05 m for the RMS (detided signal) and local amplitude respectively.

sub-sampling the data to TPJ-sampling interval reduces the accuracy of the derived changes in tidal amplitude and phase. In the case of amplitude, the observed error between the ‘true’ and derived change increases with larger non-tidal water level variation (higher RMS; Figure 2.2a). On the other hand, the accuracy of the derived phase change predominantly depends on the local amplitude of the tide in question (Figure 2.2b). In particular for amplitudes below ~ 15 cm, the derived phase change appears unreliable. Only in terms of amplitude change, the TintHA method performs more consistent than the SegHA method, with an average MedAE of 0.24 mm/year compared to 0.29 mm/year and fewer outliers. Overall, the crossover sampling improves the accuracy of both methods for both amplitudes (MedAE reduces from 0.25 mm/year to 0.18 mm/year) and phases (0.18 °/year to 0.13 °/year) for all tides (Figure 2.2c, d).

COMPARISON OF CONFIDENCE INTERVALS USING TIDE GAUGE DATA

The UTide-, GTSM- and tide gauge derived confidence intervals for the change in M_2 amplitude (a) and phase (b) are shown in Figure 2.3, following from the TintHA approach. From the figure, it appears that both the UTide and the GTSM-product subceed the confidence intervals derived from the tide gauges (i.e. more scatter points are located in the bottom right). However, statistical analysis of the results (all tides combined) show that the tide gauge derived confidence intervals correlate better with the GTSM product (correlation of 0.62 for amplitude, 0.67 for phase) than with the UTide product (amplitude: 0.53, phase: 0.55). In addition, the RMSE is lower for GTSM (0.37 mm/year and 0.05 °/year) than for UTide (0.46 mm/year and 0.08 °/year). In terms of underestimation of the confidence intervals for amplitude change, the GTSM product again performs better than UTide (0.15 mm/year versus 0.26 mm/year). Concerning the phase changes, the GTSM product performs slightly better (0.02 °/year versus 0.03 °/year). However, note that the phase results for O_1 and K_1 are based on only 11/12 tide gauges (compared to 61 and 53 for M_2 and S_2 respectively). Tide gauges where the tidal amplitude was below 15 cm were excluded, because a low tidal amplitude increases the uncertainty of phase change estimates to such an extent that any differences between the different confidence interval products become irrelevant (also shown by Figures 2.2b and 2.2d). However as these differences were typically large, they would dominate the statistics and overshadow the results that do matter.

Moreover, it was found that while the tide gauge and GTSM-derived confidence intervals are significantly lower for the crossover sampling than for the along-track sampling (respectively 25% and 27% for amplitude, 27% and 17% for phase), the UTide confidence intervals were less affected (4% for amplitude 6% for phase).

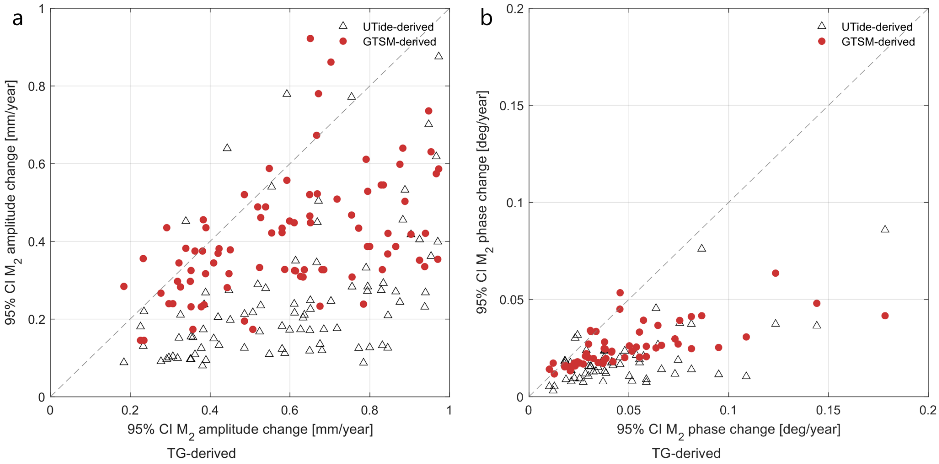


Figure 2.3: 95% confidence intervals derived from UTide, GTSM reanalysis data and the tide gauges as described in Section 2.3.3 for the change in M_2 amplitude (a) and phase (b), following from the TintHA approach and along-track sampling.

2.4.2. GLOBAL ANALYSIS

The estimated trends in amplitude at the TPJ-crossovers following from the TintHA approach, are displayed in Figures 2.4 and 2.5. The results produced by the SegHA method are very similar and incorporated shown by Figure C.5. Clearly, regions that are covered by sea ice during part of the year (above 55° N/S), have insufficient data availability for this analysis and are excluded. The distribution of locations where the estimated trend coefficients are significant, varies per tidal constituent, which is closely related to the confidence intervals (Figures 2.6, C.1 and C.2). In most cases, the GTSM-derived confidence intervals (e.g., Figure 2.6b) exceed the intervals derived by UTide (Figure 2.6a).

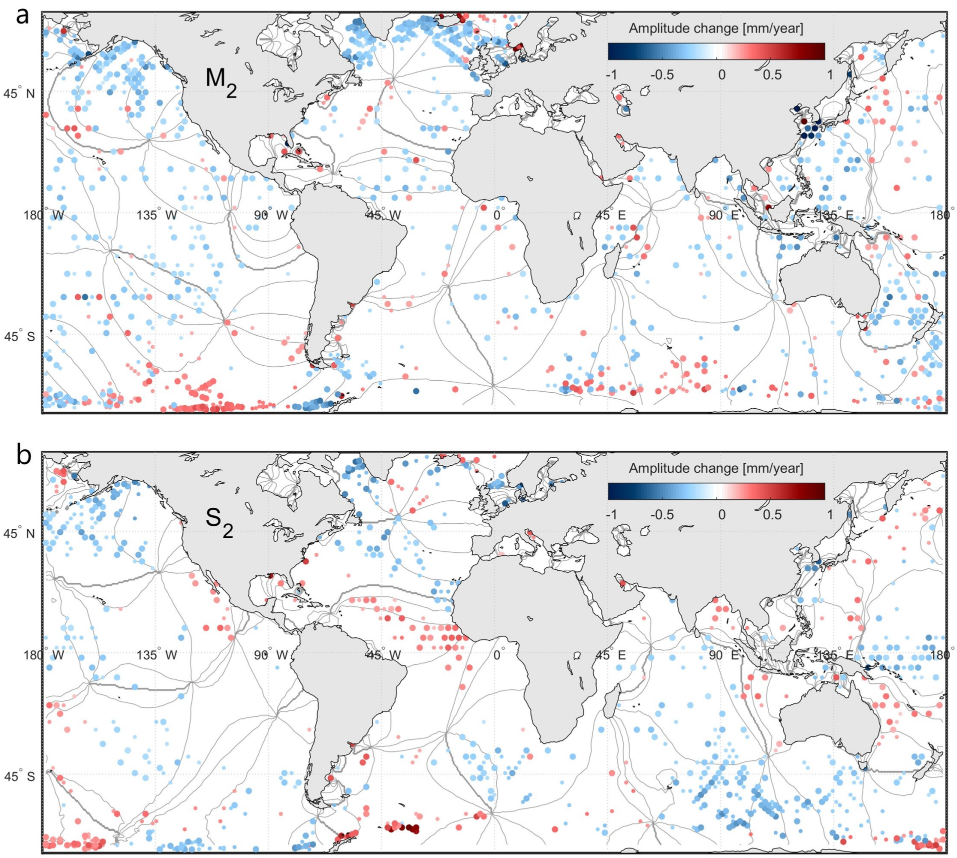


Figure 2.4: Linear change in M_2 (a) and S_2 amplitude (b) per year (1993-2020) following the TintHA approach. Locations where the post-processing criteria were not met are excluded from the figure. The smaller scatters indicate data that exceeds both the UTide and GTSM 90% confidence intervals, while the larger scatters indicate significant data at the 95% confidence level. Lines in the background depict tidal phases at 45° intervals.

As can be seen in Figure 2.4 and 2.5, all tides are subject to yearly changes of up to ±1 mm/year. The magnitude and sign of the yearly change vary largely across the globe,

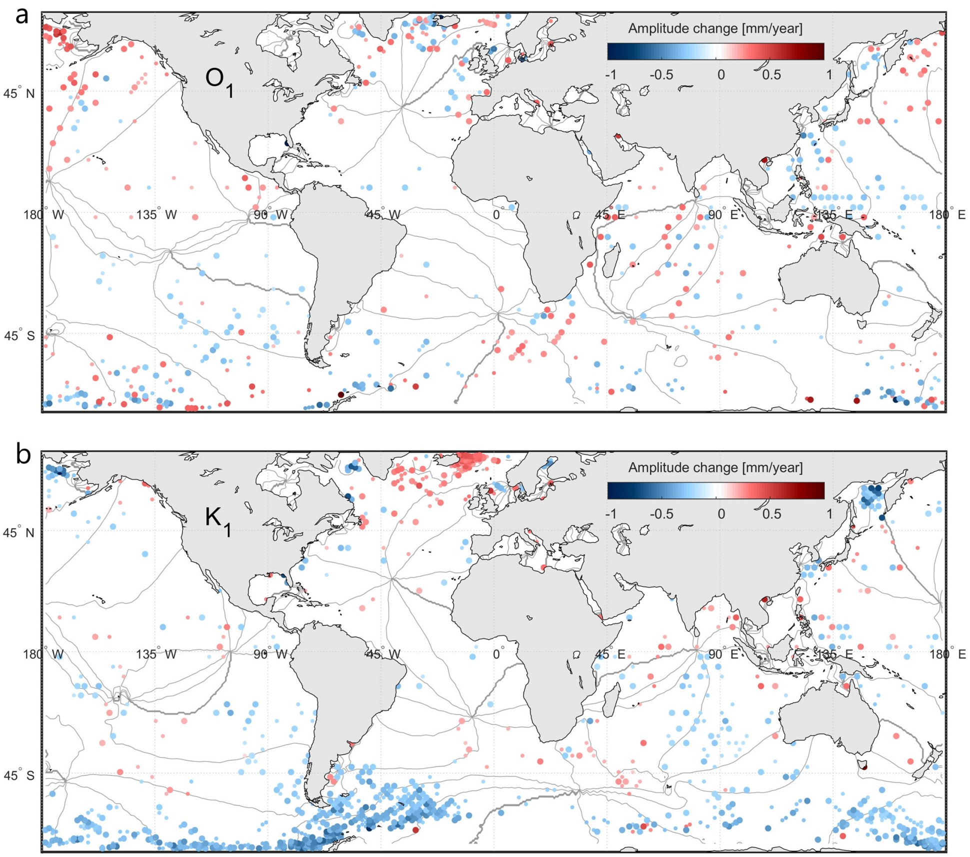


Figure 2.5: Linear change in O_1 (a) and K_1 amplitude (b) per year (1993-2020) following the TintHA approach. Locations where the post-processing criteria were not met are excluded from the figure. The smaller scatters indicate data that exceeds both the UTide and GTSM 90% confidence intervals, while the larger scatters indicate significant data at the 95% confidence level. Lines in the background depict tidal phases at 45° intervals.

while the spatial correlations of the signal vary per tide. For M_2, the change in amplitude is predominantly negative. The most obvious regions of positive change are in the south, near Antarctica and east of Iceland (Figure 2.4a). Although the overall change is rather heterogeneous, spatial correlation of the signal is stronger near the poles than at the lower latitudes. On the contrary, the change in S_2 amplitude shows more distinct regions of either positive or negative change across the globe (Figure 2.4b). Predominantly positive changes in amplitudes are observed around the equator and near the poles, while negative changes are more restricted to mid-latitudes. Differences in sign of the amplitude change appear closely related to the location of amphidromic points and co-phase lines. The change in O_1 amplitude is more similar to that of M_2, concerning the level of heterogeneity (Figure 2.5a). However, overall, the change in O_1 amplitude is a lot smaller than that of M_2 and only in a few locations, the confidence intervals are exceeded. For K_1, predominant negative changes are observed across the globe, except for the north Atlantic and the Indian Ocean (Figure 2.5b).

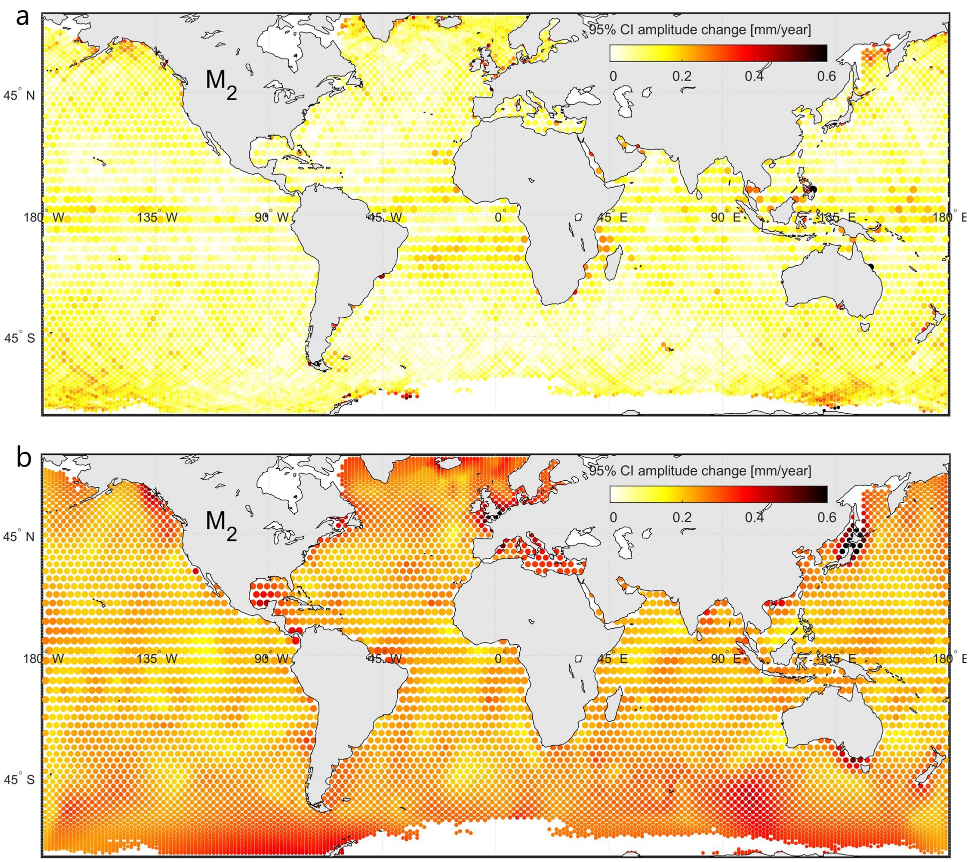


Figure 2.6: 95% confidence intervals for trend estimates derived from confidence intervals computed by UTide (a) and from standard errors derived from GTSM (b) for M_2 amplitude.

Trend estimates derived from the global tide gauge dataset are shown in Figures 2.7 and 2.8. For M_2, 41% of the differences in trend estimates from tide gauges and TPJ-crossovers were statistically insignificant considering the 95% confidence intervals derived from GTSM. For S_2 this value was 59%, for O_1 64% and for K_1 59%. However, note that in Figure 2.1b we observed a significant decrease in consistency among tide-gauge derived estimates with increased distances among the tide gauges. Given the fact that the distance between TPJ-crossovers and most tide gauges is at least 50 km, this explains part of the inconsistency between the altimeter- and tide gauge derived estimates. This in particular applies to M_2.

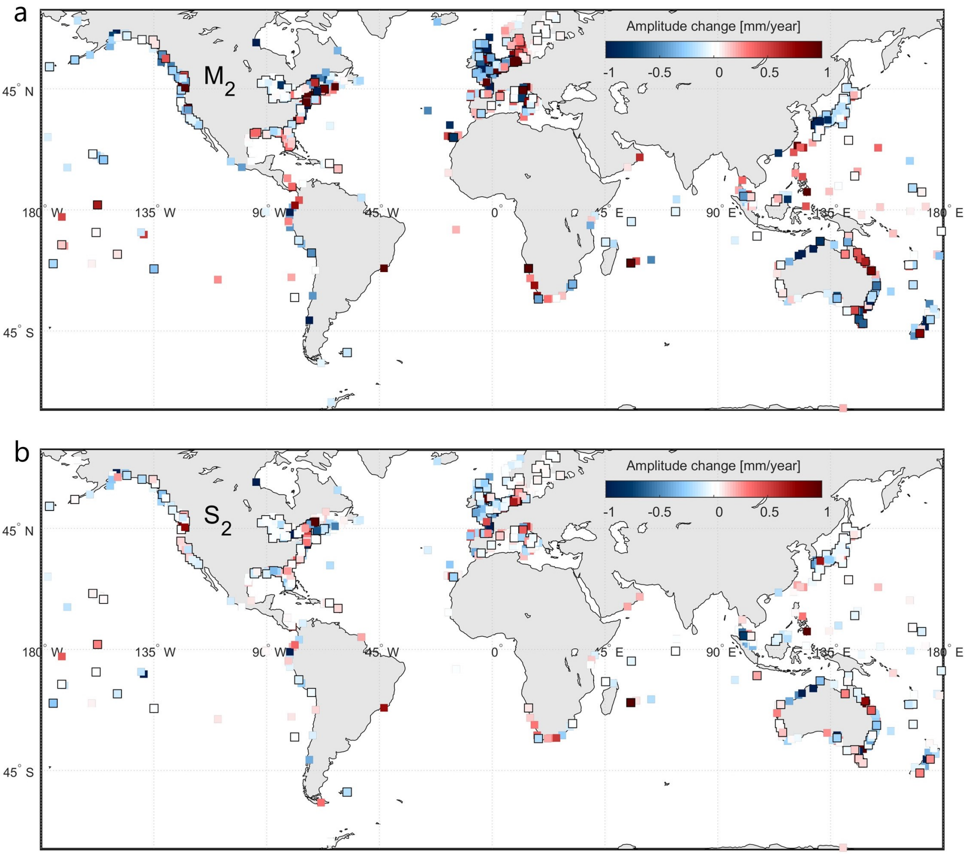


Figure 2.7: Secular trends in M$_2$ (a) and S$_2$ (b) amplitudes, derived from tide gauge records from the TPJ-period (1993-2020) (from GESLA-3; Haigh et al. (2021)). Black-outlined tide gauge locations are within 75 km of a TPJ-crossover and are used for the similarity measure as explained in Appendix B.

2.4.3. NORTH WEST EUROPEAN SHELF

A selection of results from the along-track analysis of the North West European Shelf region is displayed in Figure 2.9. Because of their relatively low amplitudes in the region (< 0.15 m), O$_1$ and K$_1$ are not included here.

The M$_2$ amplitude change derived from altimetry is predominantly negative across the domain, except for the central North Sea and the Skagerrak (Figure 2.9a). The largest change is observed towards the eastern coasts of the North Sea. Unfortunately most of the tide gauges are located along the coastline while RADS does not include coastal altimeter data. Nevertheless, the observed amplitude change at the tide gauges in the Netherlands, Germany, Denmark (and to a smaller extent the United Kingdom and Norway), is similar to that at nearby tracks. Limited similarity is observed for the tide gauges

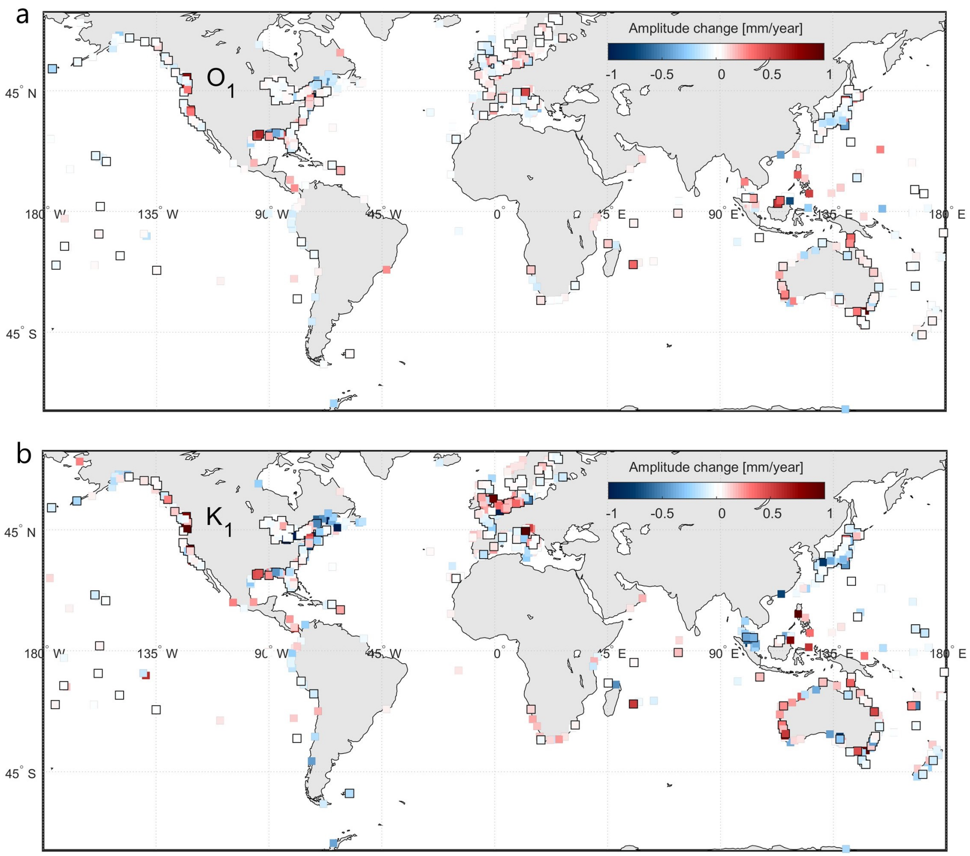


Figure 2.8: Secular trends in O_1 (a) and K_1 (b) amplitudes, derived from tide gauge records from the TPJ-period (1993-2020) (from GESLA-3; Haigh et al. (2021)). Black-outlined tide gauge locations are within 75 km of a TPJ-crossover and are used for the similarity measure as explained in Appendix B.

in the English Channel, the Irish Sea and on the west coast of Norway. The altimetry-derived change in M_2 phase is largest near the amphidromic points in the North Sea and in the northwest corner of the region (Figure 2.9c). Overall, both the sign of the phase change as derived from altimetry as well as from tide gauges, is highly variable within the domain. In addition, the availability of significant altimetry-derived phase changes near tide gauges is even more limited than was the case for the amplitude, making a comparison difficult.

The observed trends in S_2 amplitude are smaller than those in M_2 amplitude (Figure 2.9b), while the change in phase is larger (Figure 2.9d). These differences in magnitude are also observed at the tide gauges. However, tide-gauge and altimeter-derived estimates for S_2 agree in only few locations, predominantly along the Dutch coastline.

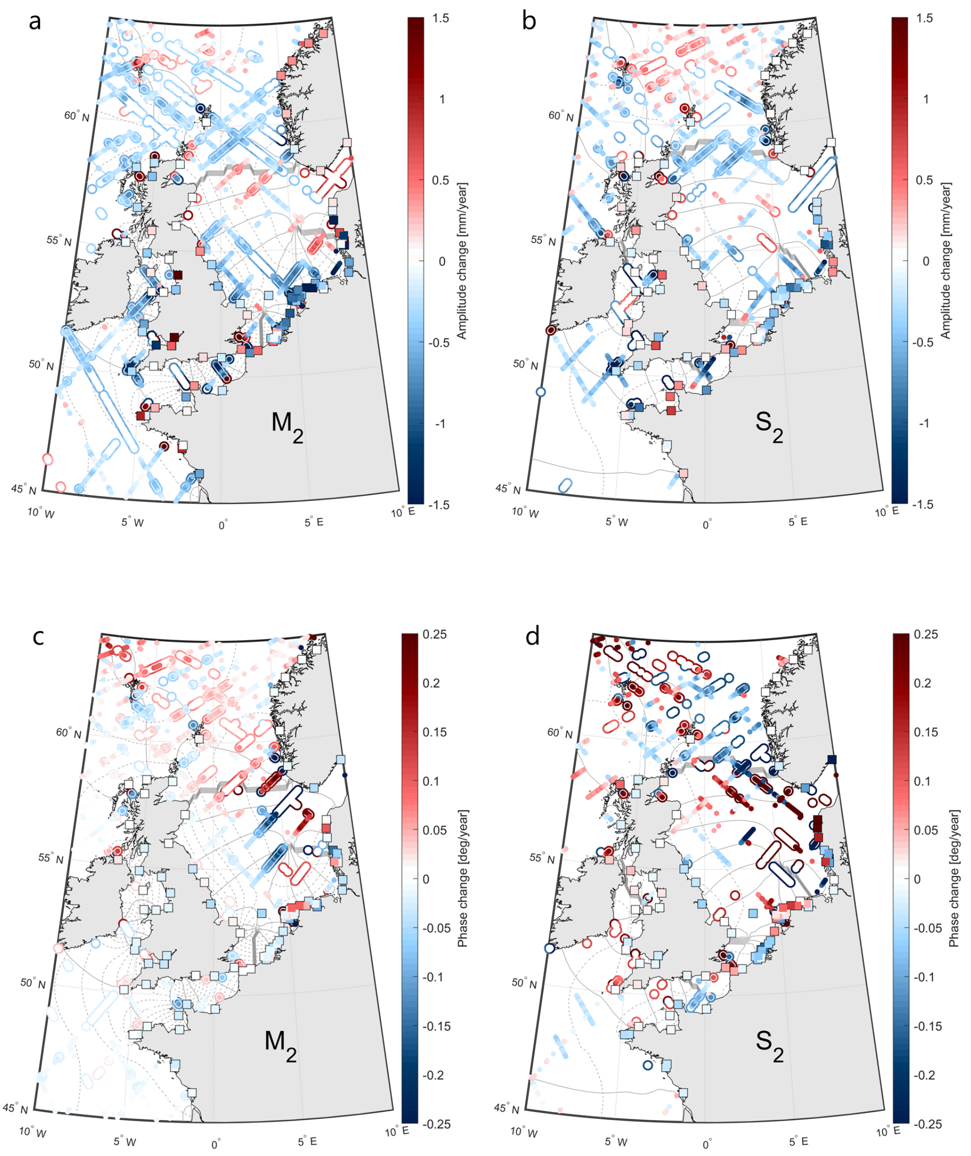


Figure 2.9: Linear change in M_2 amplitude (a), S_2 amplitude (b), M_2 phase (c) and S_2 phase (d) per year derived with the TintHA approach. The smaller solid scatters indicate significant trends given the UTide-derived 95% confidence intervals, the hollow outline indicates significance according to the GTSM-derived 95% confidence intervals (see Section 2.3). Co-tidal maps are shown in the background where the solid line indicates the phase at 45° intervals, the dashed lines show the amplitudes at 0.25 m intervals.

Both the GTSM and UTide-derived confidence intervals increase towards the coast for the amplitude change (Figure 2.10a, 2.10b, 2.11a and 2.11b) and towards amphidromic points for the phase change (Figure 2.10c, 2.10d, 2.11c and 2.11d). In all cases, the GTSM-derived confidence intervals exceed the ones computed by UTide. This is most noticeable for the S_2 amplitude change. Both GTSM- and UTide-derived confidence intervals for S_2 phase change are significantly larger than for M_2.

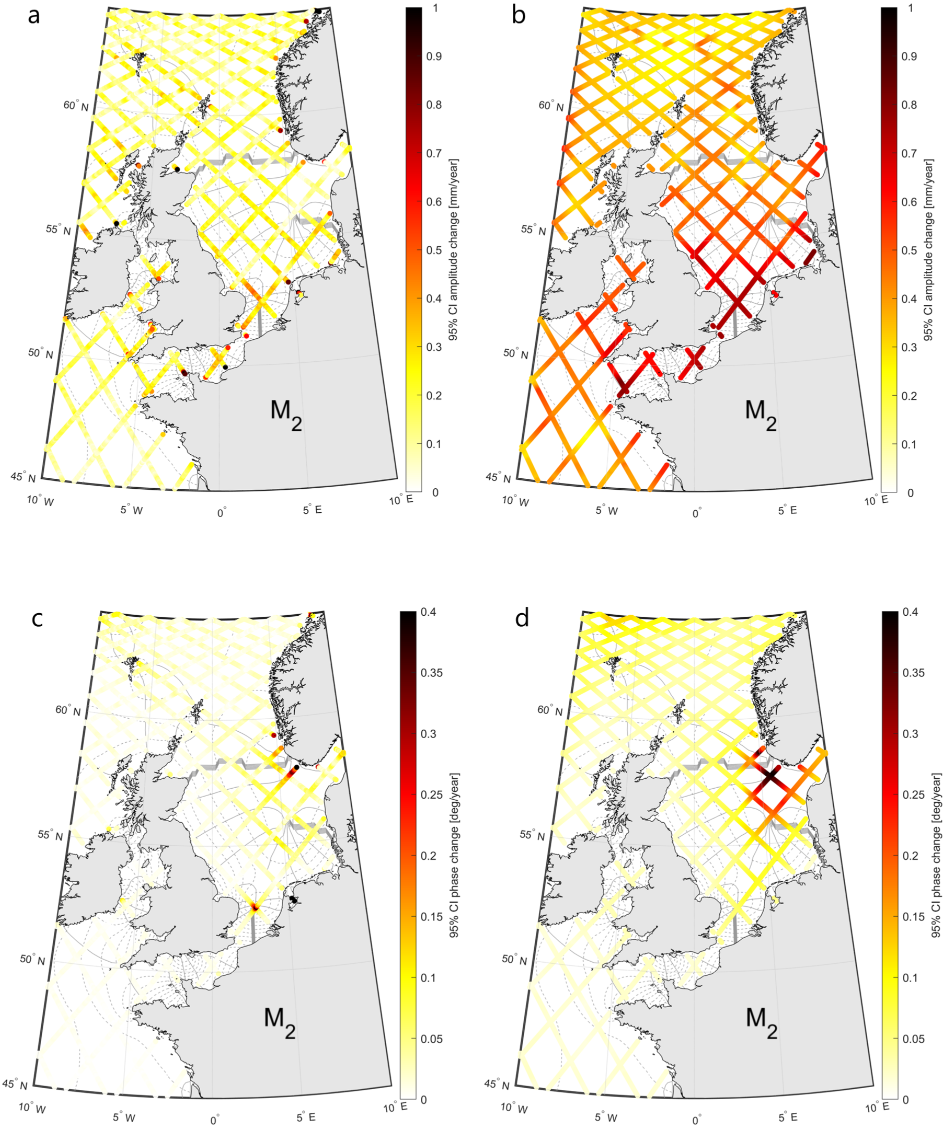


Figure 2.10: 95% confidence intervals for trend estimates derived from confidence intervals computed by UTide (a, c) and from standard errors derived from GTSM (b, d) for M_2 amplitudes (a, b) and M_2 phases (c, d).

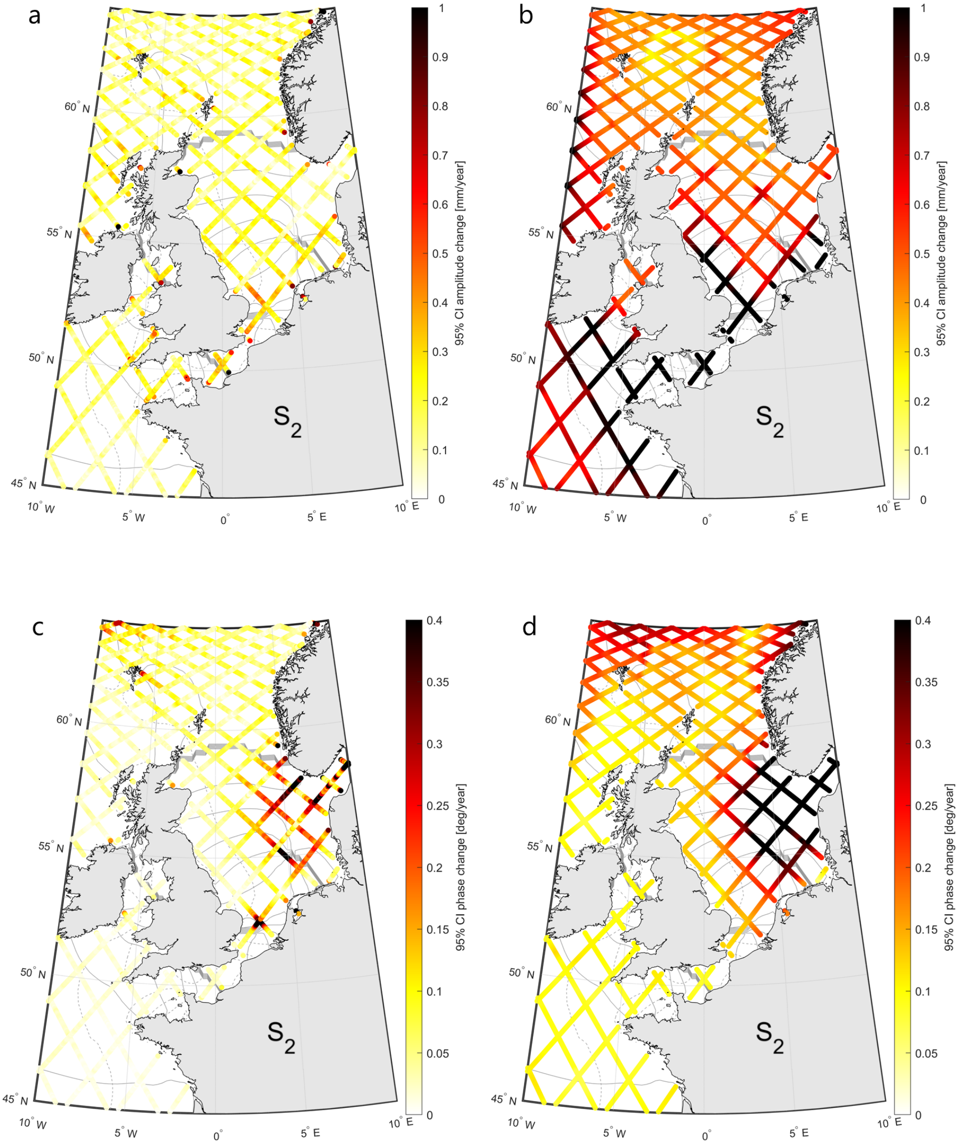


Figure 2.11: 95% confidence intervals for trend estimates derived from confidence intervals computed by UTide (a, c) and from standard errors derived from GTSM (b, d) for S_2 amplitudes (a, b) and S_2 phases (c, d).

2.5. DISCUSSION AND CONCLUSIONS

Using the full record of sea level measurements by the TOPEX/Poseidon and Jason satellites (1993–2020), a global estimate of the secular trends in M_2, S_2, O_1, and K_1 tidal harmonic constants was obtained. While satellite altimetry is routinely used for tidal analyses, this is the first time it was used to study secular trends on a global scale. With this, the presented study attempts to fill in the gaps left by earlier studies on secular changes in tides, that were predominantly based on data from tide gauges. However, compared to tide gauges, the temporal resolution of the satellite data is limited. Consequently, several years of data were required to prevent aliasing and obtain reliable tidal estimates. Therefore, the method that is typically used to study secular changes in tides from tide gauge data, by means of yearly harmonic analysis, could not be applied. In this paper two alternative approaches were implemented. The first method (SegHA) is very similar to the yearly analysis except now the time series were divided into periods of four years. Compared to the yearly analysis, this reduces the number of consecutive independent tidal estimates and hence the redundancy in trend fitting and the significance of the estimated trends. Moreover, with this approach uncertainty estimates were obtained through a simplified error propagation whereby any correlation between the amplitude and phase estimates was ignored. However, this approach can be carried out with the standard available tidal software and allows a straightforward implementation of non-linear changes. Then, in the second approach (TintHA), the linear change in tidal constants was estimated during the harmonic analysis. This way, the entire time series could be analysed at once, which reduced the issue of aliasing. In the TintHA approach, no empirical correction for a possible residual nodal modulation was derived. However, results from the SegHA approach suggest this residual to be not significant on global scale (not shown here). Moreover, both methods produced very similar results, both when applied to the sub-sampled tide gauge data (Figure 2.2) and to the actual satellite radar altimeter data. Due to the rather low magnitudes of secular trends in tides (Figure 2.4, 2.5 and 2.9, C.5 and C.6), in many regions the estimated trends just exceed the confidence levels (see Figures 2.6, 2.10, 2.11, C.1 and C.2).

2.5.1. SATELLITE-DERIVED SECULAR CHANGE IN TIDES

The main findings presented in this paper are as follows. The amplitudes of the considered tides have changed by up to 1 mm/year over the past ~ 3 decades. This implies a change of up to 10 cm per century. The change in total tidal range remains unsure because many tidal constituents are not resolvable with the available data. Whether the amplitudes were subject to an increase or a decline varies on a regional (mainly applies to S_2 and K_1) to even local basis (M_2, O_1). On the North West European Shelf, relatively large phase changes are observed close to amphidromic points (Figure 2.9c, d) which may suggest a displacement of these points. This could also be a (partial) explanation for observed variability in changes in amplitude (Figure 2.9a, b), as a displacement of an amphidromic point would reduce the tidal amplitude in the direction of the displacement and increase the amplitude in the opposite direction. This may also explain the differences in the sign of the observed changes on relatively small spatial scales. However, note that from the experiment with tide gauge data it followed that the accuracy of derived phase changes reduces strongly when tidal amplitudes are low (Figure 2.2b, d), which is the case near amphidromic points.

Results have also been compared to trend estimates from tide gauges. However, caution is required. The distance between tide gauges and the nearest TPJ-data exceeds ~ 50 km in most cases and processes that affect tides near the coast may be very different from those at open sea. Moreover, the estimated trends at crossovers are spatial averages, that may not correspond to the signal that is observed at tide gauges (being point estimates). For M_2, similar secular changes were derived from altimetry and nearby tide gauges along the coast of the Netherlands, France, Denmark and the south of Norway (Figure 2.9a, 2.9c). Discrepancies were observed along the west coast of Norway and in shallow waters around the United Kingdom (UK). The former may be related to the Norwegian coastal current and associated mesoscale activity. While the application of the 'mesoscale correction' significantly improved the consistency in trend estimates from crossing tracks, still more inconsistencies are observed in regions where mesoscale variability is also large. In addition, the capricious nature of the coastline (both Norway and UK) may have affected estimates from the altimeter data and/or explain differences between coastal (tide gauge) and shelf estimates (altimeter). Fewer similarities were observed for the S_2 tide. These were predominantly restricted to the Dutch coast. On the global scale, reasonable similarities between secular change derived from tide gauges and altimeter data were observed for S_2 (59%), $O1_2$ (64%) and K_1 (63%) while for M_2, this was only the case for 41% of the tide gauge-crossover combinations. However, such comparisons may be deceptive given the distance between crossovers and the nearest tide gauge (Figure 2.1a) and significant spatial variability in amplitude change (for M_2 in particular) observed at this distance (Figure 2.1b). Strictly speaking, the comparison between tide gauge- and TPJ-derived estimates is therefore not a validation, because the estimates refer to different locations.

2.5.2. UTIDE- VERSUS GTSM-DERIVED CONFIDENCE INTERVALS

The presented study considers both uncertainties estimates derived from UTide and computed with GTSM reanalysis data. From a comparison of both products (e.g. Figure 2.10 and 2.11) it appears that the uncertainties are most likely underestimated by UTide. In particular on shelf regions, one would expect larger uncertainties due to larger non-tidal residuals and unresolved shallow water tides, while the UTide-derived uncertainties are equally low and homogeneous in shelf regions as on the open ocean. Likely, the application of the 'mesoscale correction' in shallow water removes some tidal signal that is aliased in the SLA product that was used for this correction (Zaron and Ray, 2018). This would reduce the residuals and hence it may have caused too optimistic uncertainty estimates by UTide. Moreover, from the analysis described in Section 2.3.3, it follows that the confidence intervals estimated by UTide are relatively unaffected by changes in data availability. This also followed from an additional experiment using random subsets of the TPJ-data at crossovers (Appendix B). Comparing Figures 2.6a and B.1b, it appears that the UTide-derived uncertainties are of similar magnitude for the full crossover time series as for half the amount of data, while dividing the data in half may actually cause significant differences in trend estimate (Figure B.1a). This finding puts to question the reliability of the UTide-derived confidence intervals. On the other hand, the impact of data availability on the uncertainty is reflected by the GTSM uncertainties (e.g., see Figures 2.6b and B.1c), indicating the added value of this product. Nevertheless, the GTSM-derived uncertainties could for instance not explain all ambiguities at crossing tracks (Figure 2.9a-d). In addition, from the comparison of the confidence intervals using tide gauge data (Section 2.3), both the UTide- and GTSM-derived products appear to underestimate the uncertainties to some extent. However, one should keep in mind that these coastal results cannot directly be transferred to open ocean. For instance, GTSM does not include all physical processes that affect the water level near the coast (e.g. river outflows) and the relation between the spectrum of non-tidal water levels at the coast and on open ocean is unknown. For future studies, we recommend the use of a full 3D model that allows for direct comparison of the observed and reanalysis data. Such a model would include the mesoscale variability that is also present in the satellite radar altimeter data and allow for the same corrections to be applied to both the satellite data and reanalysis data. Moreover, from Figure 2.3b and Appendix B it follows that concerning the phase changes, the UTide- and GTSM-derived confidence intervals are more similar and both appear to underestimate the actual uncertainty. The analysis of phase changes on the North West European Shelf suggests that the GTSM-derived confidence intervals in the vicinity of amphidromic points may be too optimistic. This can be explained by the fact that the locations used for the GTSM confidence intervals do not coincide with the TPJ-tracks and/or location of amphidromic points. Ideally, the model-based confidence intervals should be derived at the exact location of the satellite data but this kind of data was not available. Finally, the uncertainty may be reduced by the inclusion of data from other satellite missions. However, given the low magnitude of the observed secular trends, even small intermission biases in the range corrections could be easily mistaken for changes in the actual tides and should thus be taken appropriate care of.

2.5.3. POTENTIAL SECULAR CHANGES INTRODUCED BY SATELLITE DATA PROCESSING

The magnitude and strong regional variability of the secular change in M_2 amplitudes corresponds to findings by other studies based on tide gauge data (e.g., Müller et al., 2011; Schindelegger et al., 2018; Woodworth, 2010). However, the altimeter-derived change in the S_2 tide differs from some documented findings (e.g., Ray, 2009; Woodworth, 2010). For instance, they found the S_2 amplitudes to have increased along the Gulf of Alaska. This contrasts both to what is derived from altimeter data at open ocean and our analysis of GESLA-3 tide gauge records (Figure 2.4b, 2.7b). This suggests that the difference may be related to the differences in considered periods. On the other hand, the inconsistencies may be associated with the atmospheric loading correction (DAC). This correction was applied to TPJ-water levels to reduce the impact of aliasing of non-tidal water level variation on the estimation of tidal harmonic constants. For the sake of consistency, the same correction was applied to the tide gauge data, which is typically not done in earlier studies on tide gauge data. Therefore, a possible S_2-like signal in DAC (for instance related to the six-hour resolution of the product) may have affected the results.

More general, it is possible that any systematic error or secular change in the atmospheric propagation (wet/dry troposphere and ionosphere) or reference frame corrections could have affected the trend estimates derived from the water levels. According to Zawadzki et al. (2018), estimation of the S_2 tide is particular sensitive to (errors in) the geophysical/range corrections. Additional trend analysis of the individual corrections suggest limited influence on M_2 and O_1 (Figures C.7, C.10). Only the wet troposphere correction shows secular change in the amplitudes of up to 0.15 mm/year. The magnitude of these changes is low compared to the trend in the water level (~ 1 mm/year) and the computed correlation is insignificant. Larger signals are observed for S_2 (Figure C.8; up to 0.25 mm/year). However, none of these signals show significant correlation with the secular change derived from the water levels. In addition, a possible effect from errors in the (model-derived) ionospheric correction on the S_2 tide was already incorporated in the confidence intervals for this tide (Appendix A). Some K_1 signal is observed in the wet troposphere and altimeter derived ionosphere correction. However, again, the correlation with water level-derived change is low (< 0.1) and positive. Since the corrections are subtracted from the range to obtain the water level, only a negative correlation would explain the secular change derived from the water levels. Finally, there may be intermission biases in range corrections that could be partly responsible for the observed trend in S_2 amplitudes, such as the CG-correction that was applied to TOPEX/Poseidon data (Beckley et al., 2021; Zawadzki et al., 2018). All in all, the analysis of the S_2 remains tricky and a more thorough analysis is deemed necessary.

2.5.4. EXPLAINING THE OBSERVED SECULAR CHANGES BY MEANS OF PHYSICAL PROCESSES

The results presented in this paper merely allow speculation about the drivers behind the observed trends. The strong local variability in some areas suggests that local processes may dominate there or that the observed change is in fact related to internal tide variability. For instance, Zhao (2016) showed that in several regions, the propagation of the internal tidal wave is subject to interannual or decadal variability. This causes temporal differences in the phase of the internal tide that increase as the internal tide propagates. Since the tidal amplitude that is observed at the surface is a combination of the barotropic and internal tide, its value depends on the phase difference between the two. As mentioned, the interannual change in these phase differences can vary as the internal tide travels further from its origin, which may cause an apparent increase in observed tidal amplitude at one crossover, but a decrease at the next. Regions where significant small-scale variability in trend estimates is observed (e.g., 30° S, 160° W; 20° N, 30° W; 15° S, 50° E), correspond to locations where the amplitude of the internal M_2 amplitude is rather high (Zhao, 2016). On the other hand, regions where the internal M_2 amplitude is low (i.e. equatorial Pacific) correspond to regions where spatial variability in yearly change in M_2 amplitude is also low.

On the other hand, part of the observed signal could be related to sea ice decline (e.g., Haigh et al., 2020). Namely, the observed changes in M_2 amplitude around Iceland (Figure 2.4a) are of opposite sign compared to the March-September amplitude differences documented by Bij de Vaate et al. (2021). This indicates that over time the annual average tide becomes closer to the September case, which is in line with interannual sea ice decline. This may also explain the increased spatial correlation in observed trends in M_2 amplitudes near the poles. Furthermore, changes in tides have been linked to sea level rise. For instance, the modelled effect of SLR on M_2 amplitudes was found to be $\sim$ 10 cm/m SLR (e.g., Pickering et al., 2017; Schindelegger et al., 2018). This, given a SLR of $\sim$ 3 mm/year since 1990, is of comparable magnitude to the TPJ-derived amplitude changes in most regions ($\sim$ 0.3 mm/year). However, the modelled M_2 amplitude change under the influence of SLR does not exhibit the large regional variability that was seen in the altimetry-derived trends, although a number of similarities can be observed on for instance the North West European Shelf. On another note, the zonal pattern in the S_2 amplitude change is striking and not as strong for the other (lunar) tides. If the observed change is in fact related to the tide and not to other non-tidal processes, this suggests the causes may be related to radiational forcing. About 15% of the S_2 tide is driven by pressure loading of the ocean (Haigh et al., 2020) and interannual variability in atmospheric pressure could translate into variable S_2 amplitudes. Given that atmospheric pressure fluctuates continuously (Lu and Tu, 2021), it may be that the secular change in S_2 amplitude cannot be accurately described by a linear trend. Finally, although we can at this stage not draw conclusions on the drivers behind the observed changes in tides, our findings could be useful for future (modelling) studies on this phenomenon.

BIBLIOGRAPHY

Beckley, B., Ray, R. D., Zelensky, N., Lemoine, F., Brown, S., Desai, S., & Mitchum, G. (2021). Integrated Multi-Mission Ocean Altimeter Data for Climate Research TOPEX/Poseidon, Jason-1, 2, & 3 User's Handbook Version 5.1.

Bij de Vaate, I., Slobbe, D. C., & Verlaan, M. (2022). Secular trends in global tides derived from satellite radar altimetry. *Journal of Geophysical Research: Oceans*, *127*(10), e2022JC018845. https://doi.org/https://doi.org/10.1029/2022JC018845

Bij de Vaate, I., Vasulkar, A. N., Slobbe, D. C., & Verlaan, M. (2021). The Influence of Arctic Landfast Ice on Seasonal Modulation of the M2 Tide. *Journal of Geophysical Research: Oceans*, *126*(5), e2020JC016630. https://doi.org/https://doi.org/10.1029/2020JC016630

Cartwright, D. E., & Edden, A. C. (1973). Corrected tables of tidal harmonics. *Geophys. J. Roy. Astron. Soc.*, *33*, 253–264. https://doi.org/doi:10.1111/j.1365-246X.1973.tb03420.x

Cartwright, D. E., & Taylor, R. J. (1971). New computations of the tide-generating potential. *Geophys. J. Roy. Astron. Soc.*, *23*, 45–74. https://doi.org/doi:10.1111/j.1365-246X.1971.tb01803.x

Cherniawsky, J. Y., Foreman, M. G., Kang, S. K., Scharroo, R., & Eert, A. J. (2010). 18.6-year lunar nodal tides from altimeter data. *Continental Shelf Research*, *30*(6), 575–587.

Codiga, D. L. (2011). *Unified Tidal Analysis and Prediction Using the UTide Matlab Functions. Technical Report 2011-01*. Graduate School of Oceanography, University of Rhode Island: Narragansett.

Codiga, D. L. (2020). UTide Unified Tidal Analysis and Prediction Functions. MATLAB Central File Exchange. [Software] [Last checked on 10/06/2020].

Dangendorf, S., Calafat, F. M., Arns, A., Wahl, T., Haigh, I. D., & Jensen, J. (2014). Mean sea level variability in the North Sea: Processes and implications. *Journal of Geophysical Research: Oceans*, *119*(10).

Devlin, A. T., Jay, D. A., Talke, S. A., Zaron, E. D., Pan, J., & Lin, H. (2017). Coupling of sea level and tidal range changes, with implications for future water levels. *Scientific Reports*, *7*(1), 1–12.

Foreman, M. G. G. (2004). *Manual for tidal heights analysis and prediction, revised*. Pacific Marine Science Rep. 77, 10.

Guarneri, H., Verlaan, M., Slobbe, D. C., Veenstra, J., & Zijl, F. (2022). Estimating tides from satellite radar altimetry in coastal seas [Manuscript in preparation]. *Marine Geodesy*, *20*, XX–XX.

Hagen, R., Plüß, A., Jänicke, L., Freund, J., Jensen, J., & Kösters, F. (2021). A combined modeling and measurement approach to assess the nodal tide modulation in the North Sea. *Journal of Geophysical Research: Oceans*, *126*, e2020JC016364.

Haigh, I. D., Marcos, M., Talke, S. A., Woodworth, P. L., Hunter, J. R., Haugh, B. S., ..., & Thompson, P. (2021). GESLA Version 3: A major update to the global higher-frequency sea-level dataset. *EarthArXiv*. https://doi.org/https://doi.org/10.31223/X5MP65

Haigh, I. D., Pickering, M. D., Green, J. M., Arbic, B. K., Arns, A., Dangendorf, S., ..., & Woodworth, P. L. (2020). The tides they are a-Changin': A comprehensive review of past and future nonastronomical changes in tides, their driving mechanisms, and future implications. *Reviews of Geophysics*, *58*(1), e2018RG000636.

Hinton, A. C. (2000). Tidal changes and coastal hazards: past, present and future. *Natural Hazards*, *21*(2), 173–184.

Jee, G., Lee, H. B., Kim, Y. H., Chung, J. K., & Cho, J. (2010). Assessment of GPS global ionosphere maps (GIM) by comparison between CODE GIM and TOPEX/Jason TEC data: Ionospheric perspective. *Journal of Geophysical Research: Space Physics*, *115*(A10).

Jin, G., Pan, H., Zhang, Q., Lv, X., Zhao, W., & Gao, Y. (2018). Determination of harmonic parameters with temporal variations: An enhanced harmonic analysis algorithm and application to internal tidal currents in the South China Sea. *Journal of atmospheric and oceanic technology*, *35*(7), 1375–1398.

Khojasteh, D., Glamore, W., Heimhuber, V., & Felder, S. (2021). Sea level rise impacts on estuarine dynamics: A review. *Science of The Total Environment*, 146470.

Li, S., Wahl, T., Talke, S. A., Jay, D. A., Orton, P. M., Liang, X., ..., & Liu, L. (2021). Evolving tides aggravate nuisance flooding along the US coastline. *Science Advances*, *7*(10), eabe2412.

Lu, E., & Tu, J. (2021). Relative importance of surface air temperature and density to interannual variations in monthly surface atmospheric pressure. *International Journal of Climatology*, *41*, E819–E831.

Müller, M. (2012). The influence of changing stratification conditions on barotropic tidal transport and its implications for seasonal and secular changes of tides. *Continental Shelf Research*, *47*, 107–118. https://doi.org/https://doi.org/10.1016/j.csr.2012.07.003

Müller, M., Arbic, B. K., & Mitrovica, J. X. (2011). Secular trends in ocean tides: Observations and model results. *Journal of Geophysical Research: Oceans*, *116*(C5).

Müller, M., Cherniawsky, J. Y., Foreman, M. G. G., & von Storch, J. S. (2014). Seasonal variation of the M 2 tide. *Ocean Dynamics*, *64*(2), 159–177. https://doi.org/https://doi.org/10.1007/s10236-013-0679-0

Naeije, M. C. (2022). Radar Altimeter Database System. [Data archive]. http://rads.tudelft.nl

Pawlowicz, R., Beardsley, B., & Lentz, S. (2002). Classical tidal harmonic analysis including error estimates in MATLAB using T-TIDE. *Computers & Geosciences*, *28*, 929–937.

Pickering, M. D., Horsburgh, K. J., Blundell, J. R., Hirschi, J. M., Nicholls, R. J., Verlaan, M., & Wells, N. C. (2017). The impact of future sea-level rise on global tides. *Continental Shelf Research*, *142*, 50–68.

Ray, R. D. (2009). Secular changes in the solar semidiurnal tide of the western North Atlantic Ocean. *Geophysical Research Letters*, *36*(29), 108–116.

Ray, R. D. (2016). On measurements of the tide at Churchill, Hudson Bay. *Atmosphere-Ocean*, *54*(2), 108–116.

Ray, R. D. (2020). Daily harmonics of ionospheric total electron content from satellite altimetry. *Journal of Atmospheric and Solar-Terrestrial Physics*, *209*, 105423.

Ray, R. D., & Zaron, E. D. (2016). M_2 Internal Tides and Their Observed Wavenumber Spectra from Satellite Altimetry. *Journal of Physical Oceanography*, *46*(1), 3–22.

Ross, A. C., Najar, R. G., Li, M., Lee, S. B., Zhang, F., & Liu, W. (2017). Fingerprints of sea level rise on changing tides in the Chesapeake and Delaware Bays. *Journal of Geophysical Research: Oceans*, *122*, 8102–8125. https://doi.org/doi:10.1002/2017JC012887

Rousseeuw, P. J., & Croux, C. (1993). Alternatives to the median absolute deviation. *Journal of the American Statistical association*, *88*(424), 1273–1283.

Savcenko, R., & Bosch, W. (2007). Residual Tide Analysis in Shallow Water-Contributions of ENVISAT and ERS Altimetry. *ESA-SP636 (CD-ROM), Proceedings of the Envisat Symposium*.

Scharroo, R., Leuliette, E. W., Naeije, M. C., Martin-Puig, C., & Pires, N. (2016). RADS Version 4: An efficient way to analyse the multi-mission altimeter database. [ESA Special Publication SP-740]. *Proceedings of the ESA Living Planet Symposium*, 9–13.

Schindelegger, M., Green, J. A. M., Wilmes, S. B., & Haigh, I. D. (2018). Can we model the effect of observed sea level rise on tides? *Journal of Geophysical Research: Oceans*, *123*, 4593–4609.

Schrama, E. J. O., & Ray, R. D. (1994). A preliminary tidal analysis of TOPEX/POSEIDON altimetry. *Journal of Geophysical Research: Oceans*, *99*(C12), 24799–24808.

St-Laurent, P., Saucier, F. J., & Dumais, J. F. (2008). On the modification of tides in a seasonally ice-covered sea. *Journal of Geophysical Research: Oceans*, *113*(11), 1–11. https://doi.org/https://doi.org/10.1029/2007JC004614

Taburet, G., Sanchez-Roman, A., Ballarotta, M., Pujol, M. I., Legeais, J. F., Fournier, F., ..., & Dibarboure, G. (2019). DUACS DT2018: 25 years of reprocessed sea level altimetry products. *Ocean Science*, *15*(5), 1207–1224.

Wahr, J. M. (1985). Deformation induced by polar motion. *Journal of Geophysical Research*, *90*(B11), 9363–9368. https://doi.org/doi:10.1029/JB090iB11p09363

Wang, X., Verlaan, M., Apecechea, M. I., & Lin, H. X. (2021). Computation-efficient Parameter Estimation for a High-Resolution Global Tide and Surge Model (GTSM). *Journal of Geophysical Research: Oceans*, *126*, e2020JC016917. https://doi.org/https://doi.org/10.1029/2020JC016917

Woodworth, P. L. (2010). A survey of recent changes in the main components of the ocean tide. *Continental shelf research*, *30*(15), 1680–1691.

Zaron, E. D., & Jay, D. A. (2014). An analysis of secular change in tides at open-ocean sites in the Pacific. *Journal of Physical Oceanography*, *44*(7), 1704–1726.

Zaron, E. D., & Ray, R. D. (2018). Aliased Tidal Variability in Mesoscale Sea Level Anomaly Maps. *Journal of Atmospheric and Oceanic Technology*, *35*(12), 2421–2435.

Zawadzki, L., Ablain, M., Carrere, L., Ray, R. D., Zelensky, N. P., Lyard, F., Guillot, A., & Picot, N. (2018). Investigating the 59-Day Error Signal in the Mean Sea Level Derived From TOPEX/Poseidon, Jason-1, and Jason-2 Data With FES and GOT

Ocean Tide Models. *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, 56*(6), 3244–3255.

Zhao, Z. (2016). Internal tide oceanic tomography. *Geophysical Research Letters, 43*(17), 9157–9164.

3

SPATIOTEMPORAL VARIABILITY IN GLOBAL STORM SURGE WATER LEVELS

Multi-mission satellite radar altimeter data have been used to study the spatial and temporal variability in global storm surge water levels. This was done by means of a time-dependent extreme value analysis applied to the monthly maximum detided water levels. To account for the limited temporal resolution of the satellite data, the data were first stacked on a $5^\circ \times 5^\circ$ *grid. Moreover, additional scaling was applied to the extreme value analysis for which the scaling factors were determined by means of a resampling method using reanalysis data. In addition to the conventional analysis using data from tide gauges, this study provides an insight in the ocean wide storm surge properties. Nonetheless, where possible, results were compared to similar information derived from tide gauge data. Except for secular changes, the satellite-derived results are comparable to the information derived from tide gauges (correlation* > 0.5*), although the tide gauges show more local variability. Where limited correlation was observed for the secular change, it was suggested that the satellites may not be able to fully capture the temporal variability in the short-lived, tropical storms, as opposed to extra-tropical storms.*

This chapter has been submitted as Bij de Vaate, I., Slobbe, D. C., & Verlaan, M. (2023). Mapping the spatiotemporal variability in global storm surge water levels using satellite radar altimetry. *Under review at Ocean Dynamics*

3.1. INTRODUCTION

Coastal zones are densely populated and subject to higher rates of population growth than other regions (McGranahan et al., 2007). At the same time, these regions are extremely vulnerable to many impacts of climate change, one of which is coastal flooding (IPCC, 2021). The latter is closely linked to the occurrence of extreme sea level events (ESLs) and observations indicate that such events have increased both in frequency and magnitude (e.g., Oppenheimer et al., 2019; Wahl et al., 2017). Projections suggest this trend to continue in the future (e.g., Oppenheimer et al., 2019). While observations indicate that on the global scale, sea level rise is the primary driver behind the increase in ESLs (e.g., Oppenheimer et al., 2019; Wahl et al., 2017), locally the increase in ESLs may be amplified or even dominated by climate change induced changes in storminess (e.g., Feng et al., 2019; Haigh et al., 2014; Rashid and Wahl, 2020).

The strong winds and low atmospheric pressure associated with (extra)tropical storms cause a rise in the water level, a so called storm surge (Resio and Westering, 2008). Several studies suggest that climate change affects the frequency and severity of (extra)tropical storms, which may exacerbate the build up of storm surges in the future (e.g., Calafat et al., 2022; Gori et al., 2022). However, as is stated in the latest IPCC report (IPCC, 2021), there is overall low confidence in any observed changes in the frequency or magnitude of storm events and their contribution to ESLs. This is mainly related to the rare and short-lived nature of the storm events and the large local variability (IPCC, 2021). The local variability in storm events is particularly important to acknowledge, given that the observational evidence on storm surges predominantly includes in-situ data from tide gauges (e.g., Muis et al., 2016). Hence, the information on storm surge water levels is restricted to coastal regions with an uneven distribution across the globe. Consequently, it is uncertain how storm surges develop on the open ocean or to what extent they are subject to temporal variability on the global scale. This also affects the validation of global surge models and any predictions that are made using such models.

In this respect, the use of satellite data may provide a significant contribution to our understanding. Over the past decades, measurements of the sea surface height (SSH) have been collected by a series of satellite radar altimeters, providing a record of instantaneous SSH with quasi-global coverage (Adebisi et al., 2021). Although the temporal resolution of satellite altimeters is relatively limited (typical revisit period is 9.9156 days (Beckley et al., 2021)), previous studies have shown that consequent issues may be overcome by combining data from multiple missions and applying data stacking (e.g., Bij de Vaate et al., 2021; Cancet et al., 2015). Up to now, only few studies have been devoted to mapping storm surges using satellite altimetry over an extended period of time (e.g., Andersen et al., 2015; Antony et al., 2014; Ji et al., 2019). These studies have shown that with adequate data processing it is possible to derive comparable storm surge properties from satellite data as from nearby tide gauges. However, in these studies the analyses were focused on specific geographic regions, and it remains uncertain if and how the ability to derive storm surges from satellite data varies with respect to, for instance, regional differences in the nature of storms (e.g., tropical versus extra-tropical) or spatial differences in data availability from satellite data. Moreover, the aforementioned examples have not studied temporal changes in storm surges.

In this paper, we present a full ocean-wide mapping of quasi-global storm surge water levels and their time-variability. We focus both on seasonal variations and secular changes in the magnitude of storm surge water levels. In addition, results are compared to similar information derived from a global tide gauge data set. In the following sections, we first give an outline of the data used in this study, before introducing the methodology and discussing the results.

3.2. DATA

3.2.1. SATELLITE RADAR ALTIMETRY

This study exploited SSH measurements from eight satellite radar altimeters, that are: TOPEX/Poseidon and Jason 1-3 (further refered to as TPJ), ERS-1 and ERS-2, Envisat-1 and SARAL. The data were combined for the full span of the TPJ-satellites, resulting in 29 years of sea level data (1993-2021). All altimeter data were obtained through the Radar Altimeter Data System (RADS; Naeije, 2022). The following geophysical and range corrections were applied (Scharroo et al., 2016): ionosphere (NIC09 for TOPEX/Poseidon and ERS-1, GIM for others), dry troposphere (ECMWF), wet troposphere (if available: radiometer, otherwise: ECMWF), solid tide (Cartwright and Edden, 1973; Cartwright and Taylor, 1971), pole tide (Wahr, 1985), ocean tide and load tide (FES2014), mean sea surface (DTU15-MSS), the sea state bias (BM3 for ERS-1 and ERS-2, Tran et al. (2018) for SARAL/AltiKa and CLS for all others) and the reference frame offset. Data from all satellites were stacked on a rectangular lat/lon grid with variable dimensions (see Section 3.3.2). Subsequently, the dynamic atmosphere correction (DAC) (MOG2d (ERA Interim forcing)) was used to detect and remove outliers in the time series. Here, data were classified as outliers when their absolute deviation from the median exceeds three times the 5-day moving maximum of the DAC. Finally, to be able to isolate the effect of variability in storminess on the ESLs, the yearly mean sea level was removed from the time series by subtracting their 365-day moving mean. Note that the remaining water level may still reflect other processes (e.g., mesoscale variability or non-linear interactions between tides and surges), but these were assumed to be negligible compared to the actual surge variability.

3.2.2. TIDE GAUGES

Alongside the altimeter data, data from a selection of tide gauges (obtained from the GESLA-3 dataset (Haigh et al., 2022)) were processed to allow for a comparison of the derived surge properties. For this purpose, we only considered tide gauge data from the period for which satellite altimeter data are available (1993-2021). The temporal resolution of the tide gauge data varies from one minute to one hour, mainly depending on the country where the stations are located and the time of data acquisition. In contrast to the altimeter data, the tide gauge data were detided using tidal constants obtained by means of a harmonic analysis using the UTide-software (Codiga, 2020) instead of using FES. This was done to remove as much of the tidal signal as possible, thus including (shallow water) tides not contained in FES, but possibly present at the location of most tide gauges. Finally, the time series were referenced to the annual mean sea level by subtracting the 365-day moving mean from the time series.

3.2.3. REANALYSIS DATA

Reanalysis data were used to compute scaling factors to account for the limited temporal resolution of the satellite-derived water levels (as described in Section 3.3.2). These data were obtained from the Global Tides and Surge Model (GTSM; Muis et al., 2016; Wang et al., 2021), forced by ERA5 reanalysis data. GTSM is a barotropic (2D) model that runs on an unstructured grid with a resolution that increases from 25 km on the open oceans to 2.5 km at the coast. GTSM has previously been used for a global reanalysis of storm surge water levels, where validation using observed time series at tide gauges resulted in an average RMSE of only 0.11 m (Muis et al., 2016). In this study, time series with a sampling rate of 10 minutes were reconstructed for the full TPJ-period. This was done for over 600 locations covering the global oceans. Subsequently, the time series were detided and referenced to the annual mean sea level in a similar way as was done with the tide gauge data.

3.3. METHODS

3.3.1. EXTREME VALUE ANALYSIS

Earlier studies that used satellite radar altimetry to observe storm surges typically used a so-called peak-over-threshold (POT) approach to identify single surge events that exceed a given magnitude (Andersen et al., 2015; Antony et al., 2014; Ji et al., 2019). While this is a common approach in extreme value analysis, its use in the presented study would have one important drawback. Namely, the POT approach relies on a threshold that is defined a priori, thus prohibiting its use for studying temporal and spatial variability in surge water levels (as this would require setting separate thresholds for different locations and periods (Butler et al., 2007)). Another approach that has been used in storm surge analyses is the so-called block-maxima approach (e.g., Butler et al., 2007; Muis et al., 2016). Such an approach relies on a relative definition of an extreme value (with respect to other water levels in the specified block) and can thus be unambiguously applied to records with varying magnitudes of variability (Butler et al., 2007). Similar to studies by Méndez et al. (2007), Menéndez and Woodworth (2010), and Lobeto et al. (2018) on ESLs, we opted for the use of monthly maxima surge water levels (*MM*) to be able to study seasonal differences. Then, as was also done in the aforementioned studies, a time-dependent Generalized Extreme Value Distribution (GEVD) was fitted to the series of *MM*. The GEVD (Equation 3.1) is a three-parameter distribution where μ is the location parameter, σ the scale parameter, and ξ the shape parameter. Depending on the shape parameter, the GEVD can belong to the Weibull, Gumbel or Fréchet family (Lobeto et al., 2018). In addition, inspired by Izaguirre et al. (2011), two scaling factors (k_1 for the location parameter and k_2 for the shape parameter) were included in the distribution to account for the reduced and globally variable temporal sampling by the satellites. This resulted in the following GEVD:

$$F(MM) = \begin{cases} \exp(-[1+\xi(\frac{MM-k_1\mu(t)}{k_2\sigma(t)})]^{-\frac{1}{\xi}}) & \xi \neq 0 \\ \exp(-\exp[-(\frac{MM-k_1\mu(t)}{k_2\sigma(t)})]) & \xi = 0 \end{cases} \tag{3.1}$$

Here the temporal variability was included in $\mu(t)$ and $\sigma(t)$ as follows, where $\omega = 2\pi$ year^{-1} and t is given in years with respect to the center time of the considered data. $\alpha_{[0,1,2,3]}$ and $\beta_{[0,1,2,3]}$ were estimated:

$$\begin{aligned} \mu(t) &= \alpha_0 + \alpha_1 t + \alpha_2 \cos(\omega t) + \alpha_3 \sin(\omega t) \\ \sigma(t) &= \beta_0 + \beta_1 t + \beta_2 \cos(\omega t) + \beta_3 \sin(\omega t) \end{aligned} \tag{3.2}$$

The model parameters were estimated by means of the maximum likelihood method as described in Méndez et al. (2007). The scaling factors (k_1 and k_2) were determined a priori based on experiments with the reanalysis data (see Section 3.3.2). In addition to satellite data, high-frequency data from tide gauges were processed in a similar manner as the satellite data, except the scaling factors which now were excluded from the GEVD.

3.3.2. DATA AVAILABILITY AND SCALING FACTORS

Although we have used satellite data from multiple missions and stacked the data over larger areas, it is likely that not all storm surges were captured in full magnitude. To account for this, two scaling factors were introduced in the GEVD. Since we considered that the predominantly coastal tide gauge data would not be representative for ocean-wide surge variability, the values of these scaling factors were determined by means of global reanalysis data. For this purpose, the model-derived time series, introduced in Section 3.2.3, were sub-sampled to satellite sampling while the start time of the sub-sampling was incrementally shifted 25 times. Subsequently, for each iteration, the MM were computed and the GEVD was fitted to the data, including the scaling factors k_1 and k_2. The optimal scaling factors were determined by minimizing the root-mean-square-error (RMSE) between the GEVD based on the original high-frequency data and the GEVD based on the sub-sampled data. Finally, the optimal, location-specific scaling factors were computed as the median of the values resulting from the 25 iterations. At the same time, the standard error (SE) of each scaling factor was computed as 1.4826 times the median absolute deviation of the 25 values. This experiment was done for two different grid sizes ($2^\circ \times 2^\circ$ and $5^\circ \times 5^\circ$) to gain insight in the impact of the size of the area over which satellite data are stacked, on the ability to accurately capture storm surge dynamics. Note that the precise satellite sampling and data availability also varies across the globe as the satellite tracks converge towards the poles, while on the other hand, the metric distance between meridians reduces towards the poles and the (seasonal) presence of sea-ice may hamper measurements of the sea level.

To assess the use of reanalysis data for the computation of the scaling factors on the open ocean, the same procedure was carried out using a selection of tide gauges that were located at least 100 km from the continental coastlines. Only tide gauges were used that have at least multiple years of gapless data. To prevent the validation being affected by temporal differences in storm surge magnitude, the reanalysis time series were first trimmed to the exact time span of the tide gauge time series. However, note that the time spans of the available data varies per tide gauge, hence truncating the time series may introduce additional spatial variation in obtained scaling factors.

3.3.3. Storm Surge Characteristics

From the estimated location, shape and tail parameters, the storm surge water levels for certain return periods of interest were calculated by means of the inverse of Equation 3.1, excluding the scaling factors:

$$MM_T = \begin{cases} \mu - \left(\frac{\sigma}{\xi}[1 - (-\ln[1 - \frac{1}{T*12}])]^{-\xi}\right) & \xi \neq 0 \\ \mu - \sigma \ln(-\ln[1 - \frac{1}{T*12}]) & \xi = 0 \end{cases}, \tag{3.3}$$

where T is the return period in years. In this paper, the surge MM water levels are shown for a return period of $\frac{1}{6}$ years (this approximates the average MM) and for 10 years; they will be referred to as $MM_{\frac{1}{6}}$ and MM_{10}, respectively (where the subscript denotes the return period). Since we considered a time-dependent GEVD, the value for μ and σ was first computed based on the fitted α- and β-parameters and t. This allowed for a reconstruction of a time-varying series of MM_T -values from which the surge characteristics of interest were derived. In this paper, the following characteristics are studied: time-averaged MM_T, seasonal range (difference between maximum and minimum MM_T in the year, excluding the secular change), seasonal phase (month with highest MM_T) and the secular change (linear trend in MM_T).

3.4. Results

3.4.1. Data Availability and Scaling Factors

The number of data from satellite altimeters that were available per grid cell roughly equalled 2.5×10^5 from 40° S to 40° N, but increased to 3×10^5 at 60° N/S, for the coarse (5° × 5°) grid. Using the fine (2° × 2°) grid, reduced the data availability by about 83%. In addition, as mentioned before, the sea ice-affected regions experience seasonal differences in data availability. This likely affected the derived storm surge properties. For this reason, all figures showing satellite data-based results will also show the maximum sea ice extent during the TPJ-period (derived from NSIDC (2022)), as a reminder to treat these specific results with caution.

The obtained optimal scaling factors for the coarse grid (5° × 5°) are shown in Figures 3.1a and 3.1d. Overall, the location parameter (μ) required more scaling in the tropics (i.e., scaling factor k_1 deviates more from 1), compared to higher latitudes. This relation is less obvious for scaling factor k_2. Moreover, the average SE for k_1 and the average RMSE of the obtained GEVD showed a clear latitudinal dependency, with larger errors at higher latitudes. In addition, as a result of the lower data availability, the analysis on the fine grid required more scaling compared to the coarse grid (Figure 3.1b, e). Nonetheless, the analysis on the fine grid resulted in significantly larger errors than the analysis on the coarse grid (Figure 3.1c). Based on these results, it was decided to use the coarse grid for the remainder of the study.

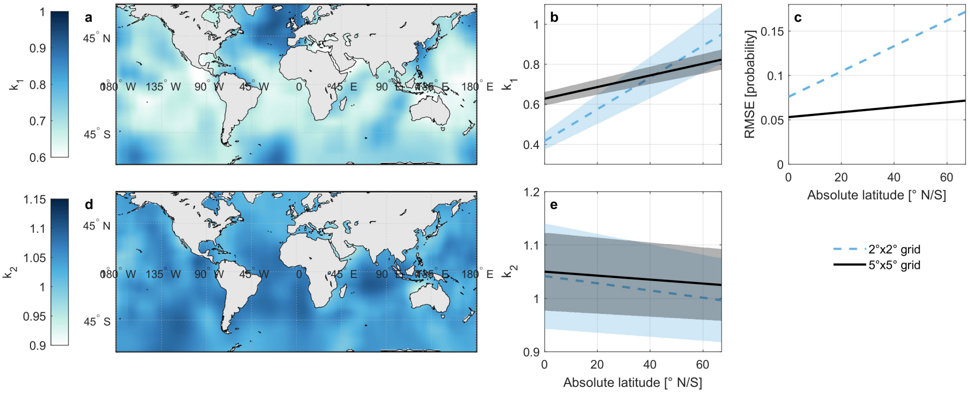


Figure 3.1: Global estimates of the scaling factors k_1 (a) and k_2 (d) for the analysis on the coarse grid (5° × 5°). Note the difference in color scales used for the two subfigures. The average scaling factors k_1 (b) and k_2 (e) including their standard error (patched area) are shown as a function of absolute latitude (linear fit), as well as the RMSE between the GEVD based on the original high-frequency and satellite-sampled data (c), for both the fine (2° × 2°) and coarse grid.

The main findings from the comparison with scaling factors computed from tide gauge data are presented in Figure 3.2. For the majority of the locations, the differences between the obtained scaling factors were within the 95%-confidence intervals of these differences and were thus considered to be insignificant (particularly applied to k_2, Figure 3.2d). For k_1, the significant differences mainly concerned higher scaling factors derived from the tide gauges compared those derived from GTSM, except for some locations in the Pacific Ocean (Figure 3.2a, c). This implies that the derived surge water levels would be scaled up more when considering the GTSM-based product (as a lower scaling factor corresponds to increased scaling). However, for k_2, all significant differences concerned lower scaling factors derived from the tide gauges compared to GTSM (Figure 3.2d, f). For the majority of the locations, better scaling results were obtained in the analysis of reanalysis data: that is, the RMSE between the GEVD based on the original high-frequency and satellite-sampled data was lower. This particularly applies to the locations where the tide gauge-derived k_1 exceeded the GTSM-derived k_1 (Figure 3.2b).

Figure 3.3 visualizes the consequences of using the coarse grid as opposed to the fine grid by comparing the obtained time-averaged $MM_{\frac{1}{6}}$ and MM_{10} to information derived from tide gauges. The message is two-fold: 1) using the coarse grid resulted in more spatial data gaps when a significant portion of a grid cell covered land and was omitted, and 2) using the coarse grids resulted in more averaging of the surge water levels. This generally resulted in lower values compared to tide gauges as the ESLs tended to increase towards the coast. The latter can be clearly seen in the North Sea. However, the opposite effect was observed in the Skagerrak (the water between Norway and Denmark) where the fine grid resulted in lower values than the coarse grid.

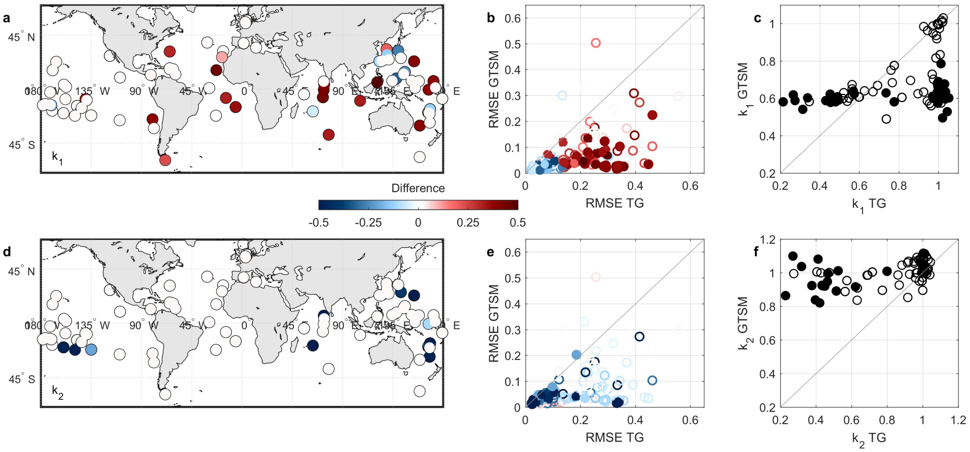


Figure 3.2: Difference between the scaling factors k_1 (a) and k_2 (d) obtained from tide gauge (TG) and reanalysis (GTSM) data for the secular analysis on the coarse grid (5° × 5°). Positive values (red) indicate the tide gauge-derived factor is larger. Insignificant differences (i.e. that do not exceed the 95% confidence interval of the difference) have been coloured white. Difference between the scaling factors k_1 (b) and k_2 (e) as a function of the RMSE between the GEVD based on the original high-frequency and satellite-sampled data. Absolute values for the scaling factor k_1 (c) and k_2 (f) derived from tide gauge and reanalysis data (GTSM). Filled scatters indicate significant differences between the two scaling factors.

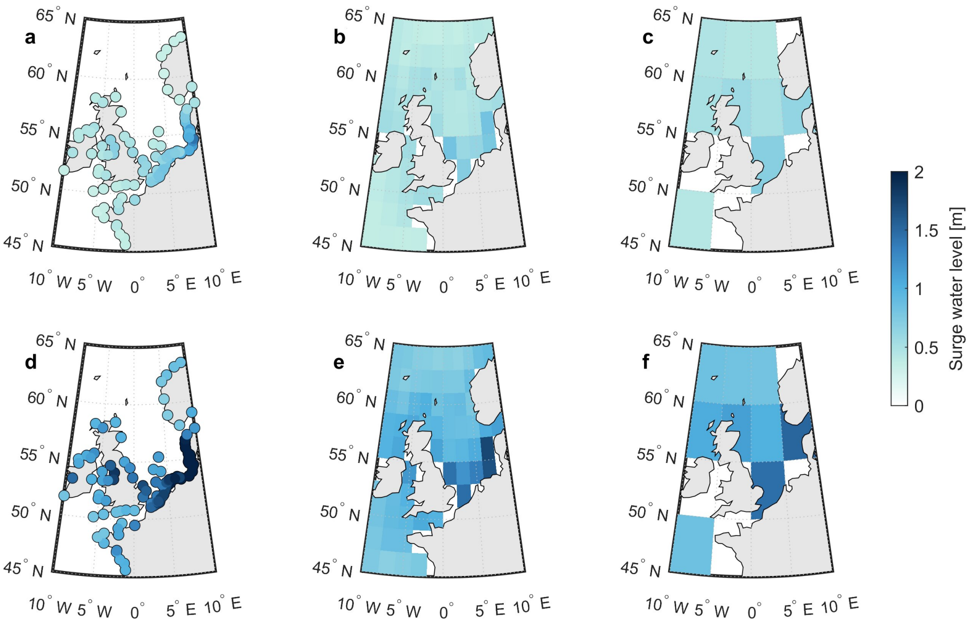


Figure 3.3: Time-averaged $MM_{\frac{1}{6}}$ (a, b, c) and MM_{10} (d, e, f) inferred from tide gauge data (a, d), the satellite data stacked on the fine grid (2° × 2° ; b, e) and the satellite data stacked on the coarse grid (5° × 5° ; c, f).

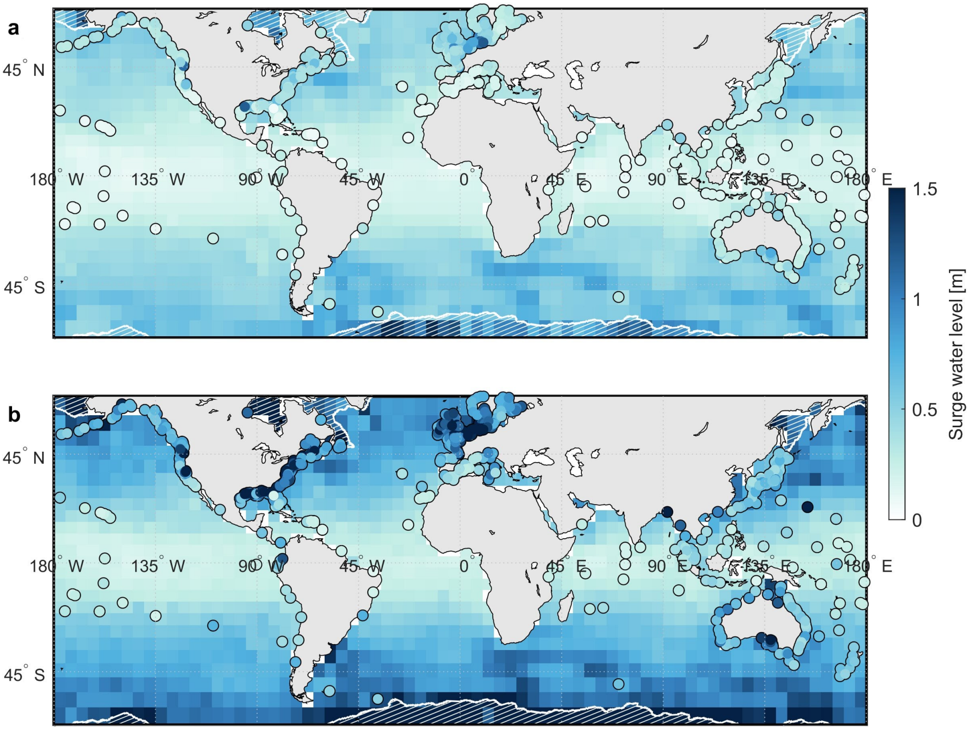


Figure 3.4: Time-averaged $MM_{\frac{1}{6}}$ (a) and MM_{10} (b) inferred from the satellite data (background) and tide gauge data (scatters). The hatched regions indicate the maximum sea ice extent during the TPJ-period.

3.4.2. TIME-AVERAGED SURGE WATER LEVELS

The time-averaged $MM_{\frac{1}{6}}$ and MM_{10} are visualized in Figure 3.4a and 3.4b, respectively. Both quantities showed a strong zonal dependency with lower values in the tropics (about 0.25 m for both return periods) and larger values at mid and high latitudes (about 0.6 and 1.2 m, respectively). This pattern was observed in both the satellite-derived and the tide gauge-derived product. More enhanced surge water levels compared to their surroundings were observed in the northwest Atlantic Ocean, the northwest Pacific Ocean, and the Southern Ocean. In several instances, the surge water level at one tide gauge exceeded that of surrounding tide gauges and the nearby satellite-derived product (e.g., in the Gulf of Mexico and the North Sea). In addition, higher surge water levels were observed in the sea-ice affected regions, particularly in the Southern Ocean. The correlation between the tide gauge-derived product and the satellite-derived product (interpolated to the locations of the tide gauges) was computed as 0.63 for $MM_{\frac{1}{6}}$, and 0.67 for MM_{10}.

3.4.3. SEASONALITY SURGE WATER LEVELS

As shown in Figures 3.5a and 3.5b, storm surges water levels on the northern hemisphere were subject to significant seasonal variation (up to 0.4 m for $MM_{\frac{1}{6}}$ and 0.75 m for MM_{10}),

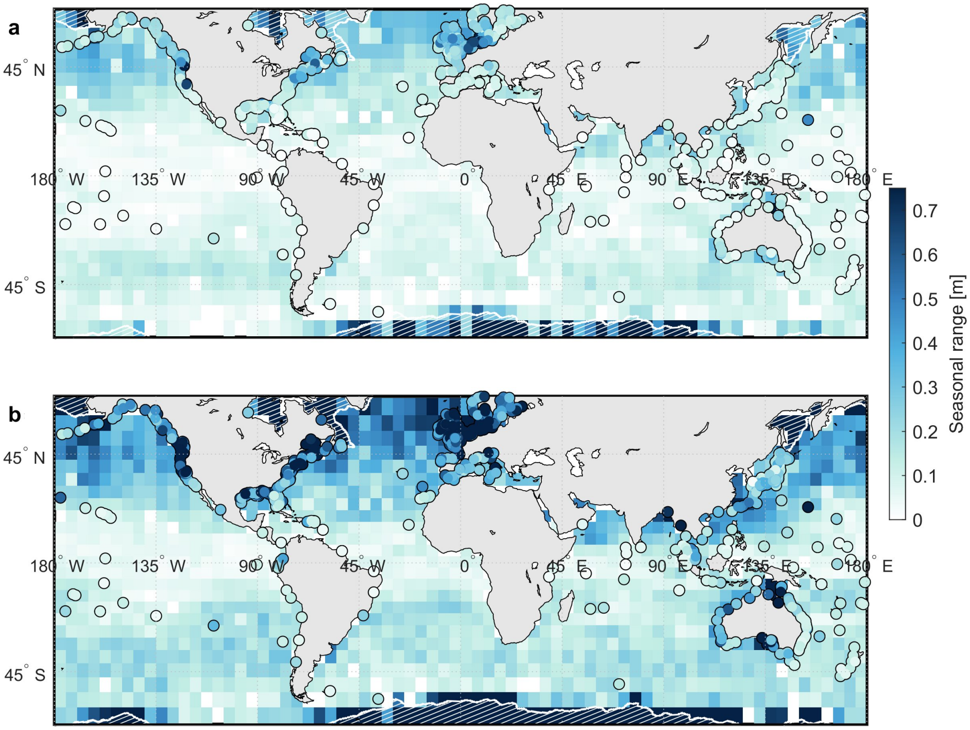


Figure 3.5: Range of seasonal variation in $MM_{\frac{1}{6}}$ (a) and MM_{10} (b) inferred from the satellite data (background) and tide gauge data (scatters). The hatched regions indicate the maximum sea ice extent during the TPJ-period.

while this was much less for the storm surges on the southern hemisphere (up to 0.2 m for $MM_{\frac{1}{6}}$ and 0.4 m for MM_{10}). Here, the sea-ice affected region in the Southern Ocean appeared to be a clear outlier, with a seasonal range in both $MM_{\frac{1}{6}}$ and MM_{10} of 0.75 m or more.

Overall, the seasonal phase (the month in which the storm surge was the highest) was found to correspond to the local winter (Figures 3.6). Exceptions were predominantly observed in the tropical zone, e.g., the Japanese and Philippine Seas (August) and the Arabic Sea (July). Another striking region was the zonal band in the Pacific Ocean around 15° N where the $MM_{\frac{1}{6}}$ was largest in March/April. Moreover, while the overall pattern in seasonal phase is comparable for $MM_{\frac{1}{6}}$ and MM_{10}, that of MM_{10} showed more variability in the tropics. Finally, the seasonal phase of the MM in the sea-ice affected regions corresponds to the month of maximum sea ice extent (March for the northern hemisphere and September for the southern hemisphere).

The computed correlation coefficient between the satellite-derived product and that from tide gauges was 0.57 (0.73) for the seasonal range of $MM_{\frac{1}{6}}$ (MM_{10}) and 0.53 (0.50) for

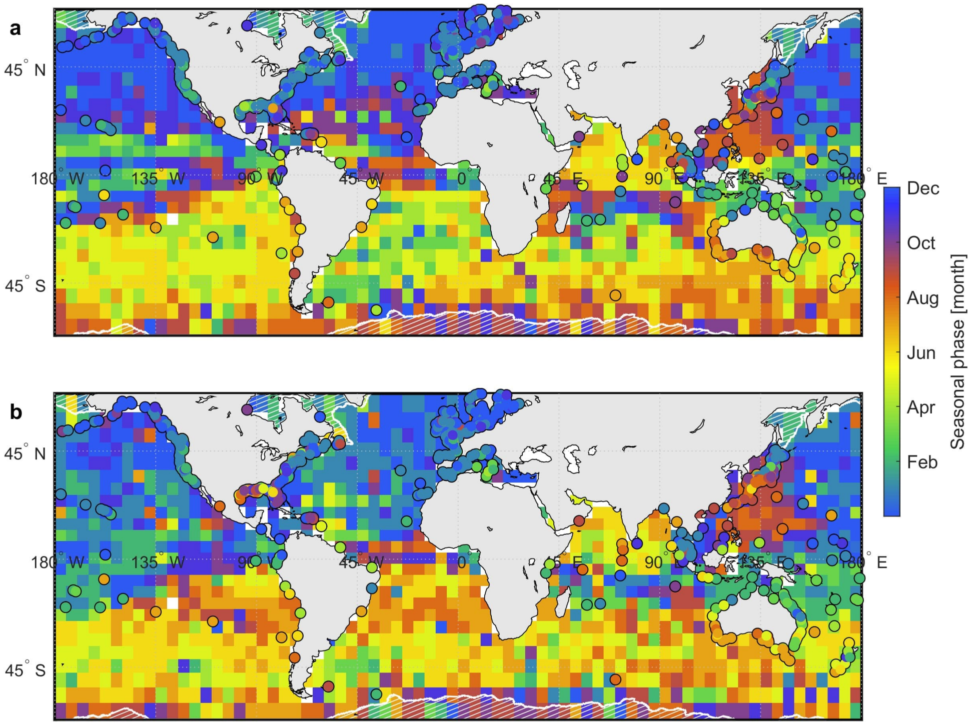


Figure 3.6: Phase (month of maximum) of seasonal variation in $MM_{\frac{1}{6}}$ (a) and MM_{10} (b) inferred from the satellite data (background) and tide gauge data (scatters). Locations with insignificant seasonality have been left white. The hatched regions indicate the maximum sea ice extent during the TPJ-period.

the seasonal phases. Differences were predominantly observed along the west coast of the USA (seasonal range of MM_{10}), in the Arabic Sea (seasonal phase) and Gulf of Mexico (all variables).

3.4.4. SECULAR CHANGE SURGE WATER LEVELS

Global changes in the $MM_{\frac{1}{6}}$ of ~ 0.25 cm/year were derived across the globe (Figure 3.7). The trend was predominantly negative, while several mid latitude regions showed an increase instead. In contrast to the variables addressed above, there was a poor agreement between tide gauge- and satellite-derived yearly trends: the computed correlation coefficient was only 0.11. An example of a region where there is a consistent mismatch between the tide gauge- and satellite-derived change in $MM_{\frac{1}{6}}$ is the east coast of the USA. As this region frequently experiences tropical storms (see Figure 3.8), the correlation coefficient was again computed considering only the tide gauges in areas that are not affected by tropical storms. This resulted in a higher correlation of 0.28. The mismatch appeared to be partly related to the high local variability in the tide gauge-derived change or isolated tide gauges (e.g., the tide gauges in the western Pacific that show pos-

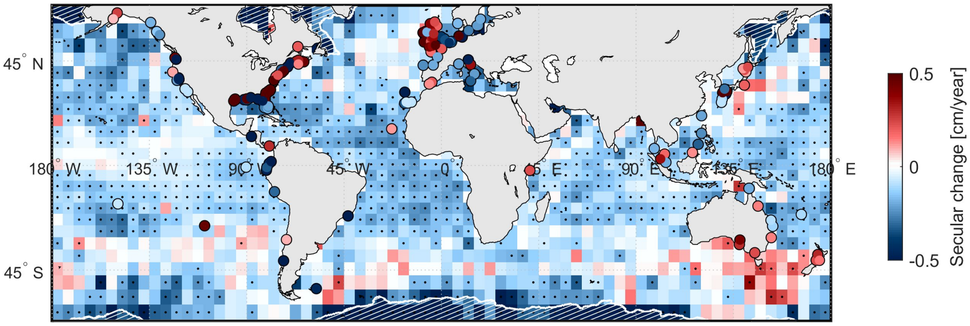


Figure 3.7: Yearly trend in $MM_{\frac{1}{6}}$ inferred from the satellite data (background) and tide gauge data (scatters). Only satellite-derived trends that exceed the 95% confidence interval have been dotted. Tide gauge-derived trends that were insignificant at the 95% confidence level were excluded from the plot. The hatched regions indicate the maximum sea ice extent during the TPJ-period.

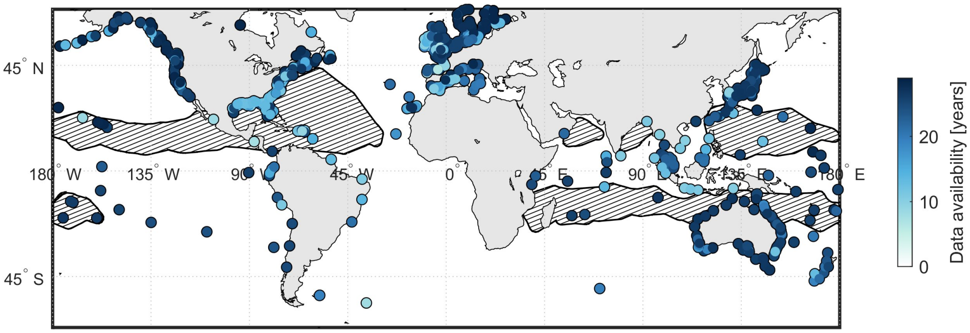


Figure 3.8: Number of years with data for tide gauges used in the study. The hatched regions indicate the areas that are affected by tropical storms with wind speeds exceeding 100 knots (derived from Knapp et al. (2022)).

itive trends). In fact, filtering out spatial outliers from the trends derived at tide gauges [1], increased the correlation up to 0.35 for a radius of 200 – 220 km. The correlation reduced again for larger distances. This relation was also observed for the time-averaged MM_T and seasonality variables, albeit with some differences in the optimal distance. Finally, large negative trends (exceeding 0.5 cm/year) were observed in the sea-ice affected regions.

3.5. DISCUSSION

For the first time, the vast amount of data available from satellite radar altimeters has been used to map the temporal variability in global storm surge water levels. This has been done by fitting a time-dependent generalized extreme value distribution (GEVD) to the monthly maximum (MM) surge water levels. However, in contrast to the in-situ data that have been widely used to study storm surges, the data from satellite altimeters

---

[1] defined as locations where the trend deviates more than two times the MAD from the median trend computed over all tide gauges within a certain distance

have a rather coarse temporal resolution. To ensure sufficient coverage of storm events, the measurements from eight different missions were combined and stacked over areas of 5° × 5°. In addition, an experiment based on reanalysis data was performed to obtain location specific scaling factors that correct for possible bias in storm surge water levels obtained by the satellites. Subsequently, the time-average storm surge water levels and their seasonal variation and secular change were derived and compared to the same information derived from high-frequency tide gauge data.

The main findings presented in this paper are as follows. The time-average $MM_{\frac{1}{6}}$ and MM_{10} showed a clear zonal pattern, with higher water levels observed at mid and high latitudes compared to the equatorial region. The largest storm surge water levels were observed in the northwest Atlantic Ocean, northwest Pacific Ocean, Southern Ocean and North Sea. These regions correspond to regions with significant mean sea level variability (e.g., Dangendorf et al., 2014; Rhines, 2001). As the short-term sea level variability has not been removed from the 'surge' water levels, it is unknown to what extent this affects the larger magnitude in aforementioned regions. However, these regions do not stand out when considering the temporal variability in storm surge water levels. The seasonal variation in the storm surge water levels was particularly large at mid to high latitudes in the northern hemisphere. The largest seasonality was observed in the sea-ice affected regions, although there are reasons to question that finding (more will follow). While the month in which the highest surge water levels were observed generally corresponds to the local winter, in particular in the tropics, there were some exceptions. Most likely, this is related to the nature of the storms that are captured. Where extra-tropical storms typically occur in winter, tropical storms (i.e., hurricanes, cyclones and typhoons) have a different seasonality (e.g., peak activity in September for the North Atlantic, in August for Japan (Camp et al., 2015)). Lastly, a predominant negative trend in storm surge water levels was observed across the domain, except for certain regions at mid latitudes.

The results derived from the satellite data have been compared to a similar product derived from tide gauges. This comparison can be used to assess the efficacy of satellite data for storm surge analysis. However, it should be kept in mind that one cannot expect a perfect agreement between the products derived from both data sources. Firstly, the tide gauge data are point estimates, while the satellite product is a spatial average. As was seen in Figure 3.3, the obtained quantities closely depend on the size of the area that was used for stacking of the satellite data. In addition, not all tide gauge data cover the full TPJ-period, nor are they gapless. Consequently, given the high interannual variability in storminess (IPCC, 2021; Weisse et al., 2014), the secular changes derived from the tide gauges may not be representative for the full TPJ-period.

Nevertheless, reasonable agreement between the tide gauge and satellite-derived product was observed for the time-average and seasonal properties. However, larger disparities were observed for the secular change. As mentioned, this may be related to the data availability at the tide gauges, although this mainly affects the tide gauges in south east Asia and the Gulf of Mexico (Figure 3.8), while discrepancies were also observed elsewhere. Hence, two alternative reasons are suggested. Firstly, the secular results appeared to be more dominated by local variability and it is plausible that the change in storminess at the coast of one island in e.g., the Pacific Ocean, is not representative for the 5° × 5° area that this island is located in. This is supported by the fact that excluding spatial outliers from the trends derived at tide gauges resulted in a better correlation. On the other hand, a positive trend in $MM_{\frac{1}{6}}$ was observed at a number of tide gauges along the east coast of the USA, yet not captured by the satellite-derived product. This leads to the second hypothesis concerning the difference between tropical and extra-tropical storms. Tropical storms are typically shorter and smaller than extra-tropical storms (Von Storch and Woth, 2008) and therefore more likely to be missed by the satellite data. This is supported by the fact that the seasonal phase of the $MM_{\frac{1}{6}}$ and MM_{10} did not everywhere correspond to the local season of tropical storms (Camp et al., 2015), even though tropical storms surges typically have a larger amplitude than extra-tropical storms (Von Storch and Woth, 2008). The discrepancies in observed trends along the east coast of the USA suggest the positive trend may reflect changes in the tropical storms as opposed to a background negative trend in extra-tropical storms. However, the regional increases in storm surge water levels around 45° N/S are consistent with the observed poleward shift in tropical storms (IPCC, 2021).

Another apparent limitation of the satellite data is related to the sea ice-affected regions. All studied storm surge properties show rather extreme and unrealistic values in the regions that are temporally covered by sea ice. In particular, the derived seasonal phase and secular change suggest that the results in these regions are affected by contamination from reflections of the sea ice rather than the actual water level. Namely, the seasonal phase corresponds to the period of maximal sea ice extent (September for the Antarctic and March for the Arctic (NSIDC, 2022)) and in this season one would not expect significant surge as the water surface can only to a limited extent be disturbed by the wind. If, however, the data is dominated by reflections from the sea ice, a higher "water" level is indeed expected during months with maximal sea ice extent. In addition, the negative trend in surge water levels in the regions suggests contamination by sea ice reflections. As the sea ice extent reduces over time, the contamination also reduces, resulting in a reduction in extreme "sea" level measurements. In fact, the sea-ice affected regions would be an interesting study area, precisely because of the significant changes to the sea ice extent. However, for such a study, a more explicit selection of the radar returns is required.

Furthermore, the method of obtaining scaling factors using reanalysis data has been assessed by comparison to a similar analysis using data from a selected set of tide gauges (Figure 3.2). Overall, a larger spread was observed in the scaling factors derived from tide gauge data compared to the reanalysis data. In most locations, the differences between the factors derived from the reanalysis and the tide gauge data appeared insignificant. For k_2, only few significant differences were found, predominantly in regions that are affected by tropical storms (see Figure 3.8). In all of these cases, the factors derived from the reanalysis data exceed those derived from the tide gauge data. This suggests that the obtained surge water levels may be scaled too little. On the other hand, in several instances, significantly lower values (thus more scaling) were computed from the reanalysis data for scaling factor k_1. Although, as this predominantly concerns locations where the scaling is significantly less successful in the experiment with tide gauge data compared to that with reanalysis data (higher RMSE, Figure 3.2b), this questions the reliability of the tide gauge-derived scaling factors.

Finally, a comparison of the results presented here to findings by earlier studies is unfortunately not straightforward, for various reasons. First of all, many studies on changes in extreme sea level events (ESLs) do not exclude the impact of variability in the mean sea level (MSL) (e.g., Wahl et al., 2017; Weisse et al., 2014). On the other hand, the study by Menéndez and Woodworth (2010) (and a follow-up by Marcos and Woodworth (2018)) specifically attempted to distinguish the impact of the MSL on the temporal variability in ESLs derived from tide gauge records. While by including the MSL an overall increase in ESLs was observed across the globe, the observed change was predominantly negative when the MSL was excluded. Although this corresponds to findings presented in our study, some contradictions were observed for individual tide gauges. Consequently, a second issue that complicates intercomparisons of results was identified: the timespan of the involved datasets. Since the study by Menéndez and Woodworth (2010) was mainly based on tide gauge data from 1970 up to 2010, it is unknown to what extent any differences in results are related to the highly variable character of storm surges on interannual time scales. On another note, the seasonal pattern found in their study, did largely correspond to what was presented in this paper. Finally, while the presented analysis provides a range of information on storm surge water levels and their temporal variability, it does not directly allow for any conclusions on changes in the frequency of storm surges. In that sense it may be of interest to extend the analysis by a type of peak-over-threshold analysis (as was done in Andersen et al. (2015), Antony et al. (2014), and Ji et al. (2019)).

3.6. CONCLUSIONS

The data from eight satellite radar altimeters (1993-2021) have been used to study the temporal and spatial variability in global storm surge water levels. The time-averaged surge water levels were dominated by a zonal pattern, with higher water levels at mid and high latitudes compared to the equatorial region. The highest water levels were observed in the Southern Ocean, northwest Pacific Ocean, northwest Atlantic Ocean and North Sea. If was found that the surge water levels on the northern hemisphere were subject to significant seasonal variability. Overall, the maximum storm surge water levels mainly occurred in the local winter months, that is, except for the tropics where the seasonal phase showed more local variability. Finally, moderate secular changes of ~ 0.25 cm/year were derived for the $MM_{\frac{1}{6}}$. The derived changes were predominantly negative, except for a few mid latitude regions with positive change.

Except for the secular changes, the satellite-derived results were comparable to the information derived from tide gauges, although the tide gauges showed more local variability. The poor correlation for the secular change may be related to the change in water levels being dominated by either changes to tropical or extra-tropical surges. It has been suggested that the satellites may not be able to fully capture the temporal variability in the short-lived tropical storms.

Nevertheless, the results provide valuable information on the spatial and temporal variability in storm surge water levels, thereby bridging the distance between the many, yet clustered tide gauges.

BIBLIOGRAPHY

Adebisi, N., Balogun, A. L., Min, T. H., & Tella, A. (2021). Advances in estimating Sea Level Rise: A review of tide gauge, satellite altimetry and spatial data science approaches. *Ocean & Coastal Management, 208*.

Andersen, O. B., Cheng, Y., Deng, X., Steward, M., & Gharineiat, Z. (2015). Using satellite altimetry and tide gauges for storm surge warning. *Proceedings of the International Association of Hydrological Sciences, 365*, 28–34.

Antony, C., Testut, L., & Unnikrishnan, A. S. (2014). Observing storm surges in the Bay of Bengal from satellite altimetry. *Estuarine, Coastal and Shelf Science, 151*, 131–140.

Beckley, B., Ray, R. D., Zelensky, N., Lemoine, F., Brown, S., Desai, S., & Mitchum, G. (2021). Integrated Multi-Mission Ocean Altimeter Data for Climate Research TOPEX/Poseidon, Jason-1, 2, & 3 User's Handbook Version 5.1.

Bij de Vaate, I., Slobbe, D. C., & Verlaan, M. (2023). Mapping the spatiotemporal variability in global storm surge water levels using satellite radar altimetry. *Under review at Ocean Dynamics.*

Bij de Vaate, I., Vasulkar, A. N., Slobbe, D. C., & Verlaan, M. (2021). The Influence of Arctic Landfast Ice on Seasonal Modulation of the M2 Tide. *Journal of Geophysical Research: Oceans, 126*(5), e2020JC016630. https://doi.org/https://doi.org/10.1029/2020JC016630

Butler, A., Heffernan, J. E., Tawn, J. A., Flather, R. A., & Horsburgh, K. J. (2007). Extreme value analysis of decadal variations in storm surge elevations. *Journal of Marine Systems, 67*(1-2), 189–200.

Calafat, F. M., Wahl, T., Tadesse, M. G., & Sparrow, S. N. (2022). Trends in Europe storm surge extremes match the rate of sea-level rise. *Nature, 603*(7903), 841–845.

Camp, J., Roberts, M., MacLachlan, C., Wallace, E., Hermanson, L., Brookshaw, A., ..., & Scaife, A. A. (2015). Seasonal forecasting of tropical storms using the Met Office GloSea5 seasonal forecast system. *Quarterly Journal of the Royal Meteorological Society, 141*(691), 2206–2219.

Cancet, M., Andersen, O. B., Cotton, D., Lyard, F., & Benveniste, J. (2015). High resolution tidal modeling in the Arctic Ocean : needs and upcoming developments. *17*, 7231.

Cartwright, D. E., & Edden, A. C. (1973). Corrected tables of tidal harmonics. *Geophys. J. Roy. Astron. Soc., 33*, 253–264. https://doi.org/doi:10.1111/j.1365-246X.1973.tb03420.x

Cartwright, D. E., & Taylor, R. J. (1971). New computations of the tide-generating potential. *Geophys. J. Roy. Astron. Soc., 23*, 45–74. https://doi.org/doi:10.1111/j.1365-246X.1971.tb01803.x

Codiga, D. L. (2020). UTide Unified Tidal Analysis and Prediction Functions. MATLAB Central File Exchange. [Software] [Last checked on 10/06/2020].

Dangendorf, S., Calafat, F. M., Arns, A., Wahl, T., Haigh, I. D., & Jensen, J. (2014). Mean sea level variability in the North Sea: Processes and implications. *Journal of Geophysical Research: Oceans*, *119*(10).

Feng, J., Li, D., Wang, T., Liu, Q., Deng, L., & Zhao, L. (2019). Acceleration of the Extreme Sea Level Rise Along the Chinese Coast. *Earth and Space Science*, *6*(10), 1942–1956. https://doi.org/doi:10.1029/2019ea000653

Gori, A., Lin, N., Xi, D., & Emanuel, K. (2022). Tropical cyclone climatology change greatly exacerbates us extreme rainfall–surge hazard. *Nature Climate Change*, *12*(2), 171–178.

Haigh, I. D., MacPherson, L. R., Mason, M. S., Wijeratne, E. M. S., Pattiaratchi, C. B., Crompton, R. P., & George, S. (2014). Estimating present day extreme water level exceedance probabilities around the coastline of Australia: tropical cyclone-induced storm surges. *Climate Dynamics*, *42*(1), 139–157.

Haigh, I. D., Marcos, M., Talke, S. A., Woodworth, P. L., Hunter, J. R., Haugh, B. S., ..., & Thompson, P. (2022). GESLA Version 3: A major update to the global higher-frequency sea-level dataset. [Dataset]. *Geoscience Data Journal*. https://doi.org/https://doi.org/10.1002/gdj3.174

IPCC. (2021). *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]* Cambridge University Press.

Izaguirre, C., Méndez, F. J., Menéndez, M., & Losada, I. J. (2011). Global extreme wave height variability based on satellite data. *Geophysical Research Letters*, *38*(10).

Ji, T., Li, G., & Zhang, Y. (2019). Observing storm surges in China's coastal areas by integrating multi-source satellite altimeters. *Estuarine, Coastal and Shelf Science*, *225*.

Knapp, K. R., Diamond, H. J., Kossin, J. P., Kruk, M. C., & Schreck, C. J. (2022). International Best Track Archive for Climate Stewardship (IBTrACS) Project, Version 4. [Dataset]. *NOAA National Centers for Environmental Information*. https://doi.org/https://doi.org/10.25921/82ty-9e16

Lobeto, H., Menendez, M., & Losada, I. J. (2018). Toward a methodology for estimating coastal extreme sea levels from satellite altimetry. *Journal of Geophysical Research: Oceans*, *123*(11), 8284–8298.

Marcos, M., & Woodworth, P. L. (2018). Changes in extreme sea levels. *CLIVAR Exchanges*.

McGranahan, G., Balk, D., & Anderson, B. (2007). The rising tide: assessing the risks of climate change and human settlements in low elevation coastal zones. *Environment and urbanization*, *19*(1), 17–37.

Méndez, F. J., Menéndez, M., Luceño, A., & Losada, I. J. (2007). Analyzing monthly extreme sea levels with a time-dependent gev model. *Journal of Atmospheric and Oceanic Technology*, *24*(5), 894–911.

Menéndez, M., & Woodworth, P. L. (2010). Changes in extreme high water levels based on a quasi-global tide-gauge data set. *Journal of Geophysical Research: Oceans*, *115*(C10).

Muis, S., Verlaan, M., Winsemius, H. C., Aerts, J. C., & Ward, P. J. (2016). A global reanalysis of storm surges and extreme sea levels. *Nature Communications*, *7*(1), 1–12.

Naeije, M. C. (2022). Radar Altimeter Database System. [Data archive]. http://rads.tudelft.nl

NSIDC. (2022). Sea Ice Index, Version 3 (G02135) [Dataset]. *NOAA National Centers for Environmental Information*. https://nsidc.org/data/g02135/versions/3#anchor-2

Oppenheimer, M., Glavovic, B., Hinkel, J., van de Wal, R., Magnan, A. K., Abd-Elgawad, A., ..., & Sebesvari, Z. (2019). Sea level rise and implications for low lying islands, coasts and communities. *In: IPCC Special Report on the Ocean and Cryosphere in a Changing Climate [Pörtner, H.-O., D.C. Roberts, V. Masson- Delmotte, P. Zhai, M. Tignor, E. Poloczanska, K. Mintenbeck, M. Nicolai, A. Okem, J. Petzold, B. Rama, and N. Weyer (eds.)]*, 321–445.

Rashid, M., & Wahl, T. (2020). Predictability of Extreme Sea Level Variations Along the U.S. Coastline. *Journal of Geophysical Research: Oceans*, *125*(9). https://doi.org/doi:10.1029/2020jc016295

Resio, D. T., & Westering, J. J. (2008). Modeling the physics of storm surges. *Physics Today*, *61*, 33–38.

Rhines, P. (2001). Mesoscale Eddies. In J. H. Steele (Ed.), *Encyclopedia of Ocean Sciences* (pp. 1717–1730). Academic Press. https://doi.org/https://doi.org/10.1006/rwos.2001.0143

Scharroo, R., Leuliette, E. W., Naeije, M. C., Martin-Puig, C., & Pires, N. (2016). RADS Version 4: An efficient way to analyse the multi-mission altimeter database. [ESA Special Publication SP-740]. *Proceedings of the ESA Living Planet Symposium*, 9–13.

Tran, N., Dibarboure, G., Picot, N., & Féménias, P. (2018). Improving the continuity of the Jason SSB time-series. *In OSTST meeting, September*.

Von Storch, H., & Woth, K. (2008). Storm surges: perspectives and options. *Sustainability Science*, *3*(1), 33–43.

Wahl, T., Haigh, I. D., Nicholls, R. J., Arns, A., Dangendorf, S., Hinkel, J., & Slangen, A. B. (2017). Understanding extreme sea levels for broad-scale coastal impact and adaptation analysis. *Nature Communications*, *8*(1), 1–12.

Wahr, J. M. (1985). Deformation induced by polar motion. *Journal of Geophysical Research*, *90*(B11), 9363–9368. https://doi.org/doi:10.1029/JB090iB11p09363

Wang, X., Verlaan, M., Apecechea, M. I., & Lin, H. X. (2021). Computation-efficient Parameter Estimation for a High-Resolution Global Tide and Surge Model (GTSM). *Journal of Geophysical Research: Oceans*, *126*, e2020JC016917. https://doi.org/https://doi.org/10.1029/2020JC016917

Weisse, R., Bellafiore, D., Menéndez, M., Méndez, F., Nicholls, R. J., Umgiesser, G., & Willems, P. (2014). Changing extreme sea levels along European coasts. *Coastal engineering*, *87*(1), 1–14.

4

SEA ICE AND THE SEASONAL MODULATION OF THE ARCTIC TIDE

Seasonal modulation of the M_2 *tide has been quantified for the entire Arctic Ocean and connected regional seas, using tidal harmonic analysis of water levels derived from Synthetic Aperture Radar (SAR) altimetry. Results are compared to numerical simulations that model the effect of two limiting cases of seasonal landfast ice cover on the* M_2 *tide. The largest seasonal modulation (up to 0.25 m) is observed along coastlines and in bays. Locally, the presence of landfast ice decreases amplitudes, but in some cases, the opposite effect was observed further afield. In most of the Arctic, winter months experience a later arrival of the tide, except for Hudson Bay where phase advance is observed. Most of the altimeter-derived seasonal modulation could be explained by the modeled impact of landfast ice. However, particularly in the Hudson Bay system there is a discrepancy between model- and altimeter-derived seasonal modulation. This suggests that other seasonal processes are important. Finally, results suggest that the consequences of variations in Arctic landfast ice are not restricted to the Arctic, but affect tidal water levels on a global scale.*

This chapter has been published as Bij de Vaate, I., Vasulkar, A. N., Slobbe, D. C., & Verlaan, M. (2021). The Influence of Arctic Landfast Ice on Seasonal Modulation of the M2 Tide. *Journal of Geophysical Research: Oceans, 126*(5), e2020JC016630. https://doi.org/https://doi.org/10.1029/2020JC016630

4.1. INTRODUCTION

Arctic sea ice is subject to large seasonal variations in extent, thickness, and mobility. Most of the Arctic region is covered by sea ice in winter months, with a total area of ~15 million km^2; however, this area declines by a factor of three, to ~4.5 million km^2, by summer (average 2011-2019; (National Snow and Ice Data Center, 2020)). Moreover, in winter there is about 1.65 million km^2 of landfast sea ice, that reduces to near zero in summer (Li et al., 2020). The presence of landfast ice can significantly affect local hydrodynamics as it prevents interaction between the atmosphere and underlying ocean (Mahoney et al., 2014).

While an ice cover shields the ocean from atmospheric forcing, it also exerts additional frictional stress on the surface. Early studies demonstrate that drag between water and ice results in tidal dampening, especially in coastal zones where ice is relatively immobile (Godin, 1986; Kowalik, 1981). Kowalik (1981) showed that the presence of sea ice can lead to tidal amplitude decay and phase delay. More recent studies have looked specifically into the modulation of tides in the Arctic region in response to a seasonally varying ice cover and suggest the effect to be substantial; in some regions changing the amplitude by up to 0.15 m (Kagan and Sofina, 2010; St-Laurent et al., 2008). This implies that it is insufficient to regard tides as constant throughout the year and ignore the influence of sea ice, which is done in most operational tide models. For accurate prediction of Arctic tidal water levels, quantification of seasonal modulation is necessary. In addition, studies have shown that Arctic tides directly affect North Atlantic tides (Arbic et al., 2004; Arbic et al., 2007), implying that the influence of Arctic sea ice on tides extends beyond the Arctic.

Müller et al. (2014) modeled seasonal variations in the M_2 tide on a global scale, forced by annual differences in stratification and sea ice cover. While they incorporated altimeter-derived sea level data in their study, only in situ data were considered for the high latitudes. Other numerical studies that focused on the Arctic were often restricted to a certain region, e.g., the Central Arctic and Russian shelf (Kagan and Sofina, 2010) and the Hudson Bay system (Kleptsova and Pietrzak, 2018; St-Laurent et al., 2008). In addition, tide monitoring in the area is limited. Hence, models that link seasonal variations in sea ice to modification of Arctic tides can only be validated locally and the large-scale signals remain uncertain.

For the first time, in this paper, the seasonal modulation of the largest tidal constituent M_2 is quantified for the entire Arctic. To achieve this, we supplement data from the sparse set of Arctic tide gauges with data from Synthetic Aperture Radar (SAR) altimeters. The latter provide water level information up to 88°N, including predominantly ice-covered areas. We then model the effect of landfast ice on the Arctic M_2 tide for summer and winter landfast ice configurations, providing insight in the extent to which landfast ice is responsible for the Arctic tidal modification. Finally, we use the model to simulate the global effect of variations in Arctic landfast ice cover, which we then relate to observed seasonal modulation at tide gauges across the globe.

4.2. METHODOLOGY

4.2.1. EVALUATING SEASONAL MODULATION FROM ALTIMETRY

There are very few tide gauges in the Arctic and only some of them have data publicly available. Furthermore, the usage of altimetry to derive tidal measurements is limited at high latitudes, because this area is not covered by the radar altimeter satellites that are typically used for this purpose (e.g., TOPEX/Poseidon and Jason (Müller et al., 2014)). Added to this, the low along-track resolution of conventional low-resolution mode (LRM) altimeters does not allow for distinguishing between measurements from the sea ice pack and alternating narrow sections of open water. Therefore, this study made use of data from two high-inclination satellite altimeter missions: CryoSat-2 (CS2; Wingham et al., 2006) and Sentinel-3 (S3; Donlon et al., 2012). CS2, launched in April 2010, has a ground coverage up to 88°N/S, and S3, launched in February 2016, reaches latitudes up to 81.5°N/S. The typical cross-track spacing of both satellites is ~ 50 km at 60°N and ~ 18 km at 80°N. Both missions are equipped with a Synthetic Aperture Radar (SAR) altimeter that is active over Arctic waters, resulting in a high along-track resolution (~ 300 m). The latter is a significant advantage, as it enables obtaining water levels from leads; i.e., open-water fractures in the sea ice pack (e.g., Schulz and Naeije, 2018; Wernecke and Kaleschke, 2015). This allows for more complete and year-round information on the sea surface height, even in areas that are usually covered by sea ice. Year-round data are necessary for proper derivation of seasonal variations in tides. Data were collected for the region north of 60°N, extended by the Hudson Bay area; 50-60°N, 100-45°W (see Figure 4.1). Acquisition times of the used data range from June 2010 to August 2019 (CS2) and from December 2016 to December 2019 (S3).

SAR returns from sea ice, leads, and open water were classified according to a multi-criteria classification scheme inspired by Poisson et al. (2018) and Schulz and Naeije (2018) and Wernecke and Kaleschke (2015). The used criteria are specified in Table 4.1. After classification, we retracked the signals of water classes by fitting the SAMOSA model (Dinardo et al., 2018; Ray et al., 2015). In contrast to conventional open water retracking, we distinguished two cases. In case of open ocean returns, the epoch t_0, significant wave height (SWH) and signal amplitude P_u were estimated, while ν (inverse of the mean-square slope of the sea surface) was set to 0 and not estimated (in line with Dinardo et al., 2018). However, for leads the sea surface is relatively smooth and we fixed SWH to 0, while ν was estimated and allowed to be infinitely large.

Compared to satellites that are conventionally used for measuring tides (e.g., TOPEX/Poseidon and Jason), CS2 and S3 have a long revisit time: CS2 has a 369 day repeat cycle (with a 30-day sub-cycle) and S3 has a repeat cycle of 27 days. This results in alias periods of 112.1 days (CS2) and 157.5 days (S3) for M_2 and does not allow for along-track tidal analysis. To overcome the low temporal resolution, we stacked the data using a 3° × 1° grid (~ 100 km), following Cancet et al. (2018). Subsequently, the data for each grid cell were combined into one timeseries assigned to the data centroid of the cell. Data were omitted when less than 1000 observations were available in a cell or when more than 30 sequential days in the year do not have data in the cell.

Table 4.1: Multi-criteria classification scheme for SAR return signals, where: sigma0 is a measure for the magnitude of the return signal, PP (Pulse Peakiness) and PPloc describe the normalized power of the peak with respect to respectively the total reflected power (after Laxon, 1994) and the 10 bins surrounding the peak, Ptail describes the normalized power in the tail of the return signal and ratioIP the ratio between the power in the leading edge and the tail. Kurtosis is only considered if a signal fulfills all criteria for a class except the number of peaks.

	Open ocean	Leads	Sea ice
Sigma0	0 - 10 dB	> 25 dB	-
PP	< 0.065	> 0.25	< 0.15
PPloc	< 0.3	-	-
Ptail	20 - 45	< 0.5	-
RatioIP	-	< 0.65	-
No. of peaks	1	1	> 1
	or:		**or**:
Kurtosis	5 - 15	-	> 12

From the full timeseries (2010-2019) of each grid cell, tidal amplitudes and phases were estimated using UTide (Codiga, 2020). UTide executes harmonic analysis for a given set of frequencies similar as in T_TIDE (Pawlowicz et al., 2002), yet it is able to deal with irregular temporal sampling. The latter is a requirement for processing altimeter-derived water levels. Tidal constants were estimated for major tides; K_1, O_1, Q_1, P_1, N_2, M_2, S_2 and K_2, and the two satellite constituents of M_2; H_1 and H_2. The latter describe the annual modulation of the M_2 tide (Zijl et al., 2016). Thereafter, a yearly signal was reconstructed based on only M_2, H_1 and H_2. At this stage, the 18.61 year nodal modulations were omitted to eliminate interference with the seasonal amplitude modulation. From the reconstructed signal, yearly and monthly average amplitudes/phases and the magnitude of annual variations were computed. Monthly average values were obtained for March and September for comparisons with model simulations.

4.2.2. EVALUATING SEASONAL MODULATION FROM MODEL

MODEL DESCRIPTION

To obtain the seasonal modulation of M_2 under the influence of a seasonally varying landfast ice extent, we used the operational Global Tides and Surge Model (GTSM) (Verlaan et al., 2015). This depth-averaged barotropic model is forced by a full tidal potential and developed in Delft3D Flexible Mesh on an unstructured grid with 25 km (open ocean) to 2.5 km (coast) resolution. To prevent a singularity at the North pole (caused by the regular grid) and an unnecessary refinement near the poles, an unstructured grid was used with the North Pole being a node in itself and the grid spanning out from there (see Verlaan et al., 2015). The model uses the GEBCO2019 gridded bathymetry dataset, which involves IBCAOv3 for the Arctic bathymetry (GEBCO Compilation Group, 2019). This bathymetry is continually developing, hence, to account for any uncertainties in the bathymetry an estimation procedure is employed (Wang et al., 2021) which uses a data assimilation framework with observations for the deep ocean tides obtained from FES2014 global tide model. The global tide model performance of GTSM was evaluated in Wang et al. (2021), and an additional analysis with specific focus on the Arctic is available in the Supporting Information (Text S1). Atmospheric forcing was not considered during the simulations.

The effect of sea ice was not included in the version of GTSM described by Wang et al. (2021). However, a sea ice cover may have a strong influence on the tidal dissipation rate (Kagan et al., 2007). This dissipation results from the friction between the sea ice and the underlying ocean and thus depends on the mobility of the sea ice. We have incorporated this dissipation in the model as an additional frictional shear stress, parameterized with a quadratic drag law in line with Kagan and Sofina (2010), Müller et al. (2014) and St-Laurent et al. (2008). The formulation used for ice-ocean shear stress (τ_i) is:

$$\tau_i = \rho_w C_{d_i} ||u_i - u||(u_i - u), \tag{4.1}$$

where u is the tidal velocity vector, u_i the ice drift velocity vector, ρ_w the density of sea water (1024 $\mathrm{kg/m^3}$). C_{d_i} is the ice-water drag coefficient whose magnitude varies from 1.32-26.8$\times 10^{-3}$ (Langleben, 1982; Madsen and Bruno, 1987; Pease et al., 1983; St-Laurent et al., 2008). In our study, it was assumed to be 5.5×10^{-3} in accordance with McPhee (1980) and Hibler (1979).

Since GTSM is not coupled to an ice model, it is not possible to obtain ice velocities at every time step. This limited our capability to compute the shear stress (Equation 4.1) for non-zero ice drift velocity. However, the state of the sea ice has two bounds: ice-free surface and fixed ice (landfast ice) (Kagan and Sofina, 2010). In the latter case, the drift velocity is equal to zero ($u_i = 0$) which makes it possible to evaluate Equation 4.1.

The ice-water frictional stress was introduced only for the regions of landfast ice cover in the domain. The remaining domain was considered to have no frictional dissipation due to this stress. The regions of landfast ice cover were modeled by means of polygons outlining their extent (Cancet et al., 2018). These polygons were created using an updated version of gridded U.S. National Ice Center sea ice charts, at 10 km resolution (U.S. National Ice Center, 2009).

QUANTIFYING SEASONAL MODULATION

We estimated the effect of landfast ice on seasonal modulation of tides by comparing M_2 amplitudes and phases from two model simulations with seasonal limiting cases of fast ice cover. We followed the approach of St-Laurent et al. (2008) and used the landfast ice covers of March and September, corresponding to the annual maximum and minimum landfast ice cover respectively (U.S. National Ice Center, 2009). Here we assumed the landfast ice induced tidal modification to be strongest in case of maximum landfast ice cover, and the change in tidal constants between the two limiting cases to be monotonic. This assumption was supported by a series of simulations based on monthly landfast ice covers of 2013 (see Supporting Information: Text S2).

To assess the impact of inter-annual variations in landfast ice cover, simulations were done for 2013 and 2017. These years respectively had, respectively, the maximum and minimum fast ice cover within the time span of the considered SAR data (2010-2019) (Li et al., 2020). In both years there was only a negligible amount of landfast ice cover in September. Therefore, we considered no fast ice in the September simulations, leading to the first limiting case of ice-free surface. The March landfast ice-extents are shown in Figure 4.1.

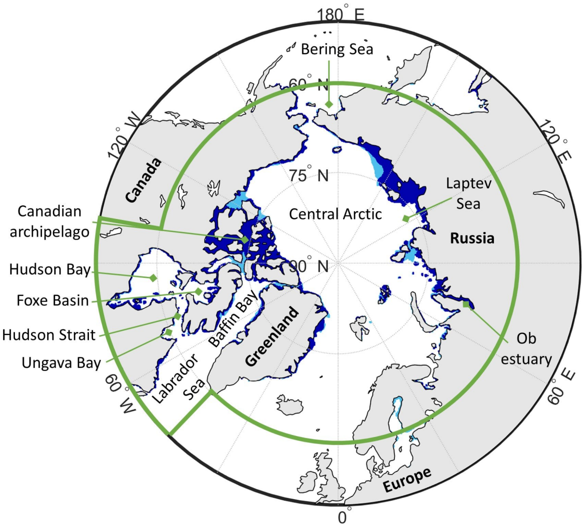


Figure 4.1: Arctic landfast ice extent for 11 March 2013 (light and dark blue) and 9 March 2017 (dark blue), obtained from updated datase of U.S. National Ice Center (2009). Note that the ice cover of 2017 completely overlaps that of 2013. The green outline indicates the domain of our study.

Simulations were done for 29.5 days, preceded by a 7-day spin up time. This period of 29.5 days corresponds to two spring-neap cycles and is sufficiently long for disentangling M_2 from the other tidal constituents involved in the analysis. Note that this period does not allow accurate separation of S_2 and K_2, and K_1 and P_1, as this requires a timespan of 182.6 days (Foreman and Henry, 1989). Although this may have a minor effect on the estimation of M_2, it will not impact our analysis since we compare amplitudes/phases between March and September. These months are half a year apart and therefore in phase with the semi-annual cycle resulting from the superposition of these constituent pairs. Hourly water levels were saved for the locations of the stacked SAR data (see 4.2.1). Subsequently, tidal constants were estimated using UTide (Codiga, 2020) for the same set of major constituents as described in section 4.2.1, but excluding H_1 and H_2. Nodal corrections were applied to eliminate amplitude differences caused by the 18.6 year nodal cycle. This allows direct comparison between estimated M_2 phases and amplitudes from the different simulations.

To compare model- and altimeter-derived seasonal modulation, root mean square errors (RMSE) and normalized root mean square errors (nRMSEs) were calculated as follows, where y represents the considered variable either derived from observations (y_{obs})

or from the model (y_{model}) and n the number of observations:

$$RMSE = \sqrt{\frac{1}{n}\Sigma_{i=1}^{n}\left(y_{obs} - y_{model}\right)^2}, \tag{4.2}$$

$$nRMSE = \frac{\sqrt{\frac{1}{n}\Sigma_{i=1}^{n}\left(y_{obs} - y_{model}\right)^2}}{\sqrt{\frac{1}{n}\Sigma_{i=1}^{n}\left(y_{obs}\right)^2}}, \tag{4.3}$$

GLOBAL EFFECT

Finally, to study the impact of seasonal variations in Arctic landfast ice on global tides, we additionally stored modeled water levels on a 1°x1° global grid (in addition to the locations of the SAR data). This was done for all simulations mentioned in the previous section; i.e. March and September, 2013 and 2017. M_2 amplitudes were estimated as described in 4.2.2.

Estimated amplitudes were then compared to March and September M_2 amplitudes derived from a global set of tide gauge records for corresponding years, obtained from the University of Hawaii Sea Level Center (Caldwell et al., 2015). Monthly amplitudes were estimated from the tide gauge records using UTide and following the same approach as was adopted for altimeter data (see 4.2.1). Tide gauge stations with large interannual variation in March-September amplitude differences (standard deviation (std) > average amplitude difference, based on 2000-2019) and those where the value of the year in question was an outlier (value > average amplitude difference ±2∗std), were excluded from the analysis.

4.3. RESULTS

4.3.1. ALTIMETER-DERIVED SEASONAL MODULATION IN THE ARCTIC

Altimeter-derived yearly average M_2 amplitudes range from near zero in the Central Arctic to over 1.5 m in Hudson Strait (Figure 4.2a). All tides in the study domain travel anticlockwise (Figure 4.2b, see Figure 4.4a for location of amphidromic points).

The total annual seasonal modulation of M_2 amplitudes ranges up to ~ 0.25 m (Figure 4.2c) with the largest ranges in Hudson Bay, Hudson Strait, Foxe Basin and along the Russian coastline (see Figure 4.1 for locations). In the Central Arctic, large annual phase modulations (up to 180°) were observed, which may be an artifact of very low amplitudes ($<$ 0.05 m, see Figure 4.2a) combined with low data availability during winter (see Figure S3), disabling proper estimation of phase variations. These data will therefore not be considered in further analysis (hatched region in Figure 4.2d; 4.5b). The largest variations in M_2 phase occur near the coasts of Russia and Canada (up to 45°).

The relative magnitude of seasonal modulation, i.e., the total annual modulation divided by the yearly average M_2 amplitude (Figure 4.3b), is largest in the Canadian

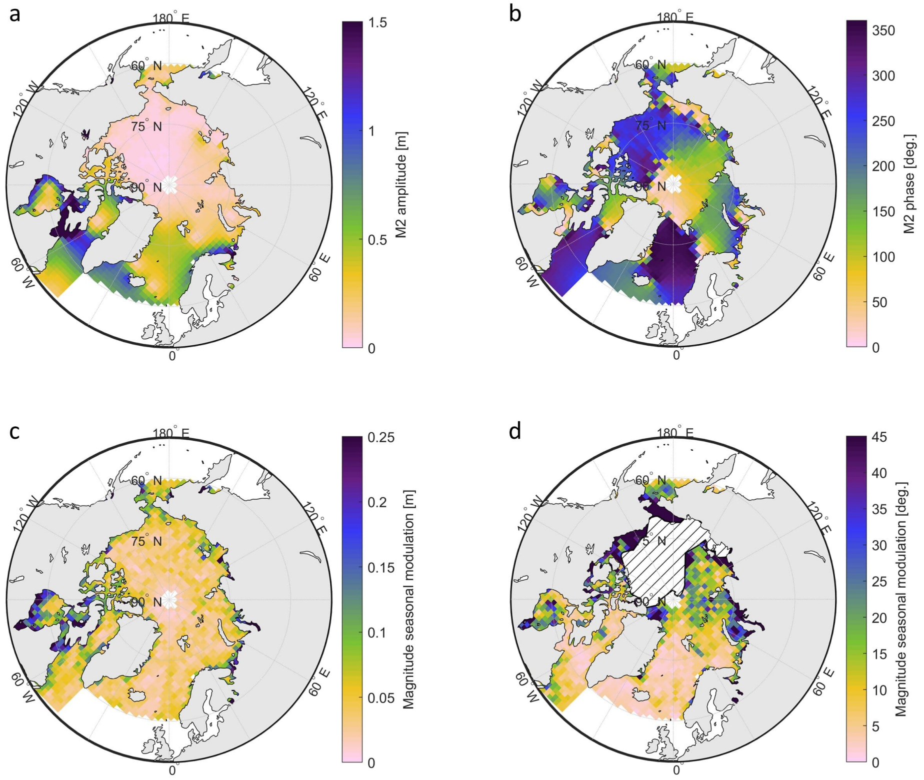


Figure 4.2: Altimeter-derived yearly average M_2 amplitude (a) and phase (b) and magnitude total annual modulation of M_2 amplitude (c) and phase (d). Hatched region indicates unreliable phase data that are not further considered. Data have been interpolated to a stereographic grid for plotting.

archipelago and Bering Sea where amplitudes are relatively low, and smallest in the European seas. The timing of seasonal modulation varies widely across the different regions. Most of the Canadian archipelago and Russian shelf experience highest M_2 amplitudes in summer months, while in Baffin Bay and the European seas, maximum M_2 amplitudes are predominantly observed in winter/spring (Figure 4.3c). In Hudson Bay, there is a near equal division between regions that experience maximum M_2 amplitudes in winter and in summer. In the Central Arctic and Bering Sea, largest amplitudes are observed in early summer.

In addition, March and September M_2 amplitudes are compared directly (Figure 4.3d, 4.5a). Amplitudes are often lower in March than September in the Canadian archipelago and on the Russian shelf (except for Laptev Sea, see Figure 4.1), while in Hudson Bay and the Bering Sea there is larger variability. In the Baffin Bay region, amplitude differences are predominantly positive, except for the Greenland fjords that are covered by landfast ice during March (Figure 4.1). Comparing Figure 4.5a with Figure 4.2c shows that

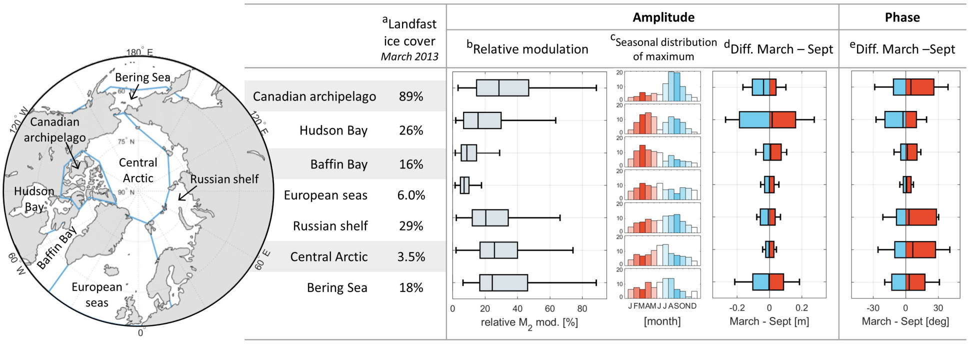


Figure 4.3: Overview of altimeter-derived seasonal modulation of M_2 for different subregions (indicated on the left). Landfsat ice cover (a) is calculated as the number of data points covered by landfast ice divided by the total number of data points in each region. Relative seasonal modulation (b) is calculated by dividing the total annual modulation (Figure 4.2c) by the M_2 amplitude (Figure 4.2a). The seasonal distribution of maximum (c) shows for each month the percentage of data points within the subregion that experience maximum M_2 amplitudes during that month. Here, Arctic summer months are displayed in blue and winter months in red, as they respectively relate to a negative (blue) or positive (red) March-September difference in amplitude (d). March-September differences (d,e) show the median, 10th and 90th percentiles of all values in the region.

in most of the Arctic the absolute March-September amplitude difference and the maximum seasonal modulation correspond closely. However, in Hudson Strait, Foxe Basin, southern Hudson Bay and the Ob estuary (Russia, see Figure 4.1), there are larger differences between the two measures (Figure 4.2c, 4.5a). This is in line with Figure 4.3c, where we showed that in several regions maximum amplitudes are observed in early summer rather than either March or September.

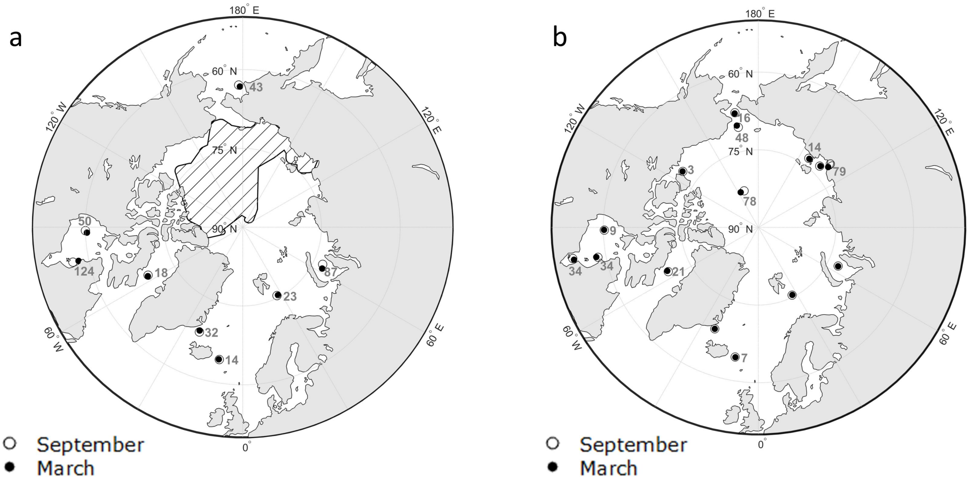


Figure 4.4: Location of amphidromic points in September and March derived from altimetry (a) and GTSM (b). Distance between each pair is indicated in gray (km). Hatched region indicates unreliable phase data that is not considered in further analysis.

In most of the Central Arctic, the Canadian archipelago and Baffin Bay phase differences are positive, indicating later arrival of the tide (phase delay) in March (Figure 4.3e, 4.5b). This is in contrast to Hudson Bay, where the phase is predominantly advanced (negative). Moreover, there are amphidromic points that shift up to 100 km between March and September: in Hudson Bay to the northeast, and on the Russian shelf to the west (Figure 4.4a).

4.3.2. MODELED SEASONAL MODULATION IN THE ARCTIC

Modeled seasonal modulation for 2013 (Figure 4.5c, d) compared with altimeter-derived modulation (Figure 4.6) shows that, although there is larger variability in altimeter-derived modulation, the predominant observed signal corresponds to the model-derived modulation in several regions. Altimeter-derived and model-derived differences in M_2 amplitude between March and September are negative in the Canadian archipelago and on the Russian shelf, positive in Baffin Bay and close to zero in the Arctic Ocean and European seas (Figure 4.5c, 4.6a). However, in Hudson Bay - and to a smaller extent, the Canadian archipelago and Bering Sea - positive differences are not captured well by the model. This results in large deviation with respect to altimeter-derived values and a large RMSE (0.21 m) and nRMSE (1.29; Figure 4.6c). While nRMSE values are large for all of the regions, this can in part be attributed to the larger variability in altimeter-derived modulation (Figure 4.6a). In addition, model-derived differences in amplitude are in most cases lower than those derived from observations, by up to 0.1 m (Figure 4.5a, c, Figure 4.6).

Model-derived phase differences correspond relatively well for Hudson Bay (predominantly negative) and the Canadian archipelago (predominantly positive) (Figure 4.5b, d, Figure 4.6d). In Baffin Bay and the European seas, both model and observations produce low phase differences, resulting in a small RMSE (≤ 15°). For the Russian shelf, Central Arctic and Bering Sea, modeled phase differences have a significantly smaller spread compared to altimeter-derived differences (up to 20°). In addition, the extent of the observed shifts in amphidromic points is not well captured by the model. In fact, in Hudson Bay the model-derived shift is in the opposite direction compared to what was observed (Figure 4.4b).

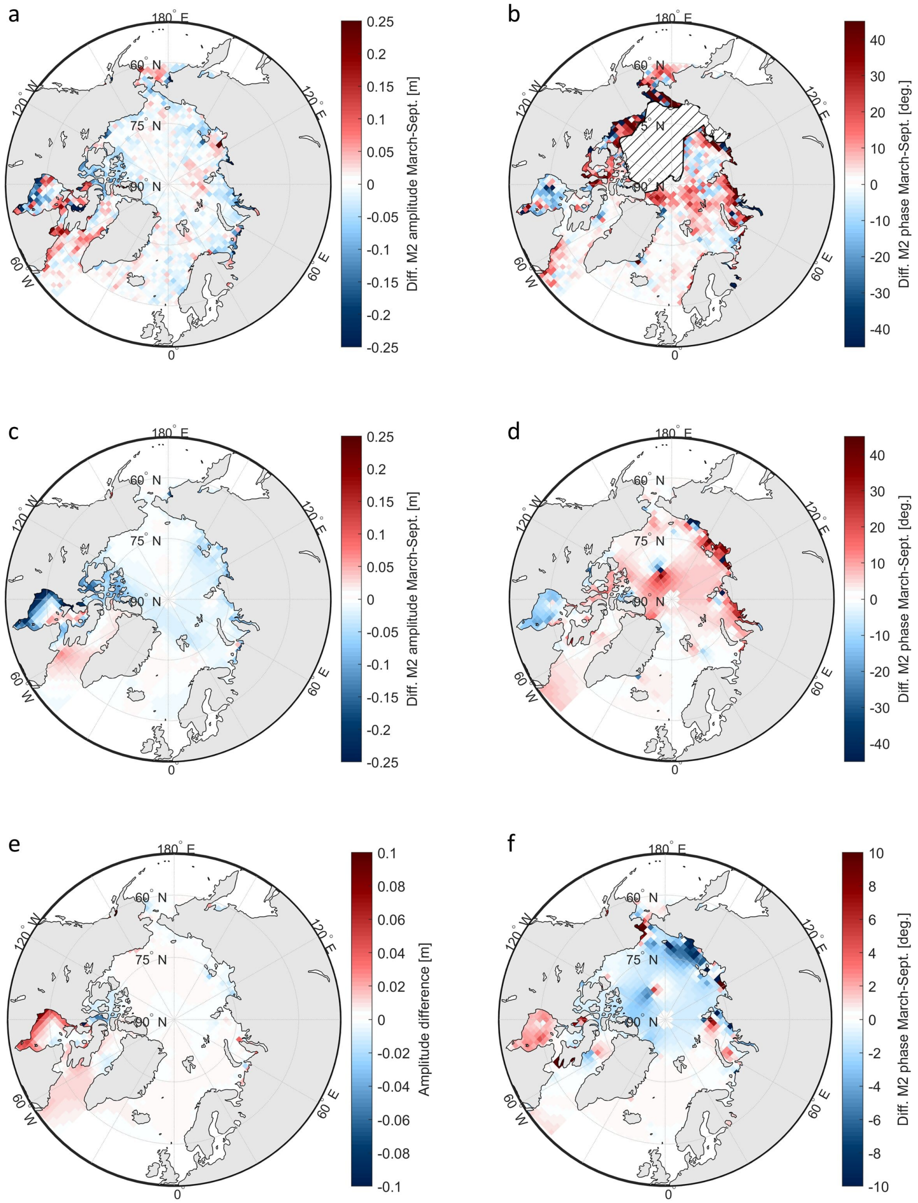


Figure 4.5: Altimeter-derived differences between March and September M_2 amplitudes (a) and phases (b) and model-derived differences in amplitude (c) and phase (d) for 2013. Bottom figures display the model-derived difference in amplitude (e) and phase (f) between March 2017 and March 2013 (2017 - 2013; different color scale). Hatched region indicates unreliable phase data that is not considered in further analysis. Data have been interpolated to a stereographic grid for plotting.

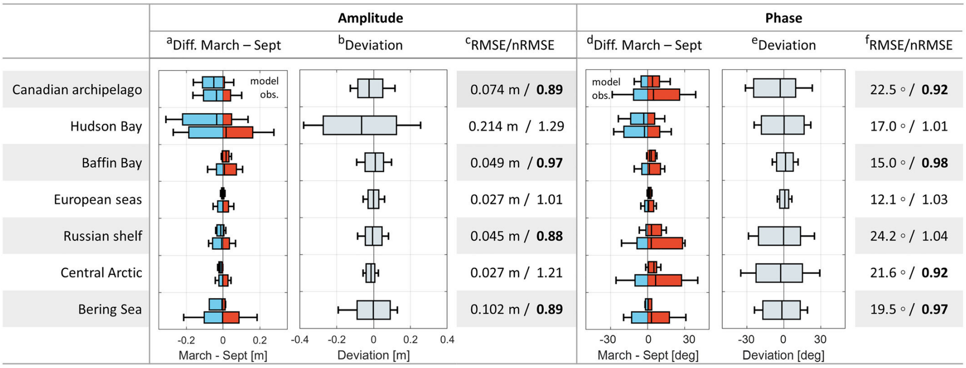


Figure 4.6: Overview of model performance with respect to altimeter-derived seasonal modulation for the subregions as in Figure 4.3. March-September differences (a,d) show first the model-derived differences for each region (2013) followed by altimeter-derived differences as in Figure 4.3. Deviation (b,e) shows the median, 10th and 90th percentiles of modelled March-September differences minus altimeter-derived values. RMSE and nRMSE (c,f) are calculated according to Equation 4.2 and Equation 4.3 respectively.

The model-derived amplitude and phase differences caused by the smaller landfast ice cover of 2017 (not shown here) display a similar pattern as in the case of 2013. In most of the domain, there is little difference between the March amplitudes of both years. However,the differences are significant in Hudson Bay and nearby regions; up to 0.05 m (~20%) (Figure 4.5e). In general, this results in reduced differences between March and September amplitudes in 2017. Negative differences in amplitude are reduced by up to 0.05 m in Hudson Bay and positive differences in Hudson Bay (east) and Baffin Bay are reduced by about 0.02 m. Seasonal modulation is larger in 2017 (up to 0.02 m) than in 2013 in the east of the Canadian Archipelago and in Labrador Sea. The east of the Canadian Archipelago is not covered by landfast ice in 2017, while it was in 2013, and has a lower M_2 amplitude in March 2017 than in 2013. This results in a larger amplitude difference in March with respect to September. In the Labrador Sea, the March amplitude of 2017 is more positive than that of 2013. In March 2017, Hudson Bay experiences less phase advance than in March 2013. In the Central Arctic, the Canadian archipelago, Baffin Bay and on the Russian shelf, phase delay is reduced as well. Nevertheless, the effect of interannual variations in landfast ice cover on model-derived seasonal modulation is smaller than the differences with respect to altimeter-derived modulation.

4.3.3. GLOBAL EFFECT OF ARCTIC LANDFAST ICE

The modeled effect of seasonal change in landfast ice extent on the seasonal modulation of tides is not restricted to the Arctic (Figure 4.7). Different March and September M_2 amplitudes are derived across the globe, although the values are significantly smaller south of ~ 50°N (up to 0.05 m) than in the Arctic. Overall, the magnitude of amplitude differences is related to the yearly average amplitude (see Figure 4.8a). However, no seasonal modulation was modeled for the Bay of Bengal (east of India, amplitudes up to 1 m) and some parts of open ocean where amplitudes are relatively large; e.g. the Indian

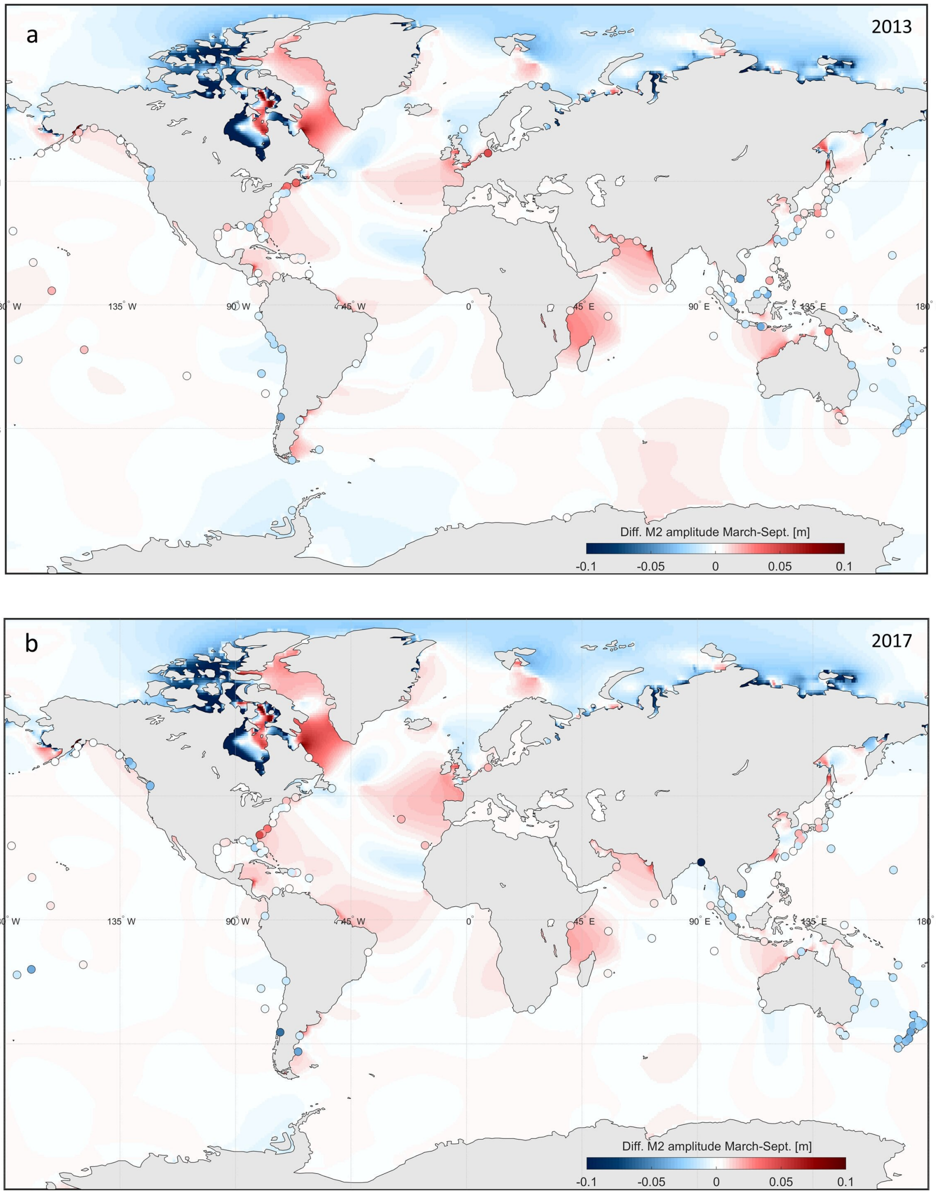


Figure 4.7: Difference between March and September M_2 amplitude as derived from tide gauge records (circles) and as modelled under the influence of Arctic landfast ice cover (background map) for 2013 (a) and 2017 (b).

and Pacific Oceans. In contrast, amplitude differences are observed in some regions with relatively small amplitudes. In these cases the amplitude differences are predominantly negative; e.g., west of northern Africa, southwest of Australia, west of Chile and east of Japan (Figure 4.7). There are small differences between the model-derived seasonal modulation for 2013 and 2017. In several regions with a positive March-September amplitude difference in 2013 (e.g., Gulf of Alaska, Gulf of Biscay and the area north of Brazil), these differences are up to 0.02 m larger in 2017. In contrast, the positive amplitude difference near Madagascar and west of India is reduced in 2017 by a similar extent. In other regions where there was a negative March-September amplitude difference in 2013 (e.g., along the coasts of Chile and Antarctica), these differences are reduced by ~ 0.02 m in 2017.

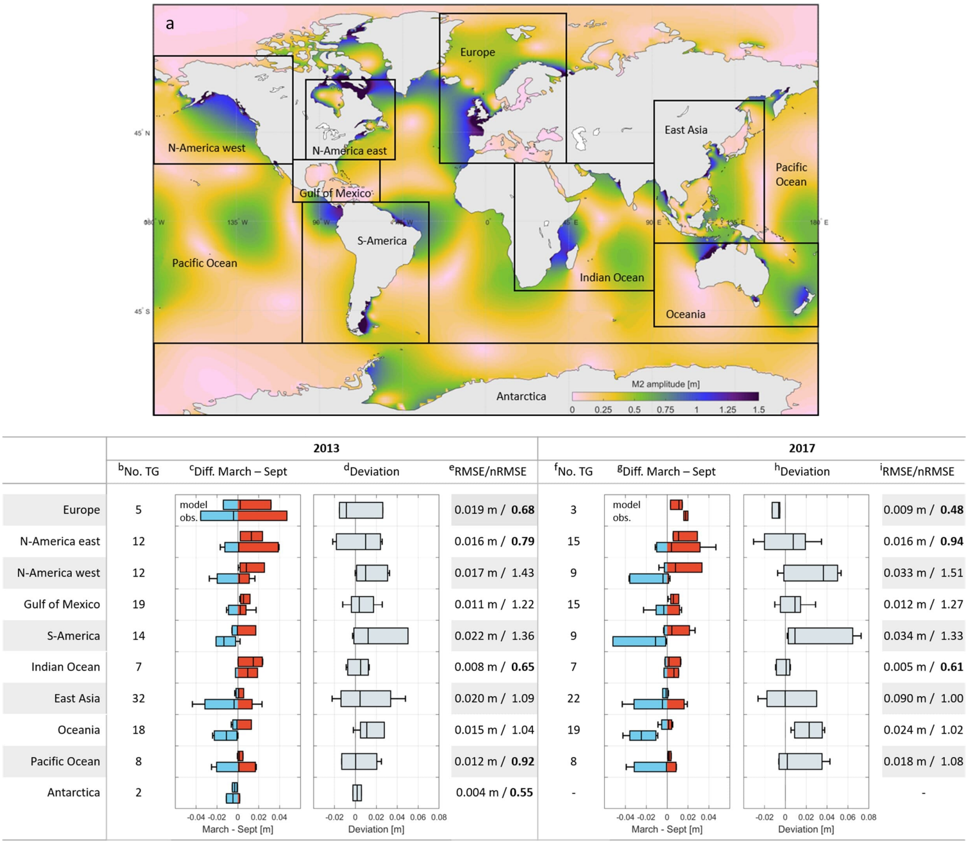


Figure 4.8: Overview of model performance with respect to tide gauge-derived seasonal modulation globally for 2013 and 2017 for regions defined in a. The background of a shows the yearly average modeled M_2 amplitude. No. TG (b, f) depicts the number of tide gauges included in each comparison (outliers excluded). March-September differences (c, g) show first the model-derived differences for each region followed by tide gauge-derived differences. Deviation (d, h) shows the median, 10th and 90th percentiles of modeled March-September differences minus tide gauge-derived values. RMSE and nRMSE (e, i) are calculated according to Equation 4.2 and Equation 4.3 respectively.

The degree of correspondence between model-derived seasonal modulation and modulation derived from tide gauge records, varies regionally (Figure 4.7) . Around Europe, New Zealand, the Arabian Sea, northern gulf of Alaska (2013), the Antarctic Peninsula (2013) and along the coast of Chile, the model-derived and observed differences in amplitude are of similar sign and magnitude. However, inconsistency is observed in other regions. For instance, the local variability as observed in the seasonal modulation derived from tide gauges in the waters around east Asia and along the east coast of North-America, is not present in the modeled signal. Regional examination (Figure 4.8) shows best agreement between modeled and tide gauge-derived seasonal modulation in Europe, N-America east, the Indian Ocean and Antarctica (2013). Overall, the agreement is slightly better in 2013 than in 2017. Nonetheless, this direct comparison highlights areas where the model largely deviates from observations and the RMSEs and nRMSEs are large (Figure 4.8e, i). Most noticeable examples are N-America west, the Gulf of Mexico, South America and East Asia. Model-derived differences between 2013 and 2017 correspond to observed differences for the tide gauges along the coast of Chile, but not for all tide gauges in the Gulf of Alaska. For the other regions we cannot compare tide gauge data of 2013 and 2017, since there was no data for either one of the two years.

We studied the seasonal modulation of the Arctic M_2 tide based on SAR altimetry-derived water levels, then linked these to seasonal variations in ice cover using the GTSM. This resulted in the first observation-based quantification of the seasonal modulation of the M_2 tide for the entire Arctic. Here, we combined data from two high-inclination satellite missions (CryoSat-2 and Sentinel-3) to ensure full (year) coverage of the Arctic waters. We stacked data from multiple tracks on a grid basis (following Cancet et al., 2018) to account for the low temporal resolution and irregular availability of the data. This approach minimized issues related to inseparable constituent pairs and aliasing, which would arise in conventional along-track tidal analysis (Cancet et al., 2018; Savcenko and Bosch, 2007).

Altimeter-derived seasonal modulation was compared to earlier studies that computed seasonal variations in M_2 amplitudes based on water level records from tide gauges or temporary pressure sensors. Comparable seasonal changes in amplitude and phase (magnitude and timing) were observed along the Russian coast (Kulikov et al., 2018); in most of Hudson Bay (Kleptsova and Pietrzak, 2018; St-Laurent et al., 2008) and in the Labrador Sea and Canadian archipelago (Kleptsova and Pietrzak, 2018). In northern Hudson Bay, altimeter-derived differences between March and September amplitudes were more positive than those derived from in situ data (St-Laurent et al., 2008). This difference may be caused by interannual variability; where our study is based on altimeter data from 2010-2019, St-Laurent et al. (2008) studied sea level data from August 2003 to August 2006.

There have been several attempts to quantify the seasonal modulation of the Arctic M_2 tide using tide models, although most of them were focused on the Hudson Bay area. These studies suggested various magnitudes of seasonal modulation. For instance, Müller et al. (2014) used a global ocean circulation and tide model (STORMTIDE) with an embedded sea-ice model and studied the effect of both ice cover variations and changes in stratification. They suggested a seasonal amplitude change up to 0.15 m (yearly max-min) in Hudson Bay. In contrast, Kleptsova and Pietrzak (2018) indicated that variations in the (total) sea ice cover alone cause a seasonal modulation that is up to tenfold larger. This is high compared to our altimetry-derived seasonal modulation in the area (0 to 0.25 m; Figure 4.2c). Kleptsova and Pietrzak (2018) attributed the deviation between their model output and in situ observations to the use of constant boundary conditions at the eastern entrance of Hudson Strait. St-Laurent et al. (2008) specifically compared M_2 March and September amplitudes (consistent with our approach), obtained from a 3D ocean model coupled to a sea-ice and snow model. According to their simulations, amplitudes differed up to 0.15 m. This is similar to the values we obtained (Figure 4.5a). However, they computed a negative amplitude difference in the western part of Hudson Bay and in Ungava Bay which contradicts the altimeter-derived values. Their outcome does largely correspond to our model-derived differences in amplitude (Figure 4.5c), even though our approach was more simplified. In addition, Kagan and Sofina (2010) modeled seasonal modulation of the M_2 tide on the Russian shelf and Central Arctic using a 3D model with an ice drift module included. Their results are comparable to our altimeter- and model-derived values for amplitude and our model-derived modulation of the phase, in terms of magnitude and sign.

In agreement with previous observations and model simulations, the results presented in this paper show that a landfast ice cover in one location typically leads to local amplitude decrease (Prinsenberg and Freeman, 1986; Sverdrup, 1927). Moreover, observed and modeled phase advance in Hudson Bay and phase delay in other regions are in line with earlier theoretical considerations (Prinsenberg and Freeman, 1986). According to the latter, increased friction results in reduced phase velocities and thus a larger phase lag. However, in Hudson Bay, the interaction between the incoming and reflected wave causes a change in phase/amplitude patterns, and the net phase is advanced (Prinsenberg and Freeman, 1986). Amplitude decrease is clearly visible in the Canadian archipelago and on the Russian shelf where the landfast ice cover is large. However, in Hudson Bay, Ungava Bay, Labrador Sea and part of Bering Sea, the opposite effect is observed. Here, areas that are not covered by landfast ice show an increase in amplitude during March, suggesting a complex relation between ice in one location and a far-field effect on tides in another. This was previously shown by Kagan and Romanenkov (2007), whose modelling study indicated that the ice-covered White Sea caused an amplitude increase in the ice-free Barents Sea. In Hudson Bay, the observed increase in amplitude at the western coast and, to a smaller extent in the southern part can be related to the observed shift of amphidromic points away from these areas. The latter was not captured by the model and contradicts earlier studies (Kleptsova and Pietrzak, 2018; Prinsenberg and Freeman, 1986). This suggests that the shift might be caused by other seasonal processes besides variations in landfast ice cover. Moreover, the observed tidal amplification in Labrador Sea is in line with Arbic et al. (2004), who linked this to ice cover presence in Ungava Bay/Hudson Strait. In addition, Arbic et al. (2010) showed that increased friction in Hudson Strait results in increased amplitudes along the North Atlantic coast. This is consistent with our simulated seasonal modulation and the modulation derived from tide gauges (Figure 4.7). Figure 4.7 even suggests such effects to be noticeable on a global scale. However, note that we observed significant differences between model- and tide gauge-derived seasonal modulation (Figure 4.8) in specific regions and that the seasonal modulation derived from tide gauge records is likely the result of multiple (regional) processes. Regardless, it has been shown that the sensitivity of the model to landfast ice reaches well beyond the Arctic. We expect that similar (yet out of phase) responses will be present due to variations in extent of landfast ice around Antarctica.

Although many of the observed differences in amplitude and phase are to a certain extent captured by our model, there is no complete agreement. In particular the positive differences in amplitude (larger amplitudes in March than in September in Hudson Bay, Hudson Strait, Bering Sea, Laptev Sea) are not captured by the model. However, the model only considers the effect of seasonal variations in landfast ice on the tide and other effects were deliberately ignored. For example, including drifting sea ice (with strong internal stresses) in the model, may lead to enhanced seasonal modulation. However, this is not likely to explain the positive amplitude differences in Hudson Strait as this was also not obtained by St-Laurent et al. (2008), who did consider drifting ice. In addition to this, for such drifting sea ice the imposition of wind forcing could result in a different frictional stress at the ice-ocean interface as opposed to the case of drifting ice without winds (Hibler et al., 2006). This might further affect tidal amplitudes and phases. Wind variations have also been connected to seasonal modulation of tides on the northwest European shelf (Zijl et al., 2016), but were not included in our model setup. Finally, as GTSM is a 2D model with assumed uniform density, the effect of spatial density differences and stratification was not considered. Spatial differences in density (i.e., baroclinic pressure gradients) influence the movement of water and therefore the mean water level (Zijl et al., 2013). Especially in the Hudson Bay system, adjustments of the mean sea level affect local tides (Pickering et al., 2017). Stratification impacts the stability of a water column and thus the amount of energy lost into turbulent processes. On the one hand, a stratified water column is more stable, which could result in more pronounced tidal transport and surface tides, compared to well-mixed conditions (Müller, 2012; Müller et al., 2014). On the other hand, stratified conditions have been associated with the generation of internal tides, which can cause local dissipation of barotropic energy (Müller, 2012; Müller et al., 2012). Considering these (local) processes could explain the larger variability in altimeter-derived results. Moreover, the fact that peak amplitudes are more often observed in July/August instead of in September (Figure 4.3c) is likely related to water properties. While the (landfast) ice cover reaches its minimum in September, average water temperature is highest and salinity is lowest in August (Steele et al., 2001).

Differences between altimeter- and model-derived seasonal modulation could also indicate shortcomings of the study/model. Firstly, model simulations are based on landfast ice cover variations of two distinct years, while the altimeter-derived seasonal modulation is calculated based on the entire water level dataset (2010-2019). This means that in the altimeter-derived seasonal modulation, interannual variation is averaged out. However, since model simulations are based on two years representing the maximum and minimum landfast ice cover within the time span of the data, the altimeter-derived seasonal modulation is expected to fall within the range of simulated modulation. Nevertheless, judging by the differences between the simulations for 2013 and 2017 (Figure 4.5e, f), the effect of interannual variations in landfast ice cover is small in most of the domain. In regions where the effect of interannual variation is significant (e.g., Hudson Bay), this does not explain the deviation with respect to altimeter-derived modulation. Secondly, in many Arctic regions, the bathymetry is still largely uncertain, while in some regions the bathymetry has a large impact on tide modelling. This might be of particular importance in Hudson Bay and Hudson Strait which display near-resonant properties (Arbic et al., 2007).

The results presented in this paper indicate the seasonal modulation of the Arctic M_2 tide can be significant (~20% of amplitude), in particular on shelves and in bays. Our work suggests that, to a large extent, this modification of the tide can be attributed to variations in landfast ice cover. The part that cannot be explained by landfast ice cover variations alone, suggests more research is needed into the impact of other seasonal processes. This entails extending the model approach by including drifting sea ice and updating the bathymetry after the newest GEBCO release, that includes the most recent IBCAO v4 dataset on the Arctic (GEBCO Compilation Group, 2020). Finally, while the Arctic sea ice cover varies seasonally, it is also subject to significant interannual variability and a declining trend (e.g., Kwok, 2018; Perovich and Richter-Menge, 2009; Serreze and Stroeve, 2015). In our study, the effect of landfast ice on tides was simulated for two years with maximum and minimum cover since 2011. Although differences in ice cover between these years are largest in the Canadian archipelago and on the Russian shelf, the difference in landfast ice cover also alters seasonal modulation of M_2 amplitudes in Hudson and Baffin Bay by up to 0.05 m. This, and the magnitude of the observed relation between seasonal variation in ice extent and tide modification, raises questions concerning the impact of a continuous Arctic sea ice decline. In particular since we showed that the effects of sea ice on tides are not restricted to the Arctic, this calls for further research involving extended use of altimetry data.

4.4. SUMMARY AND CONCLUSION

The seasonal modulation of the M_2 amplitude and phase, averaged over the period 2010-2019, was obtained for most of the Arctic based on SAR derived water levels from two high-inclination satellite missions, CryoSat-2 and Sentinel-3. Results show significant variations in tidal amplitude (up to 0.25 m) along the Russian coastline, in the Hudson Bay system, the Canadian archipelago and Bering Sea. Phase differences of ~10° are observed throughout the domain, with largest differences along the Russian and Canadian coast (up to 45°).

In addition, the M_2 amplitudes and phases were reproduced for March and September, months with, respectively, the maximum and minimum extent of Arctic landfast ice. Although an amplitude decrease was observed at most locations where landfast ice is present in March (e.g. the Russian coastline, Canadian archipelago, south of Hudson Bay), in other regions (Hudson Strait and Baffin Bay) the amplitude increased. This indicates that landfast ice not only causes local tidal dampening but has an additional far-field effect on tides. These results were subsequently compared to model simulations run for two limiting cases of Arctic landfast ice cover: 1) annually maximum landfast ice cover (March) and 2) no landfast ice cover (September). In many regions model-derived amplitude and phase differences correspond well to the values that were obtained from altimetry. However, in certain regions the positive differences in amplitude (larger in March than in September) are not well captured by the model and simulated phase differences are smaller than observed. Further research is required to fully explain the observed seasonal modulation of the M_2 tide, where other potential drivers besides variations in Arctic landfast ice should be considered.

Finally, the impact of variations in Arctic landfast ice on global tides was assessed by comparing simulated March and September amplitudes to amplitudes derived from tide gauge records. Modeled amplitude differences are considerable (up to 0.05 m) and correspond to the observed signal in several regions, but deviate in other regions. This preliminary analysis emphasizes the importance of further research in the relation between seasonal variations in sea ice and (global) modification of tides.

BIBLIOGRAPHY

Arbic, B. K., MacAyeal, D. R., Mitrovica, J. X., & Milne, G. A. (2004). Ocean tides and Heinrich events. *Nature*, *432*(7016), 460–460.

Arbic, B. K., St-Laurent, P., Sutherland, G., & Garrett, C. (2007). On the resonance and influence of the tides in Ungava Bay and Hudson Strait. *Geophysical Research Letters*, *34*(17).

Arbic, B. K., Wallcraft, A. J., & Metzger, E. J. (2010). Concurrent simulation of the eddying general circulation and tides in a global ocean model. *Ocean Modelling*, (3–4), 175–187.

Bij de Vaate, I., Vasulkar, A. N., Slobbe, D. C., & Verlaan, M. (2021). The Influence of Arctic Landfast Ice on Seasonal Modulation of the M2 Tide. *Journal of Geophysical Research: Oceans*, *126*(5), e2020JC016630. https://doi.org/https://doi.org/10.1029/2020JC016630

Caldwell, P. C., Merrifield, M. A., & Thompson, P. R. (2015). Sea level measured by tide gauges from global oceans - the Joint Archive for Sea Level holdings (NCEI Accession 0019568), Version 5.5. NOAA National Centers for Environmental Information, Dataset. https://doi.org/https://doi.org/10.7289/V5V40S7W

Cancet, M., Andersen, O., Lyard, F., Cotton, D., & Benveniste, J. (2018). Arctide2017, a high-resolution regional tidal model in the Arctic Ocean. *Advances in space research*, *62*(6), 1324–1343.

Codiga, D. L. (2020). UTide Unified Tidal Analysis and Prediction Functions. MATLAB Central File Exchange. [Software] [Last checked on 10/06/2020].

Dinardo, S., Fenoglio-Marc, L., Buchaupt, C., Becker, M., Scharroo, R., Fernandes, M. J., & Benveniste, J. (2018). Coastal SAR and PLRM altimetry in German Bight and West Baltic Sea. *Advances in Space Research*, *62*(6), 1371–1404.

Donlon, C., Berruti, B., Buongiorno, A., Ferreira, M. H., Féménias, P., Frerick, J., ..., & Nieke, J. (2012). The global monitoring for environment and security (GMES) sentinel-3 mission. *Remote Sensing of Environment*, *120*, 37–57.

Foreman, M. G. G., & Henry, R. F. (1989). The harmonic analysis of tidal model time series. *Advances in water resources*, *12*(3), 109–120.

GEBCO Compilation Group. (2019). GEBCO 2029 Grid. https://doi.org/https://doi.org/10.5285/836f016a-33be-6ddc-e053-6c86abc0788e

GEBCO Compilation Group. (2020). GEBCO 2020 Grid. https://doi.org/https://doi.org/10.5285/a29c5465-b138-234d-e053-6c86abc040b9

Godin, G. (1986). Modification by an ice cover of the tide in James Bay and Hudson Bay. *Arctic*, 65–67.

Hibler, W. D. (1979). A Dynamic Thermodynamic Sea Ice Model. *Journal of Physical Oceanography*, *9*(4), 815–846.

Hibler, W. D., Roberts, A., Heil, P., Proshutinsky, A. Y., Simmons, H. L., & Lovick, J. (2006). Modeling M2 tidal variability in Arctic sea-ice drift and deformation. *Annals of Glaciology*, *44*, 418–428. https://doi.org/10.3189/172756406781811178

Kagan, B. A., & Romanenkov, D. A. (2007). On the variability of tidal constants induced by the influence of one subsystem on another. *Atmospheric and Oceanic Physics*, *43*(3), 357–362.

Kagan, B. A., Romanenkov, D. A., & Sofina, E. V. (2007). Modeling the tidal ice drift and ice-induced changes in tidal dynamics on the Siberian continental shelf. *Izvestiya, Atmospheric and Oceanic Physics*, *43*(6), 766–784. https://doi.org/10.1134/S0001433807060114

Kagan, B. A., & Sofina, E. V. (2010). Ice-induced seasonal variability of tidal constants in the Arctic Ocean. *Continental Shelf Research*, *30*(6), 643–647. https://doi.org/https://doi.org/10.1016/j.csr.2009.05.010

Kleptsova, O., & Pietrzak, J. D. (2018). High resolution tidal model of Canadian Arctic Archipelago, Baffin and Hudson Bay. *Ocean Modelling*, *128*, 15–47.

Kowalik, Z. (1981). A Study of the M2 Tide in the Ice-Covered Arctic Ocean. *Modeling, Identification and Control*, *2*(4), 201–223.

Kulikov, M. E., Medvedev, I. P., & Kondrin, A. T. (2018). Seasonal variability of tides in the Arctic Seas. *Russian Journal of Earth Sciences*, *18*(5). https://doi.org/https://doi.org/10.2205/2018ES000633

Kwok, R. (2018). Arctic sea ice thickness, volume, and multiyear ice coverage: losses and coupled variability (1958–2018). *Environmental Research Letters*, *13*(10), 105005.

Langleben, M. P. (1982). Water drag coefficient of first-year sea ice. *Journal of Geophysical Research*, *87*(C1), 573. https://doi.org/10.1029/JC087iC01p00573

Laxon, S. (1994). Sea ice altimeter processing scheme at the EODC. *International Journal of Remote Sensing*, *15*(4), 915–924.

Li, Z., Zhao, J., Su, J., Li, C., Cheng, B., Hui, F., Yang, Q., & Shi, L. (2020). Spatial and temporal variations in the extent and thickness of arctic landfast ice. *Remote Sensing*, *12*(1). https://doi.org/https://doi.org/10.3390/RS12010064

Madsen, O. S., & Bruno, M. S. (1987). A methodology for the determination of drag coefficients for ice floes. *Journal of Offshore Mechanics and Arctic Engineering*, *109*(4), 381–387. https://doi.org/10.1115/1.3257035

Mahoney, A. R., Eicken, H., Gaylord, A. G., & Gens, R. (2014). Landfast sea ice extent in the Chuckchi and Beaufort Seas: The annual cycle and decadal variability. *Cold Regions Science and Technology*, *103*, 41–56.

McPhee, M. G. (1980). An analysis of pack ice drif in summer. In R. Pritchard (Ed.), *Sea-Ice Processes and Models* (pp. 62–75). Seatle: University of Washington Press.

Müller, M. (2012). The influence of changing stratification conditions on barotropic tidal transport and its implications for seasonal and secular changes of tides. *Continental Shelf Research*, *47*, 107–118. https://doi.org/https://doi.org/10.1016/j.csr.2012.07.003

Müller, M., Cherniawsky, J. Y., Foreman, M. G. G., & von Storch, J. S. (2012). Global M2 internal tide and its seasonal variability from high resolution ocean circulation and tide modeling. *Geophysical Research Letters*, *39*(19).

Müller, M., Cherniawsky, J. Y., Foreman, M. G. G., & von Storch, J. S. (2014). Seasonal variation of the M 2 tide. *Ocean Dynamics*, *64*(2), 159–177. https://doi.org/https://doi.org/10.1007/s10236-013-0679-0

National Snow and Ice Data Center. (2020). Charctic Interactive Sea Ice Graph [Last checked on 06/06/2019].

Pawlowicz, R., Beardsley, B., & Lentz, S. (2002). Classical tidal harmonic analysis including error estimates in MATLAB using T-TIDE. *Computers & Geosciences*, *28*, 929–937.

Pease, C. H., Salo, S. A., & Overland, J. E. (1983). Drag measurements for first-year sea ice over a shallow sea. *Journal of Geophysical Research*, *88*(C5), 2853. https://doi.org/10.1029/JC088iC05p02853

Perovich, D. K., & Richter-Menge, J. A. (2009). Loss of sea ice in the Arctic. *Annual Review of Marine Science*, *1*, 417–441.

Pickering, M. D., Horsburgh, K. J., Blundell, J. R., Hirschi, J. M., Nicholls, R. J., Verlaan, M., & Wells, N. C. (2017). The impact of future sea-level rise on global tides. *Continental Shelf Research*, *142*, 50–68.

Poisson, J. C., Quartly, G. D., Kurekin, A. A., Thibaut, P., Hoang, D., & Nencioli, F. (2018). Development of an ENVISAT altimetry processor providing sea level continuity between open ocean and Arctic leads. *IEEE Transactions on Geoscience and Remote Sensing*, *56*(9), 5299–5319.

Prinsenberg, S. J., & Freeman, N. G. (1986). Tidal heights and currents in Hudson Bay and James Bay. In I. P. Martini (Ed.), *Canadian Inland Seas* (pp. 205–216). Amsterdam: Elsevier.

Ray, C., Martin-Puig, C., Clarizia, M. P., Ruffini, G., Dinardo, S., Gommenginger, C., & Benveniste, J. (2015). SAR altimeter backscattered waveform model. *IEEE Transactions on Geoscience and Remote Sensing*, *53*(2), 911–919.

Savcenko, R., & Bosch, W. (2007). Residual Tide Analysis in Shallow Water-Contributions of ENVISAT and ERS Altimetry. *ESA-SP636 (CD-ROM), Proceedings of the Envisat Symposium*.

Schulz, A. T., & Naeije, M. (2018). SAR Retracking in the Arctic: Development of a year-round retrackers system. *Advances in Space Research*, *62*(6), 1292–1306.

Serreze, M. C., & Stroeve, J. (2015). Arctic sea ice trends, variability and implications for seasonal ice forecasting. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *373*.

Steele, M., Morley, R., & Ermold, W. (2001). PHC: A global ocean hydrography with a high-quality Arctic Ocean. *Journal of Climate*, *14*(9), 2079–2087.

St-Laurent, P., Saucier, F. J., & Dumais, J. F. (2008). On the modification of tides in a seasonally ice-covered sea. *Journal of Geophysical Research: Oceans*, *113*(11), 1–11. https://doi.org/https://doi.org/10.1029/2007JC004614

Sverdrup, H. (1927). Dynamics of tides on the north Siberian Shelf. *Geof. Publ.*, *4*, 3–75.

U.S. National Ice Center. (2009). National Ice Center Arctic Sea Ice Charts and Climatologies in Gridded Format, Version 1. Boulder, Colorado USA. NSIDC: National Snow and Ice Data Center. https://doi.org/https://doi.org/10.7265/N5X34VDB

Verlaan, M., De Kleermaeker, S., & Buckman, L. (2015). GLOSSIS: Global Storm Surge Forecasting and Information System. Auckland, New Zealand: Australasian Coast & Ports Conference.

Wang, X., Verlaan, M., Apecechea, M. I., & Lin, H. X. (2021). Computation-efficient Parameter Estimation for a High-Resolution Global Tide and Surge Model (GTSM). *Journal of Geophysical Research: Oceans*, *126*, e2020JC016917. https://doi.org/https://doi.org/10.1029/2020JC016917

Wernecke, A., & Kaleschke, L. (2015). Lead detection in Arctic sea ice from CryoSat-2: quality assessment, lead area fraction and width distribution. *The Cryosphere*, *9*(5), 1955–1968.

Wingham, D. J., Francis, C. R., Baker, S., Bouzinac, C., Brockley, D., Cullen, R., ..., & Phalippou, L. (2006). CryoSat: A mission to determine the fluctuations in Earth's land and marine ice fields. *Advances in Space Research*, *37*(4), 841–871.

Zijl, F., Verlaan, M., & Gerritsen, H. (2013). Improved water-level forecasting for the Northwest European Shelf and North Sea through direct modelling of tide, surge and non-linear interaction. *Ocean Dynamics*, *63*(7), 823–847.

Zijl, F., Verlaan, M., & Irazoqui, M. (2016). On the origins of annual modulation of M2 and M4 harmonic consituents [Last checked on 18/06/2020].

5

CLASSIFICATION METHODS FOR WATER LEVEL RETRIEVAL IN THE ARCTIC OCEAN

In the Arctic Ocean, obtaining water levels from satellite altimetry is hampered by the presence of sea ice. Hence, water level retrieval requires accurate detection of fractures in the sea ice (leads). This paper describes a thorough assessment of various surface type classification methods, including a thresholding method, nine supervised-, and two unsupervised machine learning methods, applied to Sentinel-3 Synthetic Aperture Radar Altimeter data. For the first time, the simultaneously sensed images from the Ocean and Land Color Instrument, onboard Sentinel-3, were used for training and validation of the classifiers. This product allows to identify leads that are at least 300 meters wide. Applied to data from winter months, the supervised Adaptive Boosting, Artificial Neural Network, Naïve-Bayes, and Linear Discriminant classifiers showed robust results with overall accuracies of up to 92%. The unsupervised Kmedoids classifier produced excellent results with accuracies up to 92.74% and is an attractive classifier when ground truth data is limited. All classifiers perform poorly on summer data, rendering surface classifications that are solely based on altimetry data from summer months unsuitable. Finally, the Adaptive Boosting, Artificial Neural Network, and Bootstrap Aggregation classifiers obtain the highest accuracies when the altimetry observations include measurements from the open ocean.

This chapter has been published as Bij de Vaate, I., Martin, E., Slobbe, D. C., Naeije, M., & Verlaan, M. (2022). Lead Detection in the Arctic Ocean from Sentinel-3 Satellite Data: A Comprehensive Assessment of Thresholding and Machine Learning Classification Methods. *Marine Geodesy*, *45*(5), 462–495. https://doi.org/https://doi.org/10.1080/01490419.2022.2089412

5.1. INTRODUCTION

The Arctic Ocean is highly affected by global warming. The region is subject to temperature changes of about three times the global average (IPCC, 2021), of which the Arctic sea ice decline is a major consequence. In only two decades, the perennial sea ice cover has decreased by 50%, and the remainder will likely be lost by 2050 (Kwok, 2018). Since changes in the Arctic Ocean have a global impact, the region is of great scientific interest. However, due to the remote location of the Arctic and its relatively harsh environmental conditions, the availability of observational input is limited. For instance, information on the Arctic Ocean sea surface height (SSH) is needed for many purposes; from studying the influence of Arctic glacier melt on the regional sea level (e.g., Cazenave et al., 2019; Rose, 2013) to monitoring sea ice thickness (e.g., S. W. Laxon et al., 2013; Wernecke and Kaleschke, 2015). Unfortunately, in situ data are limited to a few tide gauges at the coast and the presence of sea ice hampers measurements by satellite altimeters. In this respect, Synthetic Aperture Radar (SAR) altimetry provides a solution. SAR altimeters have a higher along-track resolution compared to conventional radar altimeters (Donlon et al., 2012), which allows measuring the SSH through fractures in the sea ice, so-called leads. However, this requires careful discrimination between measurements from sea ice and leads.

Fortunately, because of differences in the surface characteristics of sea ice and leads, these surfaces typically cause distinct SAR returns. Consequently, various classification methods have been developed that use waveform features, which describe the unique features of the SAR return signal. Empirical methods, reliant on setting thresholds for these waveform features, have been widely used to classify radar returns (e.g., S. Laxon, 1994; Peacock and Laxon, 2004; Poisson et al., 2018; Zakharova et al., 2015). More recently, machine learning-based classification methods have gained popularity (e.g., Dettmering et al., 2018; Lee et al., 2016; Muller et al., 2017; Poisson et al., 2018). Machine learning-based methods can produce higher accuracies as they can overcome shortcomings associated with the simple thresholding methods, such as failing to deal with waveform features that contain aliasing between leads and sea ice (Lee et al., 2016).

Despite the promising implementations of machine learning classifiers presented in earlier studies, some uncertainties remain. Firstly, the aforementioned classifiers and their performances cannot be directly compared as these studies involve different study areas, sensors, and validation data. For instance, it is still unknown whether unsupervised machine learning classifiers can outperform supervised learning classifiers. Secondly, the validation in previous studies was often limited, e.g., to small areas (Dettmering et al., 2018: the Greenland Sea for unsupervised classification) or few SAR data (Lee et al., 2016: 239 waveforms). The most important restriction on the extent of validation is the need to generate ground truth data, which is often done through visual inspection (Lee et al., 2016; Quartly et al., 2018). As a result, it remains unknown whether the classifier performance is location-dependent. This may, for instance, be due to regional variations in the prevailing sea ice type (e.g., first-year ice or multi-year ice). Thirdly, seasonal differences in classification performance are poorly understood, as lead detection methods are typically only tested on data from winter months. Shu et al. (2020) tested their classifier on data from spring (up to May) and showed reduced performance for May compared to earlier months. In contrast, Dawson et al. (2022) recently obtained comparable performances for the classification of data from winter and summer months.

The main objective of this paper is to provide a comprehensive assessment of lead detection methods, applied to SAR altimetry. Therefore, supervised- and unsupervised machine learning methods and a thresholding method are applied to a wide range of study areas in the Arctic Ocean, to identify the most suited classifier to be applied for SSH estimation. A key opportunity is recognized in using data from the Sentinel-3 satellites (operated by ESA and EUMETSAT), as these satellites are equipped with a Synthetic Aperture Radar Altimeter (SRAL) and the Ocean and Land Color Instrument (OLCI). Therefore, classification methods applied to data acquired by SRAL can be validated using simultaneously acquired OLCI images. This combination of temporally aligned data sources is extremely beneficial to the research as it eliminates the need to employ ice drift models to correct for the relocation of the ice in-between the measurements (Quartly et al., 2018). Although an operator-controlled selection of cloud-free images is required, most of the validation data generation process is successfully automated. In this way, a larger study area can be included in the validation. The classifiers are additionally applied to data from different years and summer months to gain insight into temporal effects. Finally, as part of the Arctic Ocean is completely ice-free during summer months, the classifiers are also tested with consideration of data from the open ocean.

In the following sections, we first expand on the Sentinel-3 satellite data that were used, the procedures that were adopted for generating the ground truth data, the specific classifiers that were implemented, and the measures that were used to assess the classifier performance. Thereafter, the different test cases are introduced, followed by the results and a discussion of the main findings.

5.2. DATA

5.2.1. SYNTHETIC APERTURE RADAR ALTIMETER (SRAL)

This study uses SAR altimetry level 1B data (non-time critical), retrieved by the Synthetic Aperture Radar Altimeter (SRAL) instrument of the Sentinel-3A and Sentinel-3B satellites (Donlon et al., 2012). In contrast to conventional altimeters, SAR altimeters obtain a relatively high along-track resolution by applying coherent processing of groups of transmitted pulses, exploiting the Delay-Doppler effect (Raney, 1998). The along-track resolution is ~ 300 m (Donlon et al., 2012). The shape of the returned signal relates to the roughness and orientation of the surface from which the signal is reflected. For instance, smooth surfaces such as leads cause specular returns, while rougher surfaces like sea ice and open ocean result in diffuse reflections (S. W. Laxon et al., 2013; Poisson et al., 2018). In this paper, the full waveforms (128 data points) were reduced to twelve waveform features (see Figure 5.1 and Table 5.1), which were then used as input for the classifiers.

Table 5.1: Description and equation of the waveform features considered in this study. Here n is the number of bins that make up the waveform, P the power of an individual bin, the $\bar{P}$ average power, $Pmax$ the maximum power, and the standard deviation of the distribution.

Waveform parameter (Abbreviation)	Description	Equation
Maximum Power (MAX)	The maximum power value of the waveform in counts	-
Kurtosis (kurt)	A measure of peakiness of the power distribution (Lee et al. 2016). Kurtosis is a fourth standardized moment.	$kurt = \frac{1}{n}\frac{\sum_{i=1}^{n}(P_i - \bar{P}^4}{)}\sigma^4$
Skewness (skew)	A measure of how slanted the power distribution is. Skewness is a third standardized moment.	$skew = \frac{1}{n}\frac{\sum_{i=1}^{n}(P_i - \bar{P}^3}{)}\sigma^3$
Pulse Peakiness	A measure of the peakiness of the waveform. It is found by dividing the maximum power by the total accumulated power of the waveform (Wernecke and Kaleschke, 2015).	$PP = \frac{P_{max}}{\sum_{i=1}^{n} P_i}$
Waveform width (ww)	The number of bins surrounding the peak with a power of at least 1% of the maximum power (Dettmering et al., 2018) (see Figure 5.1).	-
Leading edge Width (LeW)	The number of bins before the peak that has 1% to 99% of the maximum power value (see Figure 5.1).	-
Trailing edge Width (TeW)	The number of bins after the peak that has 99% to 1% of the maximum power value (see Figure 5.1).	-
Backscatter Coefficient (sigma0)	The radar backscatter coefficient describes the surface properties, radar frequency, polarization, and incident angle (Wingham et al., 2006). Sigma0 values are computed as in Satellites (2011), using the maximum power from the original waveform.	-
Pulse Peakiness Left (PPL)	Modified PP, only considering the three bins on the left of the bin belonging to the maximum (Ricker et al., 2014).	$PPL = \frac{P_{max}}{\sum_{i=imax-3}^{imax-1} P_i}$
Pulse Peakiness Left (PPR)	Modified PP, only considering the three bins on the right of the bin belonging to the maximum (Ricker et al., 2014).	$PPR = \frac{P_{max}}{\sum_{i=imax+1}^{imax+3} P_i}$
Pulse Peakiness Left (PPloc)	Modified PP, only considering the three bins on the left and three bins on the right of the bin with maximum power.	$PPL = \frac{P_{max}}{\sum_{i=imax+3}^{imax-3} P_i}$
Number of Peaks (NrPeaks)	Number of peaks with a peak prominence greater than 5% and a minimum separation distance of five bins	-

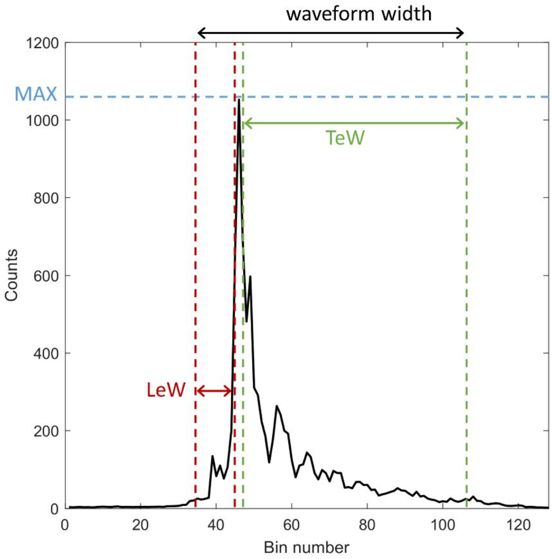


Figure 5.1: An example of Sentinel-3 SRAL level 1B waveform return from sea ice. The waveform features maximum power (MAX), leading edge width (LeW), trailing-edge width (TeW), and the waveform width (ww) are presented in the figure.

5.2.2. OCEAN AND LAND COLOR INSTRUMENT (OLCI)

Optical images taken by the Ocean and Land Color Instrument (OLCI) onboard the Sentinel-3 satellites were used to create ground truth data for the training and validation of the SAR altimetry-based classification. OLCI is a push-broom imaging spectrometer that contains 21 spectral bands (Oa1–Oa21) ranging from 400 nm to 1020 nm (Donlon et al., 2012). This study used the level 1B product, which consists of top of atmosphere radiances, calibrated to geophysical units ($Wm^{-2}\ sr^{-1}mm^{-1}$), georeferenced onto the Earth's surface, and spatially resampled onto an evenly spaced grid (Donlon et al., 2012). Pseudo-color images were constructed using three spectral bands from the OLCI data, which are; Oa3 (442.5 nm), Oa5 (510 nm), and Oa8 (665 nm). In these images, water surfaces (leads/open ocean) can be identified as darker areas (lower radiance value) compared to the brighter ice sheets. The images have a spatial resolution of 300 × 300 m (Donlon et al., 2012), which thus limits lead detection to leads that are at least 300 m wide. Nevertheless, the use of these images provides a common ground to compare different classification schemes.

5.2.3. STUDY AREAS AND DATES

The selection of SRAL tracks was based on the different experiments (see section Experimental Set-up) and the availability of cloud-free OLCI images (manually selected). In total, 35 OLCI images and 18,242 SAR waveforms were collected (see Figure 5.2). The polar nights and reduced lighting conditions experienced in most winter months restricted the use of OLCI data to images from March, April, and summer months. For the experiments with the extra open ocean class, an additional track was used that crosses the Atlantic

Ocean from 50° S to 50° N (sensed on August 19, 2021). It was assured that the ocean dataset covers various significant wave heights (SWH; Timmermans et al. (2020)) as the SWH determines the curvature of the leading edge of the SAR waveforms and thereby directly affects some of the waveform features (Fenoglio-Marc et al., 2015). Moreover, it was checked that the data were not corrupted by sea ice or land reflections and therefore, no optical data was required for creating the ground truth data. In total, 10,862 open ocean waveforms were included.

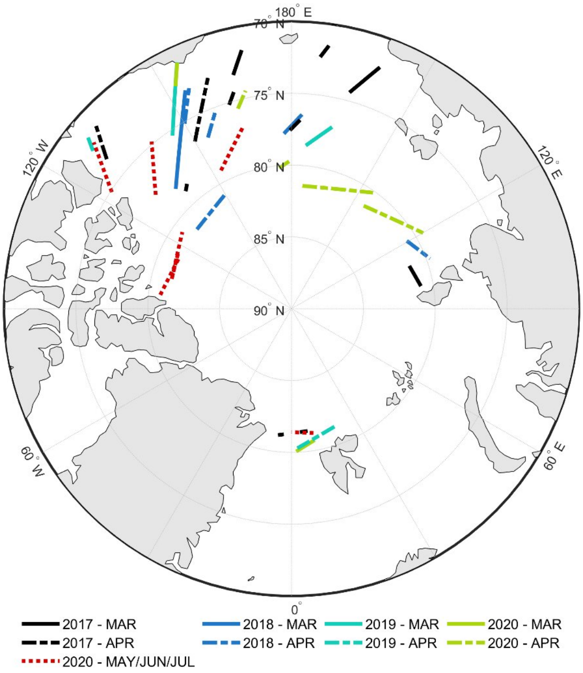


Figure 5.2: The Sentinel-3A/3B SRAL tracks from March 2017 to July 2020 that are used in this paper. Source: Authors.

5.3. METHODS

5.3.1. GENERATION OF GROUND TRUTH DATA

Before testing the classifiers, the SRAL data needed to be labeled according to the ground truth. This procedure relied on a two-step processing of the OLCI product. Firstly, the pseudo-color images were converted to binary images, where each pixel was defined as either lead or sea ice. The following procedure was adopted, inspired by Hamada et al. (2019):

1. The study areas were cropped to smaller sections surrounding the SRAL tracks spanning about 0.60° longitude. This was mainly done to reduce the impact of clouds in unused sections of the images on the image segmentation and instead focus on local along-track radiation differences.

2. The images were segmented using Kmeans clustering, considering two clusters ($K = 2$). This algorithm assigns each pixel to the nearest cluster using the radiance values, while minimizing the total distance to the mean of the clusters. This resulted in binary images.

3. Each point of the SRAL track was assigned a class based on the majority vote of the three closest pixels on the binary images. Three pixels were used to determine the class label because the SRAL data may reflect the surface properties of a combination of pixels, especially when the SRAL data point is located on the edge of a pixel. SRAL data points at the edges of the cropped images were omitted.

Because this approach to image segmentation relies only on local radiance differences, this method allows for an efficient and flexible application to different study areas. It is not necessary to adjust the approach to correct for e.g., differences in lighting conditions. In Figure 5.3, two examples of a segmented binary image (b, d) are shown next to their original images (a, c). In some instances, small-scale irregularities in radiance intensity (e.g., due to the presence of small clouds) caused the image segmentation to incorrectly label some pixels. Therefore, a second step was implemented that labels pixels based on relative along-track changes in radiance, as follows:

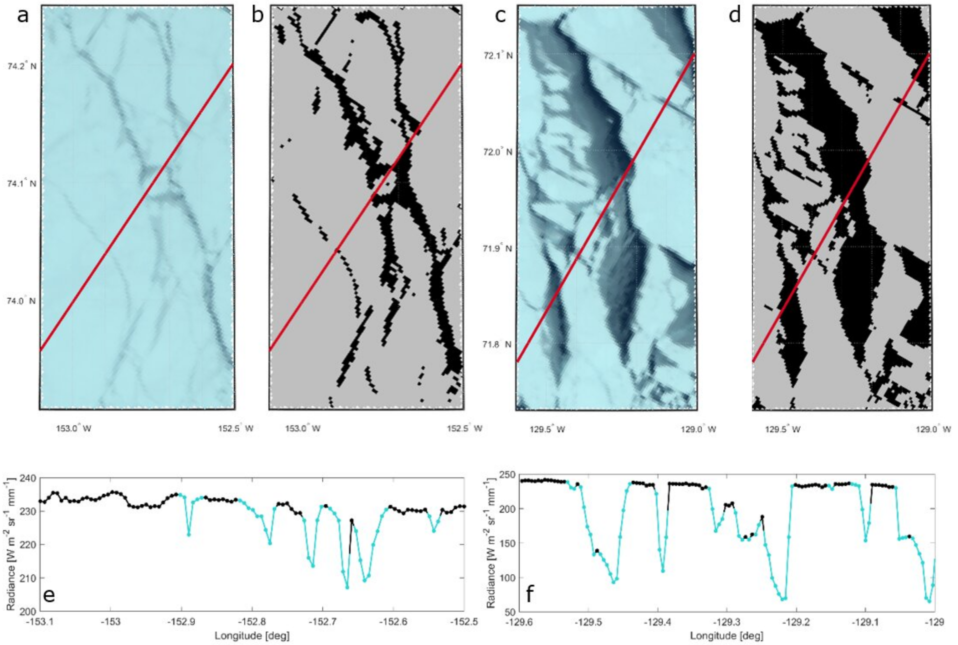


Figure 5.3: Examples of OLCI images (a, c) and their binary images after the image segmentation scheme (b, d). The red line shows the Sentinel-3 ground track. The optical images were taken on 15/04/2018 (a) and 13/04/2019 (c). For both examples, the radiance along the ground track for OLCI band Oa3 is shown below (e, f). The colored parts of the line indicate possible leads.

1. The OLCI radiances were interpolated to each SRAL data point.

2. Maximum and minimum peaks in the radiance series were identified as data samples that are, respectively, higher or lower than both neighboring samples.

3. Minimum peaks whose value does not exceed 2% (empirically determined) of the value of the preceding maximum peak were considered to be leads (see Figure 5.3e, f).

This second step helped to properly label points close to the edge of the leads. SRAL data were only labeled as leads in the ground truth data when this followed from both steps of the procedure. This resulted in a total of 11,762 sea ice and 2,961 lead data points. 3,519 waveforms were rejected.

5.3.2. CLASSIFIER CONFIGURATION AND PERFORMANCE ASSESSMENT

A total of twelve classifiers were assessed in this study, including nine supervised machine learning classifiers, two unsupervised machine learning classifiers, and one thresholding classifier (listed in the next section).

WAVEFORM CLASSIFIERS

The following classifiers were implemented in MATLAB®, using functions from the Statistics and Machine Learning toolbox and the Classification Learner application.

Supervised Machine Learning Classifiers

Supervised machine learning classifiers infer a model from labeled training data, consisting of waveform features (cohesively referred to as the feature space) and the corresponding classes. Subsequently, the model is used to predict the classes of a new set of waveform features. The following nine supervised learning classifiers were tested in this study:

- Decision tree-based classification: various tree-based models have been applied to the classification problems and have shown promising results in many remote sensing applications (Shu et al., 2020; Xu et al., 2014), including lead detection (Lee et al., 2016). One can distinguish between single decision tree classifiers and tree-based ensemble classifiers that combine many classification trees for better and more robust predictions (Hastie et al., 2009). Ensemble tree classifiers are built with many decision trees in parallel (bagging) or sequence (boosting). For this paper, we implemented four types of tree-based classifiers with varying model complexities:
  - The decision tree (DT) consists of a recursive partition of the input data in a single tree-like structure. At each split, a decision is made based on the input features and the new branches represent the possible outcomes, ultimately leading to the final class labels. DT algorithms develop conditions at each split such that the error of class labels is minimized and a meaningful relationship between a class and the values of its features can be captured

  (Quinlan, 1986). The tuning parameters of this classifier include the maximum number of splits and the type of split criterion that is used to evaluate the effectiveness of a split (Tangirala, 2020).
  - Bootstrap Aggregation (Bagged) relies on building many decision trees based on random subsets of the training data (bootstrapping). The final classification is then determined using the average of all predictions from different trees. This reduces the sensitivity to the training data, and reduces variance and over-fitting compared to DT and boosting ensembles (Breiman, 1996). The tuning parameters include the maximum number of splits, number of trees, and learning rate.
  - The Adaptive Boosting (AdaBoost) is an ensemble classifier that attempts to improve the model by iteratively combining DTs (Freund and Shapire, 1999). With each iteration, AdaBoost assigns higher weights to misclassifications. In contrast to the Bagged classifier, the boosting method increases the complexity of the model to primarily reduce the bias and reduce any under-fitting of the training data (Breiman, 1996). The tuning parameters include the maximum number of splits and the number of trees.
  - RUSBoost is another boosting method that applies random undersampling (RUS) of the data. Samples from the larger class are randomly removed to ensure a given ratio between the amount of data per class. This improves classification performance, especially for data sets with uneven class sizes (Seiffert et al., 2008). This method can be promising due to the smaller number of leads compared to sea ice data points in the dataset. Tuning parameters include the maximum number of splits, number of trees, learning rate, and class ratio.

- Artificial Neural Network (ANN): ANN is one of the most popular machine learning methods and has been earlier applied to waveform classification (e.g., Poisson et al., 2018; Shen et al., 2017). In this paper, a simple feedforward network was used that consists of an input layer (observations), a few hidden layers, and an output layer (assigned classes). Each layer consists of several so-called neurons that are connected with neurons from adjacent layers (Grossi and Buscema, 2007). The tuning parameters include the number of layers, layer size, and the activation function (e.g., ReLu, tanh, or sigmoid (Zhang et al., 2018)).

- Naive Bayes Classifier (NB): Bayesian classifiers determine the probability of the occurrence of a class based on a particular set of waveform features (Friedman et al., 1997). The NB classifier is a simple form of a Bayesian classifier that assumes all waveform features to be conditionally independent of each other. This assumption is untrue in many real-life problems (amongst which the problem at hand: see Appendix J), yet the NB classifier has performed excellently in many applications, including waveform classification (e.g., Shen et al., 2017; Zygmuntowska et al., 2013). The tuning parameters include the predictor distribution.

- Linear Discriminant (LD): In LD Analysis, the set of features that are used as input for the classifier (e.g., waveform features as in Table 5.1) is dimensionally reduced, such that variability between the classes is maximized, while variability within the classes is reduced (Qin et al., 2005). Subsequently, classes are assigned based on linear boundaries drawn within the new feature space. The tuning parameters include the type of covariance matrix (full or diagonal) that is estimated from the features.

- Support Vector Machine (SVM): in SVMs, the original feature space is non-linearly transformed into a higher dimensional feature space. Then, an optimal hyperplane that separates the data from different classes is found (Xu et al., 2014). Tuning parameters include the kernel function that is used to transform the data (i.e. linear, polynomial, or Gaussian (Savas and Dovis, 2019)), kernel scale (scaling parameter for the input data), and box constrained level (penalty factor for misclassification).

- Nearest Neighbors (KNN): KNN finds the k number of data points of which the features are closest to the point to be classified. Then, this point is given the majority class of the k closest points (Shen et al., 2017). This classifier is widely used because of its simplicity. Tuning parameters include the number of neighbors (k) and the distance metric that is used to determine the distance between data points.

Unsupervised Machine Learning Classifiers

Unsupervised machine learning algorithms do not require the input data to be labeled but cluster the dataset based on similarities in waveform features. This is particularly beneficial in case the generation of ground truth data is relatively time-consuming or restricted by data availability, such as in the problem at hand. However, unsupervised classification requires the user to manually assign a class to each cluster (Dettmering et al., 2018). For this study, the following unsupervised learning classifiers are adopted:

- Kmedoids: This type of clustering has been successfully applied to lead detection from SAR altimetry by earlier studies (e.g., Dettmering et al., 2018; Muller et al., 2017). This classifier essentially breaks up the data in K clusters while minimizing the distance of all data points to the center of the cluster (called the medoid) (Kaufman and Rousseeuw, 1987). Kmedoids clustering assumes that for each cluster, the distribution of the data in the feature space is spherical (i.e. the variance of different features is of similar magnitude) and cannot handle otherwise shaped clusters (Bindra and Mishra, 2017). The number of clusters (K) is the only tuning parameter for this classifier.

- Agglomerative Hierarchical Clustering (HC): This clustering technique initially considers all individual samples as a cluster and iteratively merges the two closest clusters. The linkage function describes the distance between the clusters. This can for instance be the smallest/furthest distance between individual elements or the group average distance between the two clusters (Murtagh and Contreras, 2012). The tuning parameters include the final number of clusters (K) and the linkage function. This type of clustering can better handle non-spherical data but is typically more time-consuming and more difficult to optimize than Kmedoids (Bindra and Mishra, 2017).

Thresholding Classification

A thresholding method uses threshold values for each waveform feature (see Table 5.1) to classify the samples. This method has been widely used for lead detection in the past (S. W. Laxon et al., 2013; Peacock and Laxon, 2004; Rose, 2013; Schulz and Naeije, 2018). The thresholding values are typically selected empirically or determined by solving an optimization problem to maximize the accuracy (Wernecke and Kaleschke, 2015). The latter approach was adopted in this study (see Appendix G for more details).

CLASSIFIER CONFIGURATION

Before assessing classification performance, the classifiers were configured by finding the optimal set of waveform features (Table 5.1) and classifier tuning parameters (see section Waveform Classifiers) using an iterative procedure. First, the classifiers were optimized individually by using all waveform features and adjusting the tuning parameters to maximize overall classification performance (see also section Classifier Performance Assessment). Then, using these optimized classifiers, the set of waveform features with the best predictive capacity was selected by running the classifiers with different combinations of features and again assessing the classification performance. Once the final waveform features were selected, the classifier settings were again tuned to maximize their performances. The final input settings (both waveform features, and classifier tuning parameters) were kept constant throughout the study. After this configuration phase, a preliminary comparison of the classifiers was conducted to limit the number of classifiers to be compared. Classifiers that performed significantly worse than the others were omitted at this stage.

CLASSIFIER PERFORMANCE ASSESSMENT

During subsequent phases of the experiment, classifier assessment consists of a training and testing phase. For the assessment of supervised classifiers (SUP), the total considered dataset is split into training data (typically 80% of the dataset) and testing data (20%). During the training phase, the total misclassification cost was minimized, where equal weights were assigned to the misclassification of all considered classes. A five-fold cross-validation technique was used to quantify the training performance (Xu et al., 2014). With this technique, the training dataset was randomly divided into five subsets, of which four were combined and acted as the training set while the remaining one acted as the testing set. This was repeated five times and the final performance score was determined by averaging the classification accuracy of the five iterations (Xu et al., 2014). The training phase for the unsupervised machine learning algorithms consisted of applying the algorithm to the training dataset. The classifier subsequently found the clusters and assigned a class to each cluster based on user input.

During the testing phase of the supervised classifiers, the trained models were applied to the testing data set. For the unsupervised classifiers, the KNN algorithm was applied to the testing dataset such that to each sample, the class of the closest cluster of the training dataset was assigned. For the threshold classifier, there was no training phase as the data were directly classified on the basis of the optimized thresholds (Appendix G).

The performance of both the supervised, unsupervised, and threshold classifiers was assessed according to the following measures:

1. The overall accuracy is defined as the total number of correct classifications as a percentage of the total number of samples.

2. The true positive rate, or sensitivity, measures the proportion of positives that are correctly classified (Flach 2016). In this study, the True Lead Rate (TLR) is used, which takes leads as the positive class. The True water Rate (TwR) is also used, which takes all water surfaces (leads and open ocean) as the positive class. Ideally, the TLR and TwR should be as high as possible.

3. The false positive rate measures the number of samples that are incorrectly classified as positive over the total number of positive data points according to the ground truth data (Flach, 2016). In this study, the False Lead Rate (FLR) and the False water Rate (FwR) were used. To reduce SSH errors due to misclassification of water surfaces, minimal FLR and FwR are pursued.

4. Receiver Operating Characteristic (ROC) graphs show a trade-off between the true positive rate and the false positive rate. The area under the curve (AUC) provides an overall performance measure. An AUC value of 1.0 suggests a perfect classifier, whereas an AUC value of 0.5 indicates the classification is equivalent to random guessing (Flach, 2016). For supervised learning classifiers, ROC graphs show the classification results as a function of the decision threshold that is used when assigning data to different classes based on the probability that the sample belongs to that class (Flach, 2016). ROC graphs for unsupervised learning and thresholding classifiers were generated by adjusting the tuning parameters (see section Classifier Configuration) (e.g., Dettmering et al., 2018; Muller et al., 2017).

5.3.3. EXPERIMENTAL SET-UP

To assess the performance of the different classifiers under varying circumstances, the tuned classifiers were applied to different data. The specific division of the data is summarized in Table 5.2. Note that the division depends on the algorithms that are used. The general performance during the winter (MAR/APR) was analyzed using the D-01 data set. Datasets D-02 and D-03 were used to assess the impact of using training data from, respectively, a different year or different areas on the classifier performance. This has practical relevance for the supervised machine learning algorithms because the availability of ground truth data is typically limited. Summer performances were studied with the D-04 data set that includes only data from summer months (MAY/JUN/JUL). Finally, the influence of the inclusion of observations from the open ocean was studied with the D-05 data set. If not otherwise specified, the training and testing data were created by randomly selecting 80% and 20% of the original data set. As unsupervised learning and thresholding classifiers do not require any labeled data, these algorithms were directly applied to the testing data, such that their performances could be compared to supervised learning classifiers.

Table 5.2: Description of data divisions for different test cases.

Name	Purpose	Description
D-01	Assess the general performance of classifiers	Randomly takes 80% data for training and 20% data for testing from all data. Excludes 2020 – MAY/JUN/JUL (see Figure 5.2) and the Atlantic Ocean track.
D-02	Analyze possible temporal biasing (interannual)	**Supervised**: uses 2017 – MAR and 2017 – APR as training data to test the rest of the data. Excludes 2020 – MAY/JUN/JUL and the Atlantic Ocean track. **Unsup./Threshold**: Apply the algorithm to the test data set.
D-03	Analyze possible regional biasing	**Supervised**: Uses data from below 80°N and between 150° E - 240° E as training to test data lying above 80°N and between 120° E - 150° E. Excludes 2020 - MAY/JUN/JUL and the Atlantic Ocean track. **Unsup./Threshold**: Apply the algorithm to the test data set
D-04	Analyze possible temporal biasing (seasonal)	Randomly takes 80% data for training and 20% data for testing from 2020 – MAY/JUN/JUL
D-05	Analyze the impact of additional open ocean class	Randomly takes 80% data for training and 20% data for testing from all data. Excludes 2020 – MAY/JUN/JUL but includes the Atlantic Ocean track.

5.4. RESULTS

5.4.1. CLASSIFIER CONFIGURATION

Following the optimization procedure described in section Classifier Configuration applied to all winter data (i.e., D-01, Table 5.2), the following waveform features were selected to be used in the remainder of the analysis: MAX, skew, ww, PP, and PPloc. Considering each of these features individually produced consistently high accuracies for all classifiers (see Appendix H). The combination of the five features further improved the performance, yet the addition of more features had little effect (Figure H.1).

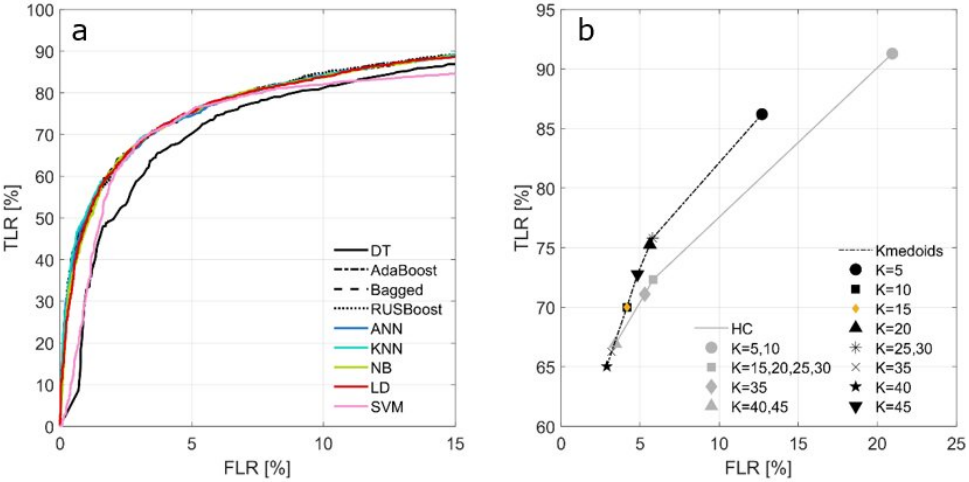


Figure 5.4: ROC graphs of supervised (a) and unsupervised (b) classifiers during the training phase of the D-01 data division. The ROC graphs of supervised classifiers are obtained by varying the discrimination threshold (see section Classifier Performance Assessment). The ROC graphs of the unsupervised classifiers are obtained by adjusting the number of clusters (K) as shown in the legend (b). Note the different limits on the axes.

With this set of five waveform features, the optimal settings for each classifier were determined (as described in section Classifier Configuration). The optimal number of clusters (K) for the unsupervised classifiers was determined based on the ROC graphs as depicted in Figure 5.4b. This figure shows that a small number of clusters ($K = 5$) relates to a high FLR and low TLR, while for a higher number of clusters ($K > 5$), there is no clear correlation between cluster size and classifier performance. The HC classifier produces the same results for some of the clustering sizes (e.g., $K = 5, 10$ or $K = 15, 20, 25, 30$), indicating this classifier is inflexible. The selected settings for all classifiers are displayed in Table 5.3.

Based on the ROC graphs obtained during classifier configuration (Figure 5.4), the DT and SVM classifiers were excluded from the analysis since they produced lower AUC values (0.88 for SVM and 0.91 for DT) than the other supervised classifiers (all 0.94). Regarding the unsupervised classifiers, the Kmedoids classifier shows the best results as it

Table 5.3: Classifier settings, where n is the size of the training dataset and f the number of predictors.

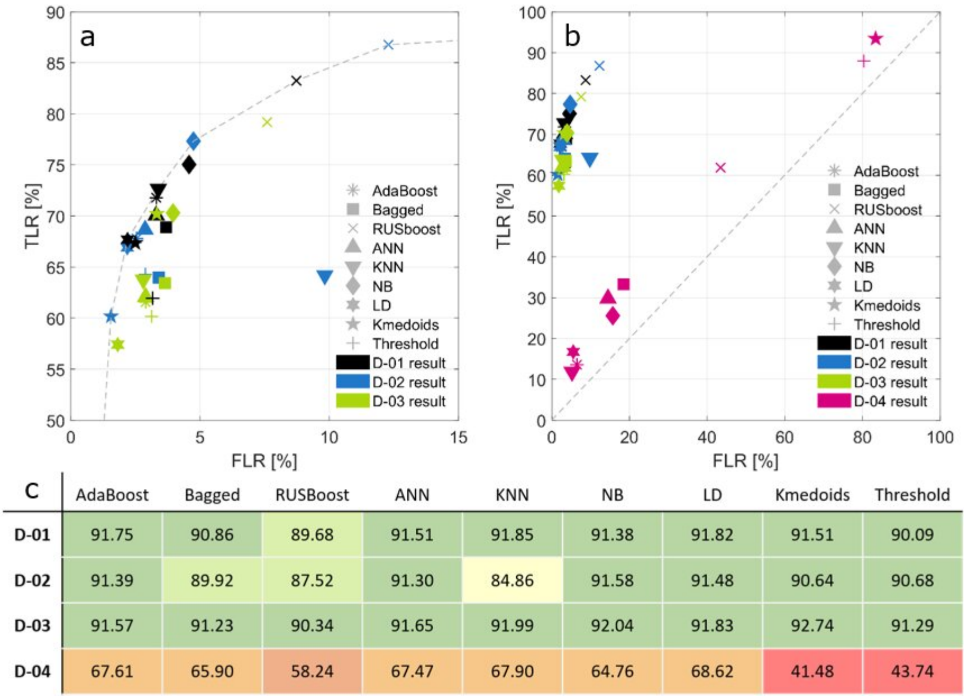
Classifier	Settings
DT	Maximum number of splits = 100, Split criterion = Gini index
Bagged	Maximum number of splits = $n-1$, Number of trees = 30
AdaBoost	Maximum number of splits = 100, Number of trees = 30, Learning rate = 0.1
RUSBoost	Maximum number of splits = 20, Number of trees = 30, Learning rate = 0.1, Class ratio = 1:1(:1)
ANN	Fully connected layers = 1, Layer size = 10, Activation = ReLU, No regularization
NB	Predictor distribution = Gaussian
LD	Covariance structure = Full
SVM	Kernel function = Gaussian, Kernel scale = $\frac{1}{4}\sqrt{f}$, Box constrained level = 1
KNN	Number of neighbors $k = 100$, Distance metric = Euclidean
Kmedoids	Number of clusters $K = 15$
HC	Number of clusters K = 40, Linkage = farthest distance
Threshold	Classify as leads if MAX > 3000 counts, PPloc > 0.55, ww < 45 bins, PP > 0.24, and skew > 7. Classify as ocean if: MAX: 500-1500 counts, PPloc: 0.2-0.35, ww: 85-110 bins, PP < 0.1, and skew: 1.5-3.5. Else: sea ice (see Table 5.1 for abbreviations).

consistently obtains higher TLR values than the HC classifier (Figure 5.4b). Therefore, the HC classifier was also excluded from further analyses.

5.4.2. CLASSIFICATION PERFORMANCES FOR WINTER DATA (D-01 - D-03)

Classification performances obtained from the test cases with data from winter months are shown in Figures 5.5a and c. In Figure 5.5a, the best performances are connected to create an optimal front. Classifiers on this front obtain the highest TLRs for a given range of FLRs and are therefore perceived to outperform classifiers that are located away from the front. The best choice out of all classifiers that are located on the front depends on user preferences regarding the FLR. KNN, AdaBoost, and LD show the best results from the general performance test (D-01) as they are all located on the optimal front. However, there is little difference between these three and most other classifiers. The Threshold and Bagged classifiers are located away from the optimal front, indicating that they perform worse than the others. While RUSBoost is located on the optimal front with the highest TLR, it also achieved the highest FLR, which explains the lower overall accuracy (Figure 5.5c).

Most of the supervised learning classifiers do not suffer significantly when they are trained with data set from another year (D-02). Most classifiers are still part of the optimal front or very close to it. However, the Bagged and KNN classifiers perform poorly when trained with data set from another year. Lastly, the test concerning possible re-



C	AdaBoost	Bagged	RUSBoost	ANN	KNN	NB	LD	Kmedoids	Threshold
D-01	91.75	90.86	89.68	91.51	91.85	91.38	91.82	91.51	90.09
D-02	91.39	89.92	87.52	91.30	84.86	91.58	91.48	90.64	90.68
D-03	91.57	91.23	90.34	91.65	91.99	92.04	91.83	92.74	91.29
D-04	67.61	65.90	58.24	67.47	67.90	64.76	68.62	41.48	43.74

Figure 5.5: ROC graph showing classification performances for D-01, D-02, D-03 (a), and D-04 (b) (see description of test cases in Table 5.2). The marker style depicts the classifier, and the colors are used to differentiate between different study cases. The grey line in (a) shows the optimal front, while the grey line in (b) simply connects FLR=TLR = 0 and FLR=TLR = 100. Note the different limits on the axes. Overall classification accuracies [%] are shown in (c).

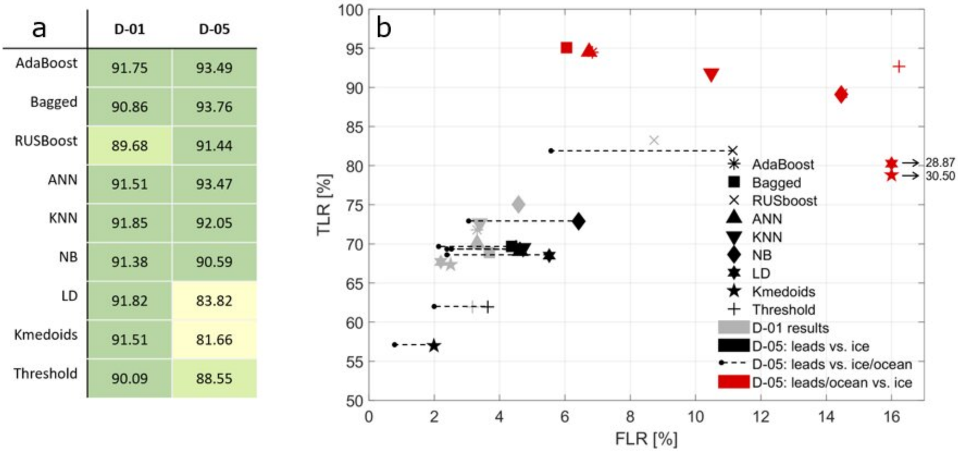
gional biasing (D-03) resulted in improved overall accuracy for most of the classifiers (up to 1.2%, Figure 5.5c). However, the performance of the supervised classifiers was reduced in terms of lead detection (Figure 5.5a). This indicates that these classifiers may perform slightly worse when they are trained with data sets from different study areas.

5.4.3. CLASSIFICATION PERFORMANCES FOR SUMMER DATA (D-04)

All classifiers perform relatively poorly when applied to summer data (Figures 5.5b and c). Classifiers in the lower FLR range (<20%) produce very low TLRs (under detection): AdaBoost, Bagged, ANN, KNN, NB, and LD, while classifiers with a higher TLR also produce high FLRs (>40%; overdetection): RUSBoost, Kmedoids, and Threshold. All classifier performances are very close to the diagonal line in Figure 5.5b, indicating that the classifiers are not much better than random guessing. Especially, the Kmedoids and Threshold classifiers produce extremely low overall accuracies (Figure 5.5c).

5.4.4. CLASSIFICATION PERFORMANCES WITH ADDITIONAL OPEN OCEAN CLASS (D-05)

Finally, to allow for increased reliability of SSH estimation by the inclusion of open water samples, the classifiers were additionally assessed with consideration of this third class. When overall accuracy is considered, most classifiers do not suffer from the addition of the open ocean class (Figure 5.6a). Only the LD and Kmedoids classifiers perform significantly worse in the D-05 test case. However, regarding the distinction between leads and sea ice waveforms, all classifiers perform slightly worse (higher FLRs; Figure 5.6b). This effect is most pronounced for LD, NB, RUSBoost, and especially Kmedoid. The TwRs are relatively high for most classifiers except for LD and Kmedoids, which also produce very high FwRs. In all cases, the real FLRs, which compares both sea ice and ocean waveforms against leads (Figure 5.6b: dashed line+dot), are lower than when ocean data are excluded from this measure. This indicates that a negligible number of ocean waveforms are misclassified as leads.



a	D-01	D-05
AdaBoost	91.75	93.49
Bagged	90.86	93.76
RUSBoost	89.68	91.44
ANN	91.51	93.47
KNN	91.85	92.05
NB	91.38	90.59
LD	91.82	83.82
Kmedoids	91.51	81.66
Threshold	90.09	88.55

Figure 5.6: Overall classification accuracies [%] for D-01 and D-05 (a), complemented by ROC graphs showing classification performances (b). The D-05 results are plotted in three ways: using the TLR and FLR that compare lead classifications to sea ice classifications (black) to allow direct comparison with the D-01 results, the actual TLR and FLR that compare lead classifications to sea ice and ocean classifications (dashed line+dot) and the TwR and FwR (red) that compare lead and ocean classifications to sea ice classifications. Note that the FwRs of LD and Kmedoids are outside plot limits.

5.5. DISCUSSION

A wide range of classification methods was assessed for lead detection in the Arctic Ocean from Sentinel-3 satellite data. The classifiers were applied to SRAL data, while simultaneously sensed OLCI images were used for creating the ground truth data. For the latter, an automatic validation process was implemented that uses Kmeans image segmentation and along-track changes in radiance. This novel approach of using OLCI imagery was particularly useful because of the perfect temporal alignment between the two involved datasets. Disadvantages of using OLCI images are the dependency on illuminated and cloud-free conditions and the relatively low spatial resolution (compared to e.g., Operation Ice Bridge imagery (Dettmering et al., 2018). Even though the spatial resolution of OLCI imagery equals the along-track resolution of SRAL data, a narrow lead may cause a specular SRAL waveform, while it would not be visible on the OLCI image. This would result in (seemingly) overdetection of leads, i.e., higher FLRs.

In total, nine supervising machine learning algorithms, two unsupervised machine learning algorithms, and a threshold classifier were applied to various test cases. This provides a comprehensive understanding of the performance of different classifiers and their applications. In this study, where the goal of lead detection is to improve the Arctic Ocean SSH estimation, a low FLR is treated as the most important classifier criterion. Misclassifications may result in large errors in the SSH estimation (see Appendix J, for an example). While the RUSBoost classifier produced results close to the optimal front for each winter test case (D-01-D-03; Figure 5.5a), its FLR values are very high compared to other classifiers. The RUSBoost classifier obtains such high FLR and TLR values because this algorithm opts to increase the correct classifications for the minor class (Seiffert et al., 2008), leading to overdetection of leads. The latter makes the classifier less suitable for SSH estimation. The Kmedoids classifier performed consistently well for the winter data (D-01-D-03). Because this is an unsupervised classifier, it can be directly applied to the data of interest, regardless of the available ground truth data. Hence, any consequences from temporal differences between training and testing data for the performance of supervised classifiers, do not apply to unsupervised classifiers. Therefore, if the ground truth data are unavailable for the testing area, using the Kmedoids classifier may be preferred. However, this classifier requires the user to manually assign waveform clusters to surface types, which is a disadvantage when a priori knowledge of the different waveform types belonging to certain classes is lacking. If sufficient ground truth is available, the supervised learning classifiers AdaBoost, LD, or ANN are preferred over the Kmedoids classifier, as their general performances were slightly better (up to 0.31%), and the classification does not require manual class assignment. Additionally, when large amounts of data are considered, the time complexity of the different classifiers favors the use of supervised classification. In the case of unsupervised classification, the time complexity is typically quadratically (Kmedoid) or cubically (HC) dependent on the number of observations (Bindra and Mishra, 2017; Whittingham and Ashenden, 2021), compared to a predominantly linear dependency for the supervised classifiers (e.g., Cai et al., 2008; Deng et al., 2016; Fleizach and Fukushima, 1998; Sani et al., 2018). Finally, while the KNN classifier produced one of the best results in the general test case (D-01) and was only marginally affected by a regional bias (D-02), its performance worsened significantly when applied to data from another period (Figure 5.5c). This indicates that the KNN classifier could be very sensitive to a change in the dataset, making the classifier unpredictable. Almost all classifiers performed slightly worse when applied to data from another period than the training data, which argues for the consideration of a training data set that spans the full period of interest. Finally, it was shown that the Bagged classifier did not produce high enough TLRs in any of the winter test cases.

From the fourth test case (D-04) it appears that all classifiers perform poorly when applied to data from summer months (Figure 5.5b). This is likely related to the physical transformation of the sea ice during this period. Most of the altimetry return signals are specular, even when ground truth data suggest that they originate from sea ice (see Figure 5.7e). The presence of melt ponds on the surface of the sea ice may be the cause of this increase in specular returns during summer months. Unevenness in the color of the sea ice on the OLCI images (Figure 5.7d) may indicate melt ponds. This must however be confirmed with images from sensors with higher resolution. Additionally, more diffuse signals are returned from what appear to be leads based on the ground truth data (Figure 5.7f). This may be related to the presence of waves on widening leads, or the presence of separated ice floes smaller than the resolution of the OLCI images. The sole use of SAR waveform features for lead detection in summer months is deemed unsuitable and auxiliary information is required. For instance, in the study by Dawson et al. (2022) local variations in elevation were successfully used in addition to waveform features, to distinguish between SAR returns from leads and melt ponds. However, this does require preliminary retracking of the data before the classification and constrains the variety of leads that can be detected (Dawson et al., 2022). On another note, further research should show to what extent, reduced classifier performance impacts the uncertainty associated with SSH estimates from summer data. In this respect, it should be noted that the occurrence of sea ice, and thus leads, is significantly reduced in summer, resulting in more possibilities regarding SSH estimation from open water and reducing the need for accurate lead detection.

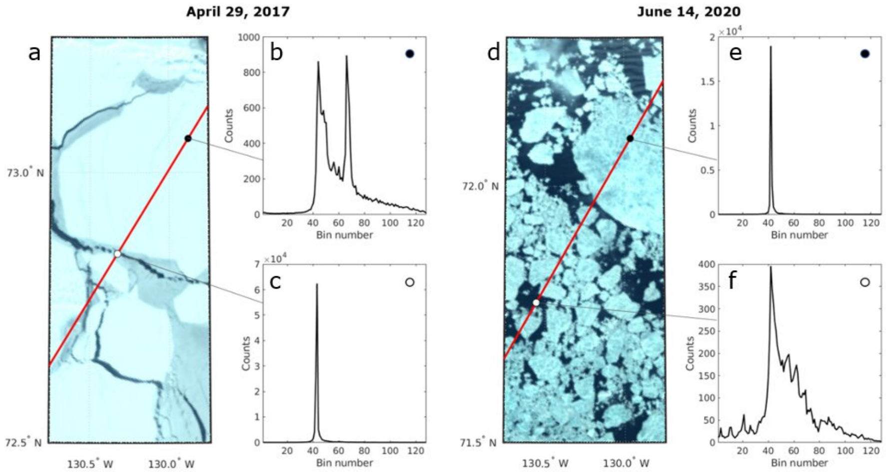


Figure 5.7: Example OLCI images with two typical SAR waveforms for a winter date (a–c) and a summer date (d–f). In both examples, the first waveform (b, e) belongs to sea ice and the second (c, f) to a lead according to the OLCI image, while in the summer case (d–f), the SAR waveforms were incorrectly classified as the opposite class by all classifiers.

Finally, to improve SSH estimation by reliable detection of open water areas, the addition of a third class is required. The addition of the open ocean class had little impact on the overall accuracy of most classifiers, except for LD and Kmedoids (Figure 5.6a). However, the performance of all classifiers decreased slightly in terms of lead detection (Figure 5.6b; TLR and FLR), while the impact on the Threshold and Bagged classifiers was the smallest. If one is purely interested in obtaining as many good water level measures as possible (considering TwR and FwR), the AdaBoost, Bagged and ANN classifiers perform best. The reduced performance of the Kmedoids classifier was associated with the fast increase in data, which cluttered the clusters and complicated manual class assignment. Moreover, the Threshold classifier performs poorly when TwR and FwR are concerned.

The results produced by this study were compared to results from other studies that tested different classification methods for lead detection from altimetry (Figure 5.8). However, caution is advised when comparing the classifier performances found in different studies. Differences in input data (e.g., SAR or conventional radar, different study dates or study areas), different settings for the classifiers, or different methods for ground truth data generation, impact the obtained classifier performances. For instance, Lee et al. (2016) applied two tree-based supervised machine learning classifiers to SAR altimetry data from CryoSat-2: DT and Random Forest (RF). The obtained classification results show extremely high accuracies and high TLR values compared to the results obtained in this paper. However, they tested their classifiers using only 239 waveforms, hence the classifiers may have been overfitted to this small dataset. Their findings show that the ensemble tree classifier (RF in their case) outperforms the DT classifier, which agrees with the findings presented here. Moreover, Dettmering et al. (2018) applied the unsupervised Kmedoids classifier to Cryo Sat-2 SAR altimetry data and used images from the NASA Operation Ice Bridge mission for validation. They obtained TLRs that were significantly lower than those produced by the Kmedoids classifier in this study. This may be because the resolution of the images from Operation Ice Bridge is 1m (Dettmering et al., 2018), compared to the 300 m along-track resolution of CryoSat-2. This difference most likely resulted in the underdetection of leads from the altimeter data. Furthermore, Wernecke and Kaleschke (2015) classified CryoSat-2 data with threshold optimization, using MODIS images for validation. Their TLRs and FLRs are comparable to the results from this study, however, they only validated the classification with data from the Beaufort Sea.

Where the comparison of classifier performances presented by prior studies may be misleading, this study provides a comprehensive assessment of the relative classifier performances. Nevertheless, there are more classifiers that could be applied to lead detection in the Arctic. For instance, the results from Lee et al. (2016) suggest that the RF classifier performs well, which has not been tested in this study. Furthermore, the use of OLCI images for validation appeared very useful because of their perfect temporal alignment with the SRAL data. However, while the generation of ground truth data has been largely automated in this study, a manual check was required to reject images that were deteriorated by small clouds. This process is time-consuming and future studies may benefit from an improved algorithm for cloud rejection that would also detect small and thin clouds, allowing more ground truth data to be generated. A solution may be the com-

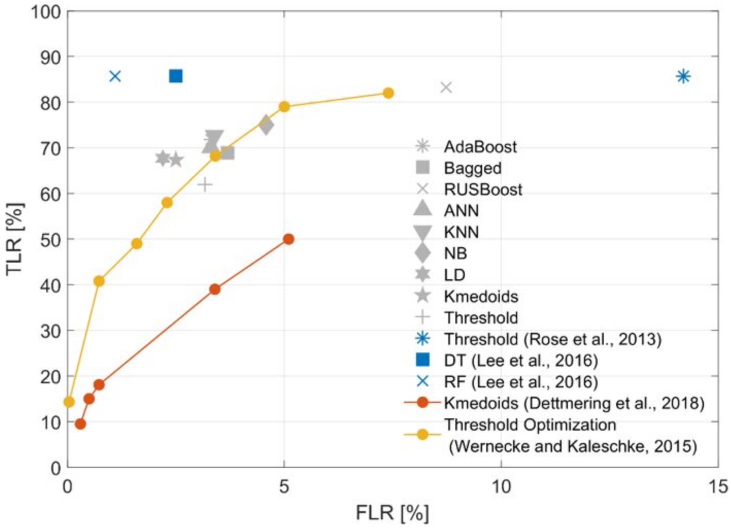


Figure 5.8: ROC graphs showing classification results from previous studies and general performance (D-01) results obtained in this study.

bination of OLCI images and Sea and Land Surface Temperature Radiometer (SLSTR) data, which have recently been used for cloud detection (Fernandez-Moran et al., 2021). This synergy could also be exploited to better distinguish between leads and melt ponds. Likewise, for specific test cases or when additional (sub)classes are considered, the classification may benefit from a different/extended set of waveform features. For instance, it was found that sea ice waveforms sometimes resemble open ocean waveforms (see Appendix I), which would cause large errors in the SSH estimation. For most of the data, a clear regional separation between sea ice and the open ocean can be assumed. Therefore, the initial set of waveform features may be extended by a certain along-track history parameter. For instance, the addition of the moving standard deviation of the pulse peakiness (see Table 5.1) appeared to significantly improve the three-class classification performances (Appendix I). Finally, it should be acknowledged that the predictors used in this study are to some degree correlated with each other (see Appendix J). While this is something that should generally be avoided in any statistical model (including classifiers) (Blalock, 1963), Figure J.1 shows that the correlation within the set of predictors is largely consistent across different data divisions and is therefore expected to have a limited effect on the quality of the classifiers.

5.6. CONCLUSIONS

This paper provides a thorough assessment of twelve different waveform classification methods, applied to Sentinel-3 SRAL data. Here, the perfect temporal alignment between SRAL and OLCI data was successfully exploited for generating the ground truth data. In addition to assessing the general classifier performance, the classifiers were applied to different test cases to analyze the impact of possible regional or temporal biasing and the impact of additional classes.

It was shown that all classifiers performed relatively well on data from March and April (2017–2020). Overall, the AdaBoost and ANN classifiers showed the most robust results throughout the analysis. However, supervised learning requires labeled training data, and thus ground truth data must be available. This study showed that the usage of training data from another study area or a different year slightly worsens the performance of some classifiers, hence the use of a comprehensive training dataset is recommended. Alternatively, the unsupervised machine learning Kmedoids classifier does not require the ground truth data and consistently showed excellent results but performed poorly when tracks that (partly) cover open ocean were considered. Additionally, the interpretation of classifications by Kmedoids is sensitive to differences in user knowledge. Moreover, if large amounts of data are considered, the supervised classifiers may be preferred over unsupervised classifiers as they typically have lower time complexities. Finally, the thresholding method performs worse than the machine learning-based methods yet may still be preferred due to its simplicity in application.

BIBLIOGRAPHY

Bij de Vaate, I., Martin, E., Slobbe, D. C., Naeije, M., & Verlaan, M. (2022). Lead Detection in the Arctic Ocean from Sentinel-3 Satellite Data: A Comprehensive Assessment of Thresholding and Machine Learning Classification Methods. *Marine Geodesy*, *45*(5), 462–495. https://doi.org/https://doi.org/10.1080/01490419.2022.2089412

Bindra, K., & Mishra, A. (2017). A detailed study of clustering algorithms. *2017 6th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO)*, 371–376.

Blalock, H. M. (1963). Correlated independent variables: The problem of multicollinearity. *Social Forces*, *42*(2), 233–237.

Breiman, L. (1996). Bagging predictors. *Machine learning*, *24*(2), 123–140.

Cai, D., He, X., & Han, J. (2008). Training linear discriminant analysis in linear time. *2008 IEEE 24th international conference on data engineering*, 209–217.

Cazenave, A., Hamlington, B., Horwath, M., Barletta, V. R., Benveniste, J., Chambers, D., Döll, P., Hogg, A. E., Legeais, J. F., Merrifield, M., et al. (2019). Observational requirements for long-term monitoring of the global mean sea level and its components over the altimetry era. *Frontiers in Marine Science*, *6*, 582.

Dawson, G., Landy, J., Tsamados, M., Komarov, A. S., Howell, S., Heorton, H., & Krumpen, T. (2022). A 10-year record of Arctic summer sea ice freeboard from CryoSat-2. *Remote Sensing of Environment*, *268*, 112744.

Deng, Z., Zhu, X., Cheng, D., Zong, M., & Zhang, S. (2016). Efficient knn classification algorithm for big data. *Neurocomputing*, *195*, 143–148.

Dettmering, D., Wynne, A., Müller, F. L., Passaro, M., & F., S. (2018). Lead detection in polar oceans-a comparison of different classification methods for CryoSat-2 SAR data. *Remote Sensing*, *10*. https://doi.org/doi:10.3390/rs10081190

Donlon, C., Berruti, B., Buongiorno, A., Ferreira, M. H., Féménias, P., Frerick, J., ..., & Nieke, J. (2012). The global monitoring for environment and security (GMES) sentinel-3 mission. *Remote Sensing of Environment*, *120*, 37–57.

Fenoglio-Marc, L., Dinardo, S., Scharroo, R., Roland, A., Sikiric, M. D., Lucas, B., Becker, M., Benveniste, J., & Weiss, R. (2015). The German Bight: A validation of CryoSat-2 altimeter data in SAR mode. *Advances in Space Research*, *55*(11), 2641–2656.

Fernandez-Moran, R., Gómez-Chova, L., Alonso, L., Mateo-Garcıa, G., & López-Puigdollers, D. (2021). Towards a novel approach for sentinel-3 synergistic olci/slstr cloud and cloud shadow detection based on stereo cloud-top height estimation. *ISPRS Journal of Photogrammetry and Remote Sensing*, *181*, 238–253.

Flach, P. A. (2016). Roc analysis. In *Encyclopedia of machine learning and data mining* (pp. 1–8). Springer.

Fleizach, C., & Fukushima, S. (1998). A naive bayes classifier on 1998 kdd cup. *Fleizach C. & Fukushima S.(1998). A naive Bayes classifier on.*

Freund, Y., & Shapire, R. (1999). A short introduction to boosting-proceedings of the sixteenth international joint conference on artificial intelligence.

Friedman, N., Geiger, D., & Goldszmidt, M. (1997). Bayesian network classifiers. *Machine learning*, *29*(2), 131–163.

Grossi, E., & Buscema, M. (2007). Introduction to artificial neural networks. *European journal of gastroenterology & hepatology*, *19*(12), 1046–1054.

Hamada, M. A., Kanat, Y., & Abiche, A. E. (2019). Multi-spectral image segmentation based on the k-means clustering. *Int. J. Innov. Technol. Explor. Eng*, *9*, 1016–1019.

Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (Vol. 2). Springer.

IPCC. (2021). *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]* Cambridge University Press.

Kaufman, L., & Rousseeuw, P. (1987). Clustering by means of medoids. statistical data analysis based on the l1-norm and related methods, y. dodge ed.

Kwok, R. (2018). Arctic sea ice thickness, volume, and multiyear ice coverage: losses and coupled variability (1958–2018). *Environmental Research Letters*, *13*(10), 105005.

Laxon, S. (1994). Sea ice altimeter processing scheme at the EODC. *International Journal of Remote Sensing*, *15*(4), 915–924.

Laxon, S. W., Giles, K. A., Ridout, A. L., Wingham, D. J., Willatt, R., Cullen, R., Kwok, R., Schweiger, A., Zhang, J., Haas, C., et al. (2013). CryoSat-2 estimates of Arctic sea ice thickness and volume. *Geophysical Research Letters*, *40*(4), 732–737.

Lee, S., Im, J., Kim, J., Kim, M., M. Shin, Kim, K., & L., Q. (2016). Arctic Sea Ice Thickness Estimation from CryoSat-2 Satellite Data Using Machine Learning-Based Lead Detection. *Remote Sensing*, *8*. https://doi.org/doi:10.3390/rs8090698

Muller, F. L., Dettmering, D., Bosch, W., & Seitz, F. (2017). Monitoring the arctic seas: How satellite altimetry can be used to detect open water in sea-ice regions. *Remote Sensing*, *9*, 1–20. https://doi.org/doi:10.3390/rs9060551

Murtagh, F., & Contreras, P. (2012). Algorithms for hierarchical clustering: An overview. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *2*(1), 86–97.

Peacock, N., & Laxon, S. (2004). Sea surface height determination in the Arctic Ocean from ERS altimetry. *Journal of Geophysical Research: Oceans*, *109*, 1–14. https://doi.org/doi:10.1029/2001JC001026

Poisson, J. C., Quartly, G. D., Kurekin, A. A., Thibaut, P., Hoang, D., & Nencioli, F. (2018). Development of an ENVISAT altimetry processor providing sea level continuity between open ocean and Arctic leads. *IEEE Transactions on Geoscience and Remote Sensing*, *56*(9), 5299–5319.

Qin, A. K., Shi, S., Suganthan, P. N., & Loog, M. (2005). Enhanced direct linear discriminant analysis for feature extraction on high dimensional data. *AAAI*, 851–855.

Quartly, G. D., Rinne, E., Passaro, M., Andersen, O. B., Dinardo, S., Fleury, S., & Tsamados, M. (2018). Review of radar altimetry techniques over the Arctic ocean: Recent progress and future opportunities for sea level and sea ice research. *The Cryosphere Discussions*, 1–51.

Quinlan, J. R. (1986). Induction of decision trees. *Machine learning*, *1*(1), 81–106.

Raney, R. K. (1998). The delay/doppler radar altimeter. *IEEE Transactions on Geoscience and Remote Sensing*, *36*(5), 1578–1588.

Ricker, R., Hendricks, S., Helm, V., Skourup, H., & Davidson, M. (2014). Sensitivity of CryoSat-2 Arctic sea-ice freeboard and thickness on radar-waveform interpretation. *The Cryosphere*, *8*(4), 1607–1622.

Rose, S. K. (2013). Measurements of sea ice by satellite and airborne altimetry.

Sani, H. M., Lei, C., & Neagu, D. (2018). Computational complexity analysis of decision tree algorithms. *International Conference on Innovative Techniques and Applications of Artificial Intelligence*, 191–197.

Satellites, C. L. (2011). Surface topography mission (stm) sral/mwr l2 algorithms definition, accuracy and specification. *Ramonville St-Agne, France*.

Savas, C., & Dovis, F. (2019). The impact of different kernel functions on the performance of scintillation detection based on support vector machines. *Sensors*, *19*(23), 5219.

Schulz, A. T., & Naeije, M. (2018). SAR Retracking in the Arctic: Development of a year-round retrackers system. *Advances in Space Research*, *62*(6), 1292–1306.

Seiffert, C., Khoshgoftaar, T. M., Van Hulse, J., & Napolitano, A. (2008). Rusboost: Improving classification performance when training data is skewed. *2008 19th international conference on pattern recognition*, 1–4.

Shen, X., Zhang, J., Zhang, X., Meng, J., & Ke, C. (2017). Sea ice classification using Cryosat-2 altimeter data by optimal classifier–feature assembly. *IEEE Geoscience and Remote Sensing Letters*, *14*(11), 1948–1952.

Shu, S., Zhou, X., Shen, X., Liu, Z., Tang, Q., Li, H., Ke, C., & Li, J. (2020). Discrimination of different sea ice types from cryosat-2 satellite data using an object-based random forest (orf). *Marine Geodesy*, *43*(3), 213–233.

Tangirala, S. (2020). Evaluating the impact of gini index and information gain on classification using decision tree classifier algorithm. *International Journal of Advanced Computer Science and Applications*, *11*(2), 612–619.

Timmermans, B., Gommenginger, C., Dodet, G., & Bidlot, J.-R. (2020). Global wave height trends and variability from new multimission satellite altimeter products, reanalyses, and wave buoys. *Geophysical Research Letters*, *47*(9), e2019GL086880.

Wernecke, A., & Kaleschke, L. (2015). Lead detection in Arctic sea ice from CryoSat-2: quality assessment, lead area fraction and width distribution. *The Cryosphere*, *9*(5), 1955–1968.

Whittingham, H., & Ashenden, S. K. (2021). Hit discovery. In *The era of artificial intelligence, machine learning, and data science in the pharmaceutical industry* (pp. 81–102). Elsevier.

Wingham, D. J., Francis, C. R., Baker, S., Bouzinac, C., Brockley, D., Cullen, R., ..., & Phalippou, L. (2006). CryoSat: A mission to determine the fluctuations in Earth's land and marine ice fields. *Advances in Space Research*, *37*(4), 841–871.

Xu, L., Li, J., & Brenning, A. (2014). A comparative study of different classification techniques for marine oil spill identification using radarsat-1 imagery. *Remote Sensing of Environment*, *141*, 14–23.

Zakharova, E. A., Fleury, S., Guerreiro, K., Willmes, S., Rémy, F., Kouraev, A. V., & Heinemans, G. (2015). Sea ice leads detection using SARAL/AltiKa altimeter. *Marine Geodesy*, *38*, 522–533. https://doi.org/doi:10.1029/2001JC001026

Zhang, H., Weng, T.-W., Chen, P.-Y., Hsieh, C.-J., & Daniel, L. (2018). Efficient neural network robustness certification with general activation functions. *Advances in neural information processing systems*, *31*.

Zygmuntowska, M., Khvorostovsky, K., Helm, V., & Sandven, S. (2013). Waveform classification of airborne synthetic aperture radar altimeter over arctic sea ice. *The Cryosphere*, *7*(4), 1315–1324.

6

CONCLUSIONS AND RECOMMENDATIONS

6.1. CONCLUSIONS

The main goal of this thesis was to assess the spatiotemporal variability of global tide and storm surge characteristics by exploiting the wealth of water level data from satellite radar altimeters. In this chapter, the findings will be recapitulated by (1) addressing the research questions that were posed in Chapter 1 and (2) providing a final reflection on the main research objective.

RQ 1: What is the radar altimetry-derived secular variability in global tides, and the related uncertainty, and how does this compare to the secular variability derived from tide gauge data?

While tide gauges typically measure the sea surface (sub)hourly, the revisit period of satellite radar altimeters varies from 9.9156 (TPJ series) to 369 days (CS2). The relatively low temporal resolution of the data affects the possibilities in using altimeter-derived water levels for tidal analysis. Namely, any signal with a frequency of half the sampling frequency or higher is aliased onto a lower frequency. Hence, tidal analysis of radar altimeter data is essentially always based on an analysis at the alias frequencies. To separate between these frequencies, long data records are required. Consequently, studying secular variability in the tidal signal cannot be done by means of yearly tidal analysis and subsequent curve fitting, as is commonly done with data from tide gauges.

In the study presented in Chapter 2, two alternative methods were implemented to obtain secular trends in four major tides from LRM satellite radar altimeter data (TPJ). The first method (referred to as SegHA) was based on the yearly analysis but treats the time series in blocks of four years instead. In the second method (TintHA), estimation of the linear change in tidal constants was integrated in the tidal harmonic analysis. The first method was relatively easy to implement using the available standard software and allowed for a straightforward implementation of non-linear changes. However, this came at the price of a simplified propagation of the uncertainty. In addition, by processing the data in blocks of four years, the trend fitting was done based on only seven samples, which resulted in reduced significance of estimated trends. Moreover, the aliased signal of one of the major tides (K_1) could not be separated from other tidal signals using only four years of data. The second method, however, allowed processing of the full time series at once and hence was able to estimated trends for all tides of interest. Although both methods obtained very similar results when applied to synthetic and actual satellite altimeter data, the TintHA method resulted in lower uncertainties and, therefore, was favoured over the SegHA method. In addition, the uncertainties were further reduced by combining the data from two crossing tracks (crossovers). The presented study considered both the uncertainties obtained from the tidal software and an alternative computed by means of subsampling model timeseries. From a comparison of both products it appeared that the uncertainties derived from the tidal software were relatively unaffected by changes in data availability and likely underestimated the actual uncertainty.

The analysis of almost 30 years of satellite radar altimetry showed that the amplitudes of the major tides were subject to changes of up to 1 mm/year. In many regions, the estimated trends just exceeded the (model-derived) confidence levels. Whether the secular change concerned an increase or decline varied on a regional to even local basis. The magnitude of changes in the phases varied largely across the domain, with the largest changes observed in the vicinity of amphidromic points. The observed changes were most significant for the M_2 tide, and these generally corresponded to trends observed at nearby tide gauges. However, caution is advised when comparing between trends derived from altimetry and from tide gauges. That is, where changes in tidal dynamics are often the result of local processes (Haigh et al., 2020), we cannot necessarily expect the tidal trends at tide gauges to correspond to those about 50 km offshore, at the location of a satellite crossover. In an attempt to bridge the gap between the two data sources, satellite data were additionally processed along-track for a case study on the northwest European Shelf. Again, similar changes were observed for M_2 for the majority of the domain, while discrepancies were observed near the coasts of Norway and the UK. Fewer similarities were observed for the S_2 tide.

While the comparison of tidal trends from tide gauges and satellite altimetry cannot be stringent, it does provide insights in the added value of using satellite data to study secular changes in global tides. Clearly, in regions where data from tide gauges are limited or lacking (e.g., on the open ocean), the results may contribute to our understanding of the drivers behind the observed changes. For instance, from the altimetry-derived trends it appeared that observed changes in amplitudes may be related to displacement of amphidromic points, a finding impossible to infer from tide gauge records alone. However, the study also showed the weak spots in using satellite data. In particular the low temporal resolution of the data introduced several issues that significantly affected the uncertainty of the estimated trends. Although the issue of aliasing was dealt with by processing sufficiently long periods of data, it followed that the uncertainty of obtained tidal constants related closely to the magnitude of non-tidal water level variability. Therefore, this variability was removed as much as possible by applying the dynamic atmospheric correction (DAC) and mesoscale correction. Nevertheless, it is likely that this did not remove all of the signal and results were deteriorated in regions with significant non-tidal activity (e.g., coastal areas or at the location of strong ocean currents). This was particularly evident in the along-track analysis and comparison to trends derived from nearby tide gauges (e.g., along the coast of UK and Norway). Finally, in the discussion of Chapter 2, it was suggested that the tidal trends derived from satellite data may to some extent be introduced by systematic errors in the processing of the data. For instance, any (erroneous or real) change in the range corrections may translate to the retrieved water levels and be interpreted as changes in the tides. While the range corrections are an important and acknowledged potential error source in water level retrieval from satellite altimetry, no correlation was found between tidal signals in respectively the corrections and the retrieved water level.

RQ2: What is the radar altimetry-derived spatiotemporal variability in global storm surge characteristics, including their time-averaged magnitude and seasonal and secular variability, and how does this compare to the analysis of tide gauge data?

Storm surge analysis from satellite radar altimetry is hampered by the low temporal resolution of the data. With a revisit period of at least 9.9156 days, entire storms may not be captured by the satellite. Therefore, as presented in Chapter 3, the data from eight satellite radar altimeters were stacked on a $5° \times 5°$ grid (up to ~ 66° N/S). Then, a time-dependent generalized extreme value distribution (GEVD) was fitted to the monthly maximum (detided) water levels (that were referenced to the yearly mean sea level). This allowed to study both the spatial variability in time-averaged storm surge water levels, their seasonal cycle, and secular change. Additional scaling was applied to the GEVD to account for the limited (and spatially variable) temporal resolution of the data. This was done using scaling factors that were obtained with a synthetic experiment using reanalysis data. The same analysis was applied to tide gauge data (excluding the stacking and scaling) to allow for a comparison of the results.

The time-averaged surge water levels showed a clear zonal dependency, with higher water levels at higher latitudes. The observed seasonal variability was particular strong on the northern hemisphere but overall, higher surge water levels were observed in the local winter. However, in the tropics, more (local) variability in the seasonal phase was found. This likely reflects the difference between tropical and extra-tropical storms. Finally, we derived moderate secular changes in the average monthly maximum surge water levels (< 1 cm) across the globe. The derived changes were predominantly negative, except for few mid-latitude regions with positive change.

Except for the secular changes, the satellite-derived results showed a good correlation to the information derived from tide gauges (correlation > 0.5), although the tide gauges showed more local variability. The poor correlation for the secular change may be related to the change in water levels being dominated by either changes to tropical or extra-tropical surges. The lower correlation may be an unavoidable consequence of stacking the satellite data over a large area, as it was found that the satellite-derived results correlated better when spatial outliers were removed from the tide gauge-product. In addition, it has been suggested that the satellites may not be able to fully capture the temporal variability in the short-lived tropical storms. Nevertheless, the use of satellite data for storm surge analysis could certainly be a valuable contributor to increasing our understanding of the spatial and temporal variability in storm surge water levels. This in particular applies to the large-scale patterns and dynamics on the open ocean, where tide gauge data are simply not available.

RQ3: What is the SAR altimetry-derived seasonal variability in the Arctic tide and to what extent can this be linked to changes in the Arctic sea ice cover?

The analysis presented in Chapter 1 did not include the Arctic, for the reason that the SAR altimetry-derived record of Arctic water levels was not long enough (about 12 years) to study secular changes in the tides. Instead, Chapter 4 presented a study on the seasonal modulation of the major Arctic tide (M_2) using data from the CryoSat-2 and Sentinel-3 missions. This study combined SAR altimetry-derived water levels from leads and open ocean, that were obtained using a threshold classifier and a surface specific implementation of the SAMOSA retracker. The data were stacked on a $1° \times 3°$ grid to compensate for the poor temporal resolution of the satellites. Thereafter, the seasonal modulation of the tide was derived by processing the full record of data and including satellite constituents of M_2 in the tidal analysis, rather than processing the data on a monthly basis. Together with M_2, the satellite constituents configure into a seasonally modulating M_2 signal. The results showed large variations in tidal amplitude, approaching 0.25 m in certain areas, and phase, up to 45°. The observed change was largely comparable to in-situ observations, although these are particularly limited in the Arctic region (both in spatial and temporal sense), rendering a conclusive comparison almost impossible. This also clearly demonstrates the added value of using satellite data for studying temporal variability in Arctic tides.

The observed seasonal modulation in the M_2-tide were linked to seasonal variation in the Arctic sea ice cover. This was done by means of model simulations whereby the Global Tides and Surge Model (GTSM) was forced by respectively two limiting cases of sea ice. These two cases were the (March) maximum landfast sea ice cover and the (September) absence of landfast sea ice. Subsequently, monthly average M_2 amplitudes and phases were obtained. In addition the average March and September amplitudes/phases were computed from the altimeter-derived modulating M_2-signal. A comparison of the altimetry-derived amplitudes showed that the presence of landfast sea ice generally caused a local amplitude decay but off-site amplitude amplification. In many regions, this corresponded to the modelled effect of the sea ice on tides. However, in certain regions, the positive differences in amplitude (larger in March than in September) were not captured by the model and the modelled phase differences were smaller than observed. This may suggest that although the seasonal variation in the Arctic sea ice is an important contributor to the seasonal modulation of the Arctic tide, other processes play a role too. However, this cannot be stated with certainty as these processes were not included in the study.

Moreover, the study presented in Chapter 4 assessed the impact of variations in the Arctic sea ice on global tides by comparing simulated March and September amplitudes to monthly amplitudes derived from tide gauge records. In several regions, the modelled amplitude differences were of considerable magnitude (up to 0.05 m) and in line with the seasonality observed at tide gauges.

RQ4: Can machine learning classification methods contribute to better lead detection from SAR altimetry in the Arctic Ocean?

A key component of the processing applied to obtain the Arctic water level dataset used in Chapter 4 is a classification of the waveforms. This classification aimed to remove any SAR returns from surfaces other than leads or open ocean. In the study presented in Chapter 4, a traditional threshold classifier was used. Recent studies, however, suggested that machine learning-based methods may result in higher accuracies, although a comprehensive comparison of different classifiers was lacking. Therefore, in Chapter 5, twelve different threshold and machine learning-based classifiers were assessed based on their general performance and in terms of regional and temporal uniformity in application. This study made use of SAR data provided by Sentinel-3 and simultaneously sensed optical images for the generation of ground truth data. In total, one optimized threshold classifier, nine supervised and two unsupervised machine learning classifiers were assessed. In addition, the classification performance using twelve different waveform features was compared.

It was shown that all classifiers performed relatively well on winter data but poor on summer data. Overall, the AdaBoost and ANN (supervised machine learning) classifiers showed the most robust results. However, these supervised classifiers require labeled training data and thus the availability of ground-truth data. It was shown that using training data from another region or period slightly reduced the performance of the classifier. Hence it is recommended to use a training dataset that includes all years of interest and covers a variety of study areas. On the other hand, the unsupervised machine learning Kmedoids classifier does not require ground truth data and also performed well on the classification of sea ice and lead SAR returns, but to a lesser extent on the classification of open ocean returns. Although all classifiers performed poor on lead detection from summer data, it remains unclear to what extent the reduced classifier performance impacts the uncertainty associated with water level estimates from summer data. As the occurrence of sea ice, and thus leads, is significantly reduced in summer, we can derive more water levels directly from open water, which reduces the need for accurate lead detection. In the case of winter data, using one of the better performing machine learning-based classifiers (i.e., AdaBoost or ANN) instead of the traditional threshold classifier would increase the data availability over leads, while the number of false lead detections remains comparable.

6.1.1. REFLECTION ON MAIN RESEARCH OBJECTIVE

Based on the analyses presented in this thesis, it follows that the use of satellite radar altimetry provides a significant contribution in mapping the spatiotemporal variability in global storm surge and tide characteristics.

One the one hand, the wealth of global (LRM) altimeter data has been exploited, for the first time, to study the ocean-wide temporal variability in tidal and storm surge water levels. This has provided additional benefits compared to the analysis of observations from the sparsely distributed tide gauges. Firstly, the use of satellite data filled up data gaps in regions where in-situ data are limited (e.g., along the African coastlines). Secondly, as the satellite data have regular (quasi-)global coverage, it enabled to observe the large scale patterns. This information is valuable in gaining better understanding of the relation between observed changes and climatic processes rather than local factors.

However, it was also shown that in particular the analysis of long-term changes in tides and surges was affected by the low temporal resolution and limited record length of the satellite data. Consequently, the secular analysis was associated with significant uncertainties that in several regions render the satellite data useless. This predominantly affected the tidal analysis in regions with large non-tidal water level variability (e.g., in shallow water or in regions with strong mesoscale variability). As regards storm surges, it was suggested that the satellites may not be able to fully capture the magnitude and temporal dynamics of the short-lived and small-scale tropical storms.

Another recurring issue with the use of satellite altimetry in this thesis concerns the validation of the results. While the novel use of the data in question naturally calls for a validation, there is simply no comparable product available. In this thesis, derived tide/surge characteristics have been compared to observations at tide gauges. However, the distance between a tide gauge and a satellite track (that is, the section of the track where the retrieved water levels were deemed reliable) was often at least 30 km. Where in some instances, an apparent disagreement is observed between, for example, the satellite- and tide gauge-derived secular change, it is uncertain whether this reflects a real spatial difference in secular variability or an issue with the data processing.

On the other hand, the use of SAR altimetry has filled a significant data gap in the Arctic region. Although the record of SAR altimetry (2010-2022) is currently not long enough to study secular variability in tidal or surge water levels, it has been used successfully in the study of the seasonal variability in the Arctic tide. In combination with model simulations, this study has provided an important contribution towards better understanding of the relation between sea ice and tides.

However, the study on the Arctic region, only considered the M_2 tide. Because CryoSat-2 and Sentinel-3 are in sun-synchronous orbits, their data cannot be used to derive information about the solar tides (e.g., S_2). Finally, where we did not achieve to study the (seasonal variability in) Arctic storm surges, it remains uncertain whether the current record of SAR altimeter-derived water levels can be used for this purpose.

6.2. RECOMMENDATIONS

While the work presented in this thesis has made a significant contribution to gaining more insight in the temporal variability in global storm surge and tidal water levels, more is still to gain. The most important recommendations are listed below:

1. Firstly, it is worthwhile to further improve the processing of altimeter-derived water levels, and possibly reduce the uncertainties associated with the derived secular trends, by means of improved removal of non-tidal water level variability. Alternatives for the surge correction and mesoscale correction (introduced in Chapter 2) should be explored. For instance, in coastal waters, it is recommended to use high resolution hydrodynamic models to correct for surge variability instead of the DAC that was applied in this thesis.

2. The study of secular changes in tides may also benefit from the inclusion of data from other satellite missions (e.g., ERS-1, ERS-2, Envisat and SARAL). The resulting increase in data availability would likely reduce the associated uncertainties. However, given the low magnitude of the observed secular trends, even small intermission biases in the range corrections could be easily mistaken for changes in the actual tides and should thus be appropriately accounted for.

3. In addition, it may be of interest to derive temporal changes in tides and surges from coastal altimeter data products (such as X-TRACK;Birol et al., 2017 or X-TRACK/ALES; Birol et al., 2021). In contrast to the conventionally processed altimeter data, these products include data up to 5–10 km from the shoreline. Considering these data would thus allow for a better validation of satellite-derived tide and surge characteristics by means observations from tide gauges.

4. As shown in Chapter 5, the variety of classifiers used for lead detection performed poorly on data from summer months, in terms of the distinction between SAR return signals from leads versus sea ice. Although a vast extent of the Arctic sea ice disappears in summer and hence the majority of the data actually originates from open ocean, it is still recommended to improve the lead detection from summer data. This may for instance be done by including a secondary data source (such as the Sentinel-3 SLSTR data). In addition, it is recommended to implement one of the better performing machine-learning based classifiers that followed from Chapter 5 (e.g., AdaBoost or ANN) to further increase the number of Arctic water level estimates.

5. Furthermore, a pivotal assumption in the presented study on secular changes in tides (Chapter 2) is that long-term changes are linear. However, this does not need to be the case, depending on the driving processes. For example, it was suggested that the long-term change in S_2 may be related to periodic meteorological forcing and hence contains significant non-linear variability on different time scales.

Therefore, we propose to conduct a tidal analysis including a time-varying trend in the harmonic constants, e.g., using the state space method framework (e.g., Harvey, 1990).

6. Only a first step has been made in attributing the observed temporal changes in global tides to physical processes (that is, variations in the Arctic sea ice extent, Chapter 4). It is therefore of interest to extend this analysis and assess the impact of sea ice decline, sea level rise and other physical processes on temporal changes in global tides. Including these processes in a hydrodynamic model may allow for better interpretation of the observed temporal changes. In a similar sense, the impact of temporal changes in Arctic tides and surges on global (coastal) water levels may be assessed by for instance including them as boundary conditions in hydrodynamic models. Such a set-up can also ultimately be used to make long-term predictions on the impact of Arctic sea ice decline on tidal and surge water levels.

7. On another note, a topic that has not been covered by this thesis is the Arctic surge. From Chapter 3, it became clear that the data from RADS cannot be used to study surge water levels in sea ice-affected regions. However, SAR altimetry-derived water levels (Chapter 4) could possibly be used for this purpose. Although the SAR data record is likely too short for an analysis of the secular changes, the analysis could still provide insight in the seasonal variation in Arctic surge water levels and its relation to the sea ice cover. In addition, obtaining more information on the (time-independent) surge water levels would be useful for the validation of hydrodynamic models used for storm surge forecasting.

8. In addition, where this thesis considered the extreme still water levels, the contribution of waves was disregarded. Wave activity has strong spatial and temporal variability and is also affected by climate change, typically through processes that are also associated with changes in storm surge water levels; e.g., changes in the atmospheric circulation and increased fetch in the absence of sea ice (Dodet et al., 2019). Where changes in the mean wave height of about 5-10% are projected across the globe (Morim et al., 2019), it is recommended to include this effect in ESL projections. Recently, there have been successful efforts in deriving waves from satellite altimetry (e.g., Passaro et al., 2021), thereby overcoming the need for in-situ data.

9. Finally, where Chapter 4 assessed the effect of changes in the Arctic sea ice cover on tides, the impact of Antarctic sea ice variability is as of yet unknown. While the focus was on the Arctic Ocean because of its pronounced susceptibility to global warming and poor availability of in-situ data, the potential impact of Antarctic sea ice variability on tidal and storm surge water levels should not be disregarded. Furthermore, it is unknown whether the findings from Chapter 4 can be applied to the Antarctic sea ice as the latter differs in many aspects from that of the Arctic (e.g., thickness, mobility and orientation with respect to coastlines).

6.3. SCIENTIFIC AND SOCIETAL VALUE

Extreme sea levels pose a direct threat to the population in coastal areas that is nowadays mitigated by the use of water level forecasting models and coastal engineering. The occurrence and magnitude of extreme water levels are subject to change under the influence of climate change and this should be taken into account during the design of coastal defenses and risk management strategies. However, the specifics, such as the extent of change, spatial patterns, and driving mechanisms, are encompassed by uncertainty. In particular, the contribution of changes in tides and surges is not well understood on a global scale. In that light, the work presented in this thesis is anticipated to form a significant contribution to better understanding and forecasting of, and ultimately, protection to future ESLs. The following key contributions were identified.

The research that is presented in Chapters 2, 3, and 4 contributes to the assessment of the impact of climate change on ESLs. In Chapter 4, seasonal variation in the Arctic tide was related to differences in sea ice extent. These findings can be used to assess the impact of Arctic sea ice decline on future tides. Chapter 2 and 3 showed, for the first time, the secular changes in tidal and surge water levels with full ocean coverage. Where these full-scale products provide us with a better overview of the dynamics, one could now attempt to link the observed changes to physical processes by means of model simulations.

In addition, the findings from Chapters 2, 3, and 4 can be used to improve the models that are used for forecasting ESLs (such as the GTSM). Where these models typically assume tidal constants to be stationary in time, Chapters 2 and 4 have demonstrated that this is not sufficient. Significant seasonal differences in the amplitude of the M_2 tide were observed in the Arctic (up to 0.25 m) and it has been suggested that the seasonal variation in the Arctic sea ice extent alone could cause amplitude changes up to several centimeters on the global scale. Moreover, significant secular changes in tides were observed across the globe, which add up to several centimeters per tidal constituent over the past three decades. Furthermore, the findings from Chapter 3 could be used to validate surge models on the open ocean, which is currently lacking. This may contribute to improved model performance.

Finally, the assessment in Chapter 5 can be used to improve radar altimeter-derived Arctic sea level variability or estimates of the sea ice thickness (relying on freeboard height). In addition, where there was prior to this work no cohesive dataset of instantaneous Arctic water levels available, the data processed for the study of Chapter 4 is publicly available and can potentially be applied in future studies on the Arctic.

A

SENSITIVITY ANALYSIS OF ESTIMATED CHANGE IN S_2 TIDAL HARMONIC CONSTANTS TO THE IONOSPHERIC ALTIMETER CORRECTION

One of the geophysical corrections that is applied to the TPJ altimetry data is the ionospheric correction. This can be either an altimeter-derived correction or a modeled correction (NIC09 for TOPEX/Poseidon and GIM for Jason1-3) (Scharroo et al., 2016). As demonstrated by Zawadzki et al. (2018), any error in these corrections would cause a signal at the alias frequency of S_2. In addition, they showed that replacing the altimeter-derived correction with the modeled correction alters the amplitude of the S_2 signal by up to 3 mm.

To assess the sensitivity of the analysis described in the paper to the choice of ionospheric correction, an additional experiment was carried out. For this purpose, the TPJ data of ~ 500 random crossovers across the globe were corrected by the model-derived ionospheric correction and the radar-derived correction respectively. Consequently, the data were processed and analyzed as described in the paper (following both the SegHA and TintHA approach). The resulting linear change in S_2 amplitudes was compared by computing the median absolute deviation (MAD). Note that crossovers close to land or sea ice were ignored because the radar-derived ionospheric correction is likely deteriorated there (Fernandes et al., 2014).

It was found that the impact of the ionospheric correction was largest near the equator (MAD of up to 0.08 mm/year) and reduced at higher latitudes (< 0.02 mm/year). This is in line with Figure 7 from Zawadzki et al. (2018). The differences were subsequently interpolated to all crossover locations, multiplied by 1.48 to obtain the standard error, and then combined with the standard errors of the S_2 amplitudes that were computed by UTide (Figure C.1a) and the model-based alternative (Figure C.2a), before multiplying them by the z-score to obtain the final confidence intervals. To get an idea of the impact of the contribution of the ionospheric correction to the uncertainty, one could compare Figures C.1a and 2.6a). As the initial confidence intervals outputted by UTide are the same for all semi-diurnal tides (e.g., S_2 and M_2), the difference between Figure C.1a and 2.6a is solely due to the ionospheric error.

B

ASSESSMENT OF CONFIDENCE INTERVALS USING RANDOM SUBSETS OF TPJ-DATA

To compare the respective performances of the confidence intervals obtained from UTide and those obtained by GTSM, the following experiment was performed. Firstly, the TPJ-data at 500 random crossovers across the globe were processed. Then, these crossover time series were randomly divided into two time series with half the number of measurements. Subsequently, the linear change in tidal constants was estimated from both time series and the absolute difference was computed. This was done 50 times, resulting in 50 differences for each crossover, tidal constituent, and tidal constant. These differences were then compared to the 95% confidence intervals for the differences, following from error propagation of the 95% confidence intervals for the trend estimates. This was done for respectively the product that was computed by UTide (following from this experiment) and the confidence intervals obtained from GTSM. In the case of the latter, the time series were sampled in a similar way as the TPJ-data in this experiment.

An example of the outcome of this experiment is shown in Figure B.1 (for the M_2 amplitude change). As can be seen in Figures B.1b and B.1c, the confidence intervals obtained by GTSM are at least twice as large as those that followed from UTide. For the M_2 amplitude change, on average 60% of the differences in trend estimates are significant using the UTide confidence intervals, while this is only 6% when the GTSM confidence intervals are applied (see Figures B.1d and B.1e for global variability). Similar percentages were obtained for S_2: 60% for UTide and 4% for GTSM. For the diurnal tides, the UTide confidence intervals turned out slightly more appropriate (46%), while the GTSM product was consistently accurate (6%).

Concerning the phase change, both confidence intervals perform mediocrely: 54% of the differences are significant following the UTide product, while this is 53% for GTSM (all tides combined). As was also hypothesized in the paper, the fact that the locations of the GTSM reanalysis data are not aligned with the TPJ-data and the location of amphidromic points may affect the quality of the computed confidence intervals, in particular for phase change.

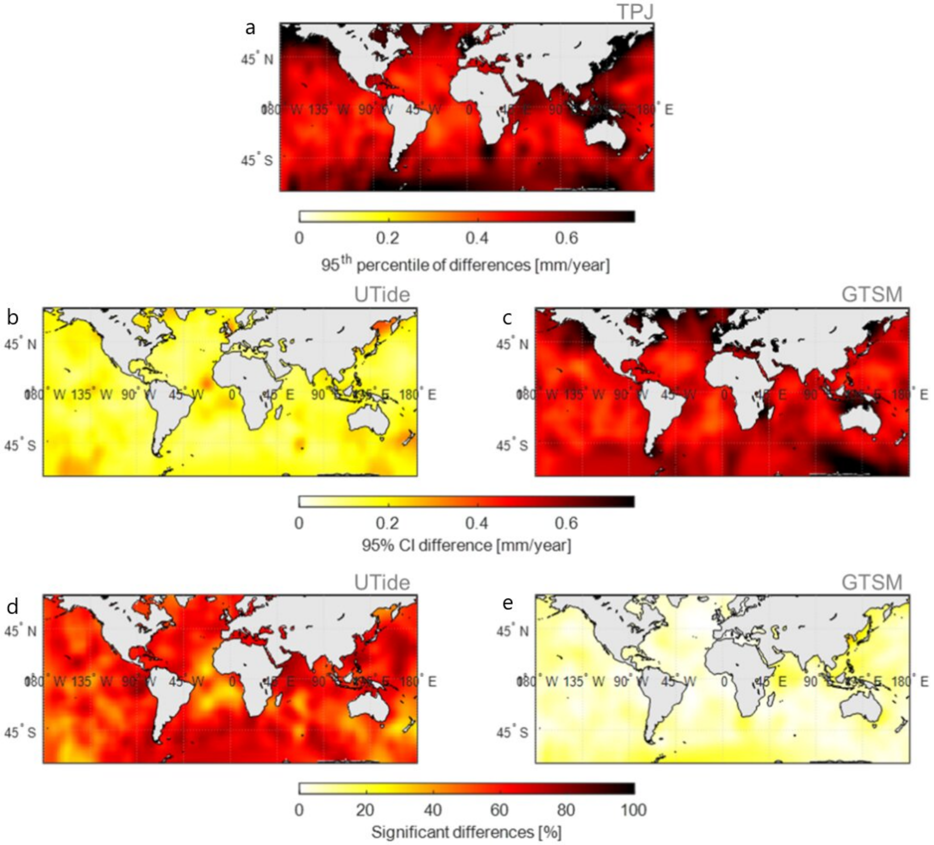


Figure B.1: 95th percentile of the differences between M2 amplitude change estimated from the 50 data division pairs (a). 95% confidence intervals for these differences, as obtained from UTide (b) and GTSM (c). Percentage of the differences that are significant following the UTide (d) and GTSM confidence intervals (e).

C

SUPPORTING FIGURES FOR CHAPTER 2

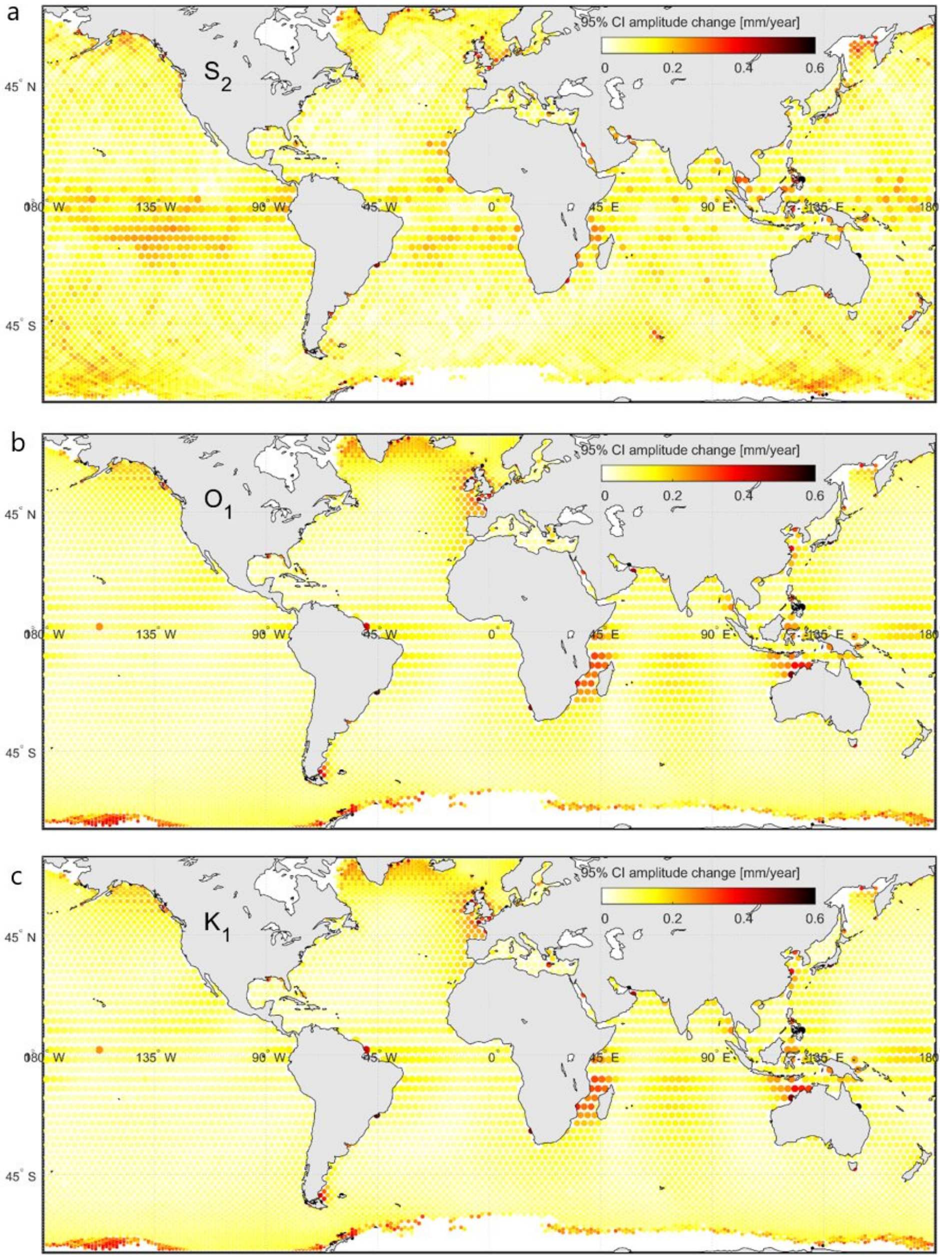


Figure C.1: 95% confidence intervals for trend estimates derived from amplitude standard errors computed by UTide following the TintHA approach, for S_2 (a), O_1 (b) and K_1 (c). The intervals for S_2 include the sensitivity to the ionospheric correction as explained in Appendix A.

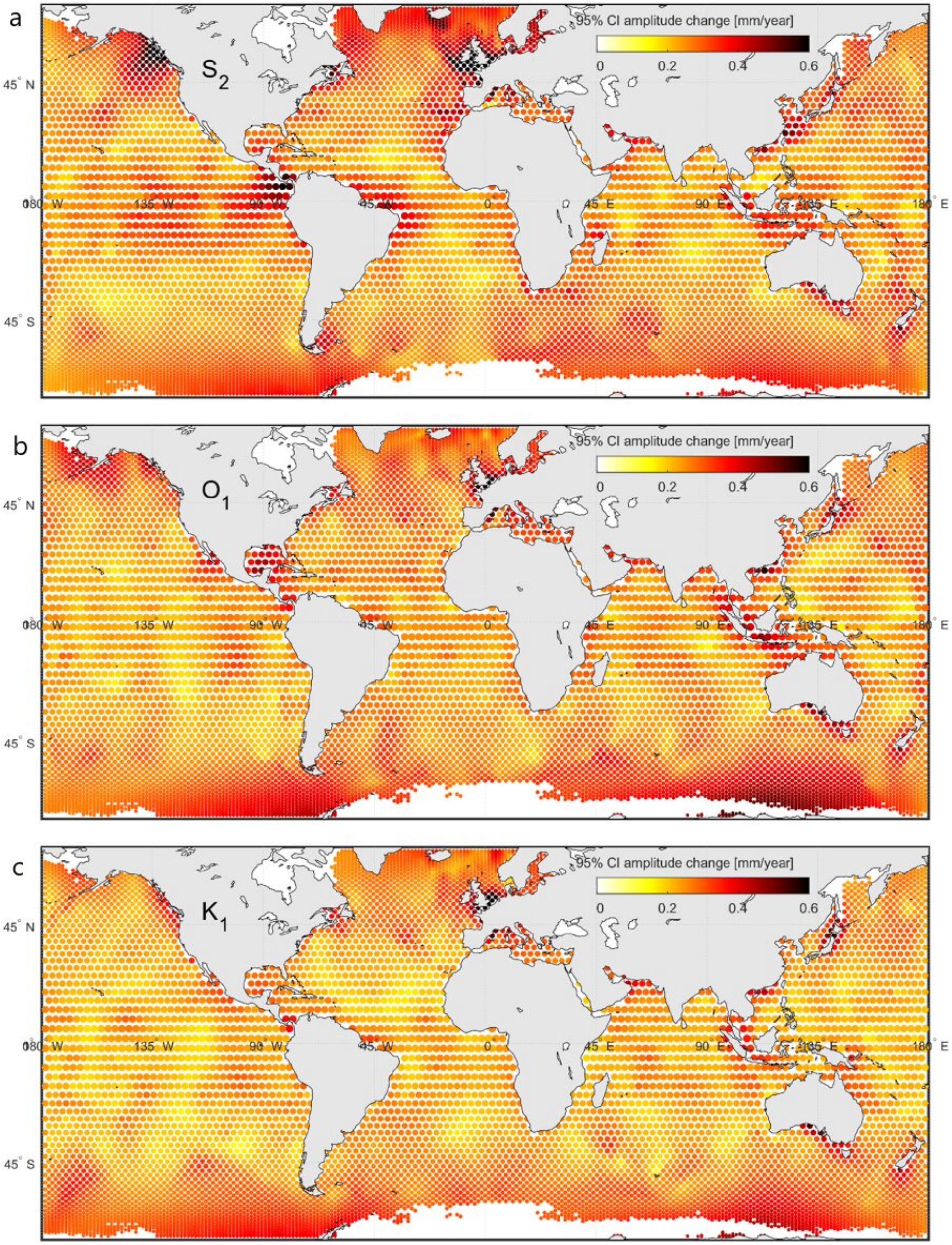


Figure C.2: 95% confidence intervals for trend estimates derived from amplitude standard errors derived from model approach described in paper following the TintHA approach, for S_2 (a), O_1 (b) and K_1 (c). The intervals for S_2 include the sensitivity to the ionospheric correction as explained in Appendix A.

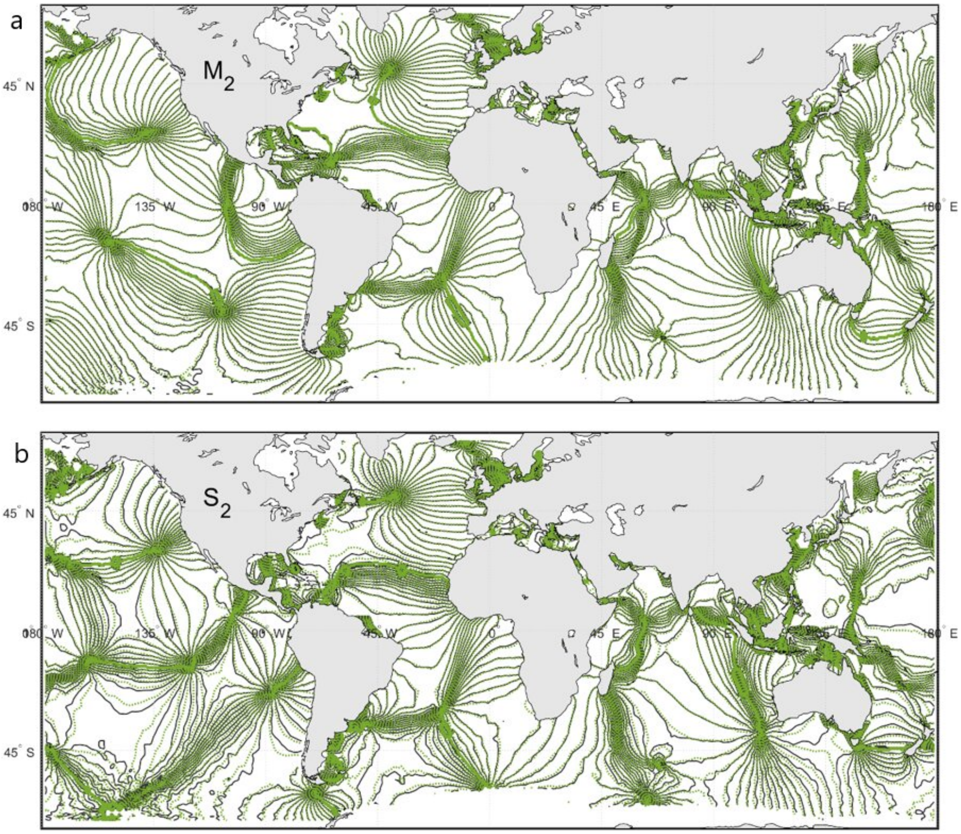


Figure C.3: Cotidal phase lines at 10-degree intervals, for 1993 in black and for 2020 in green: M_2 (a), S_2 (b), derived by TintHA approach.

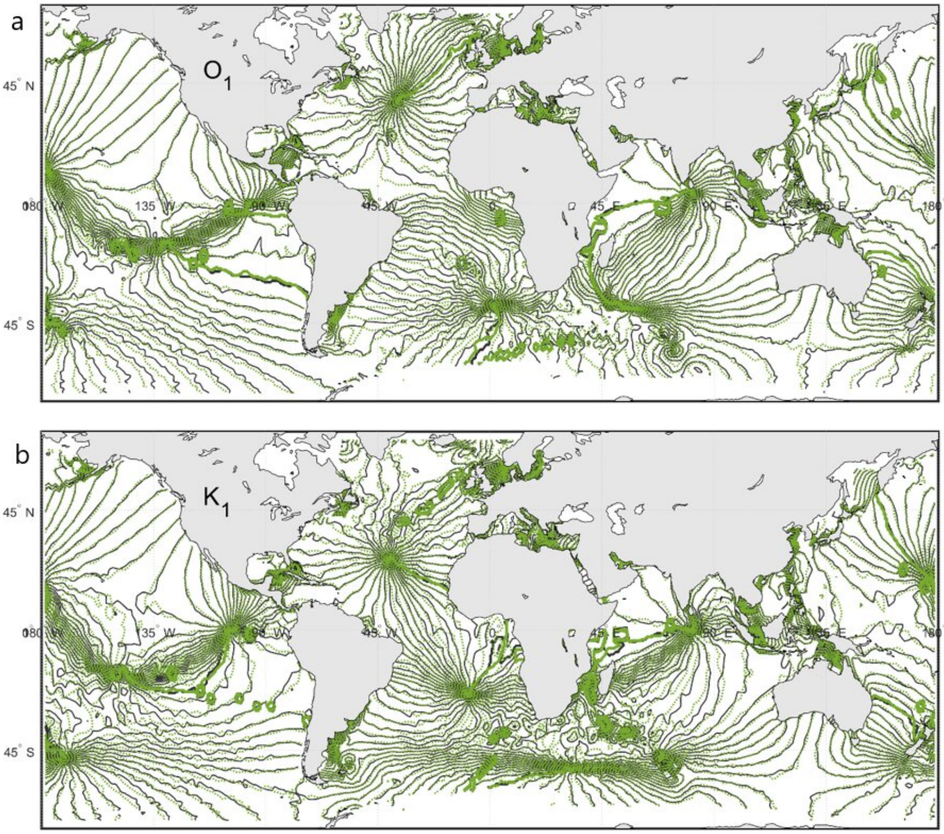


Figure C.4: Cotidal phase lines at 10-degree intervals, for 1993 in black and for 2020 in green: O_1 (a), K_1 (b), derived by TintHA approach.

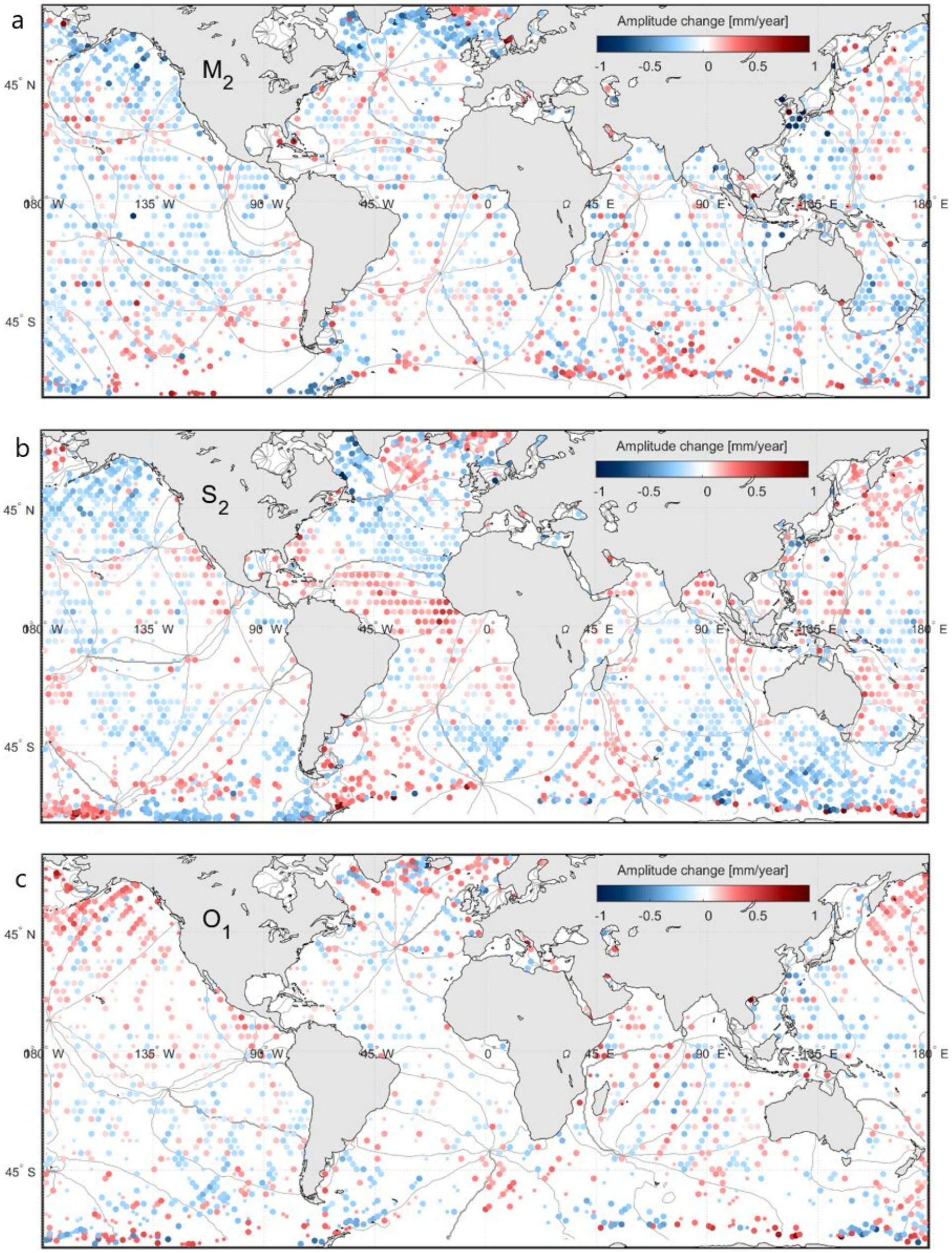


Figure C.5: Linear change in M_2 (a), S_2 (b), and O_1 amplitude per year (1993-2020) following the SegHA approach. The smaller scatters indicate data that exceeds both the UTide and GTSM 90% confidence intervals, while the larger scatters indicate significant data at the 95% confidence level. Lines in the background depict tidal phases at 45° intervals.

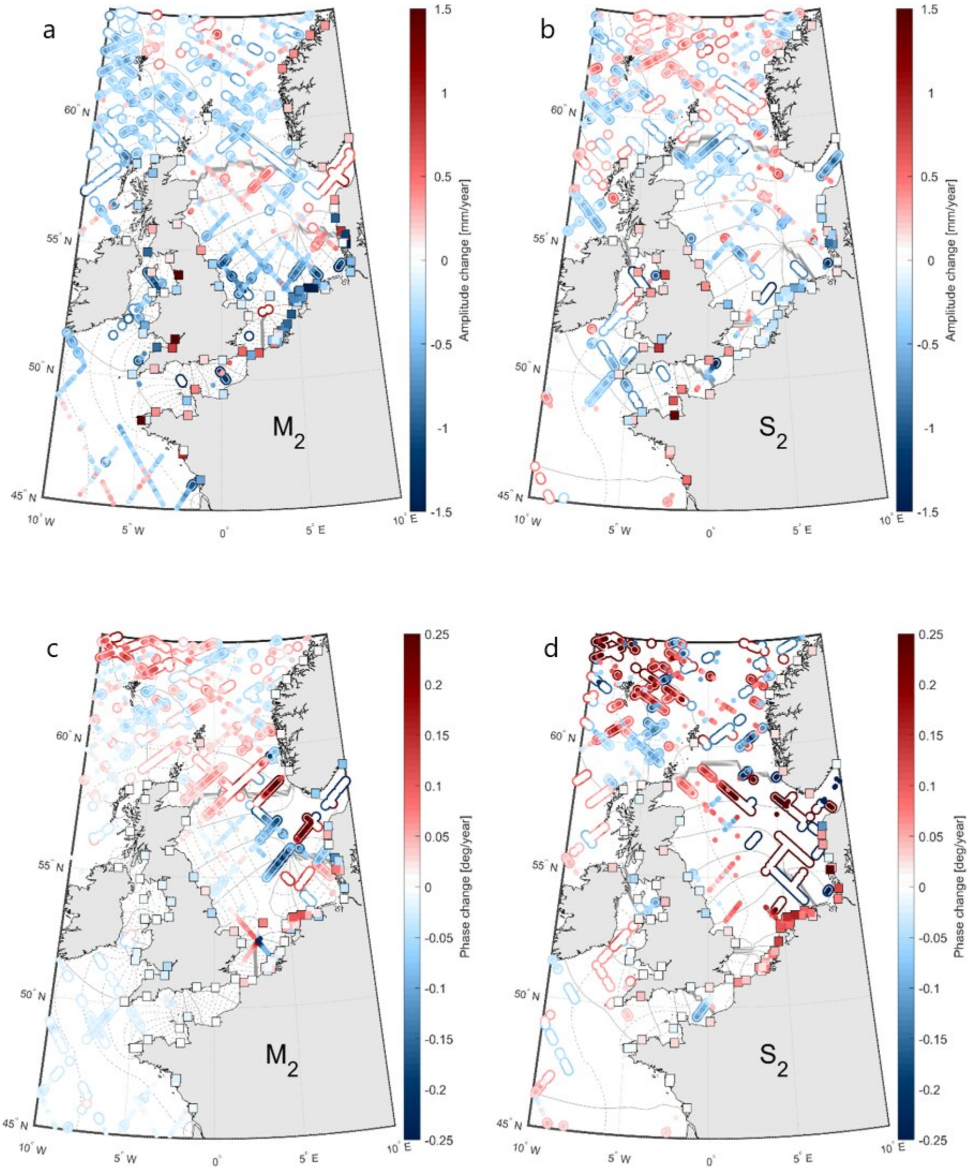


Figure C.6: Linear change in M$_2$ amplitude (a), S$_2$ amplitude (b), M$_2$ phase (c) and S$_2$ phase (d) per year derived by the SegHA approach. The smaller solid scatters indicate significant trends given the UTide 95% confidence intervals, the hollow outline indicates significance according to the GTSM 95% confidence intervals as described in the paper. Co-tidal maps are shown in the background where the solid line indicates the phase at 45° intervals, the dashed lines show the amplitudes at 0.25 m intervals.

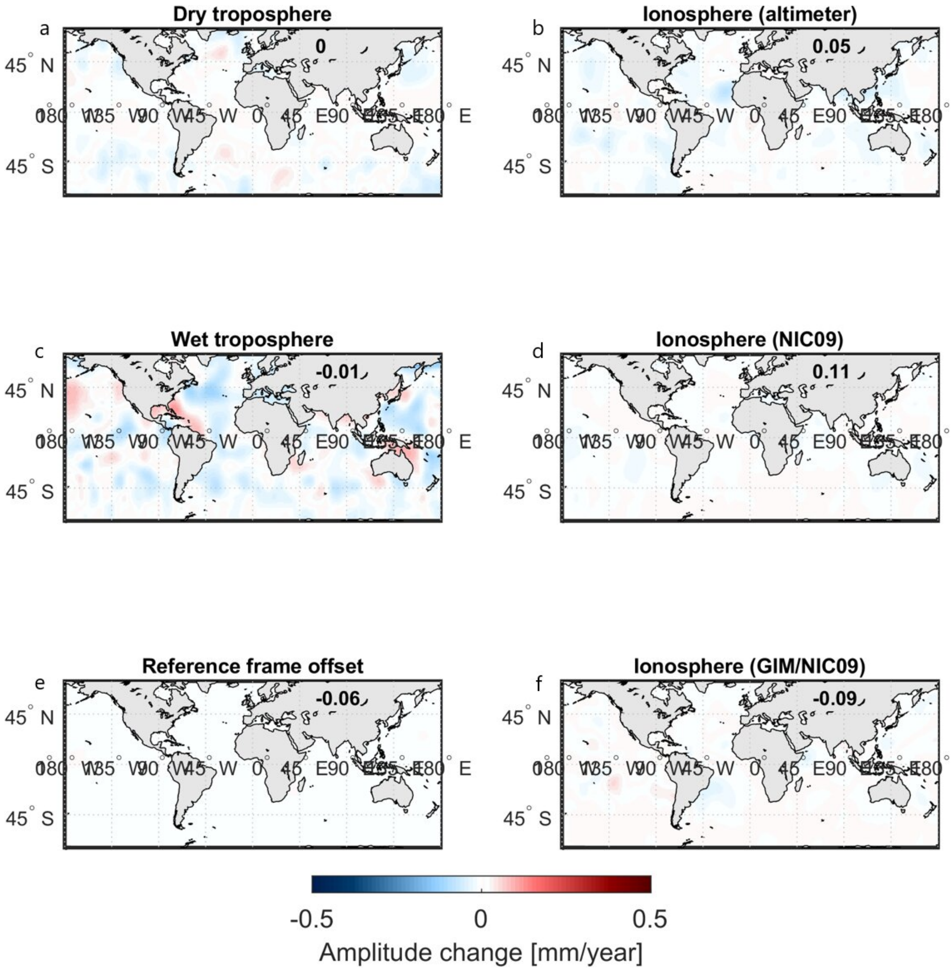


Figure C.7: Linear change in M_2 amplitude in the atmospheric propagation corrections (a, b, c, d, f) and reference frame offset correction (e) that were applied to the TPJ data. Data has been interpolated for visualization purposes. The value on top of Russia depicts the correlation coefficient between the change depicted in the figure and the change in M_2 amplitude as presented in the paper (Figure 2.4a). Following from the TintHA approach.

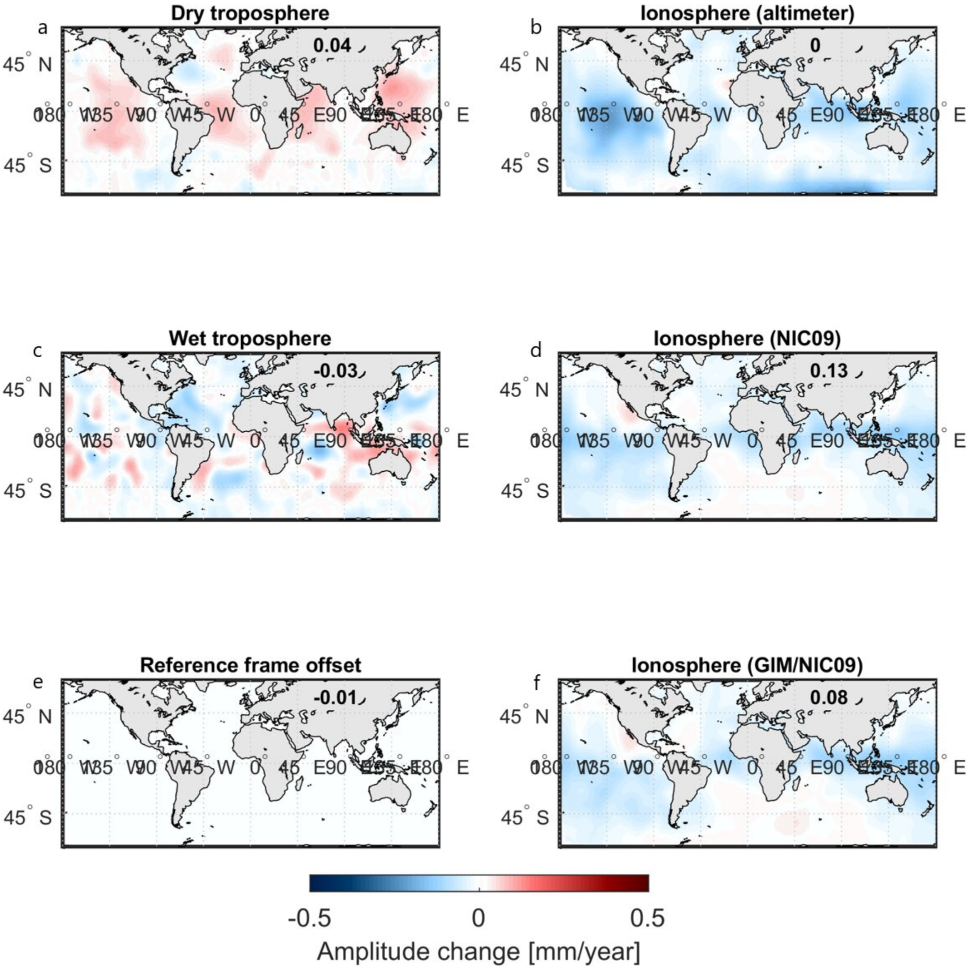


Figure C.8: Same as Figure C.7 but for S_2.

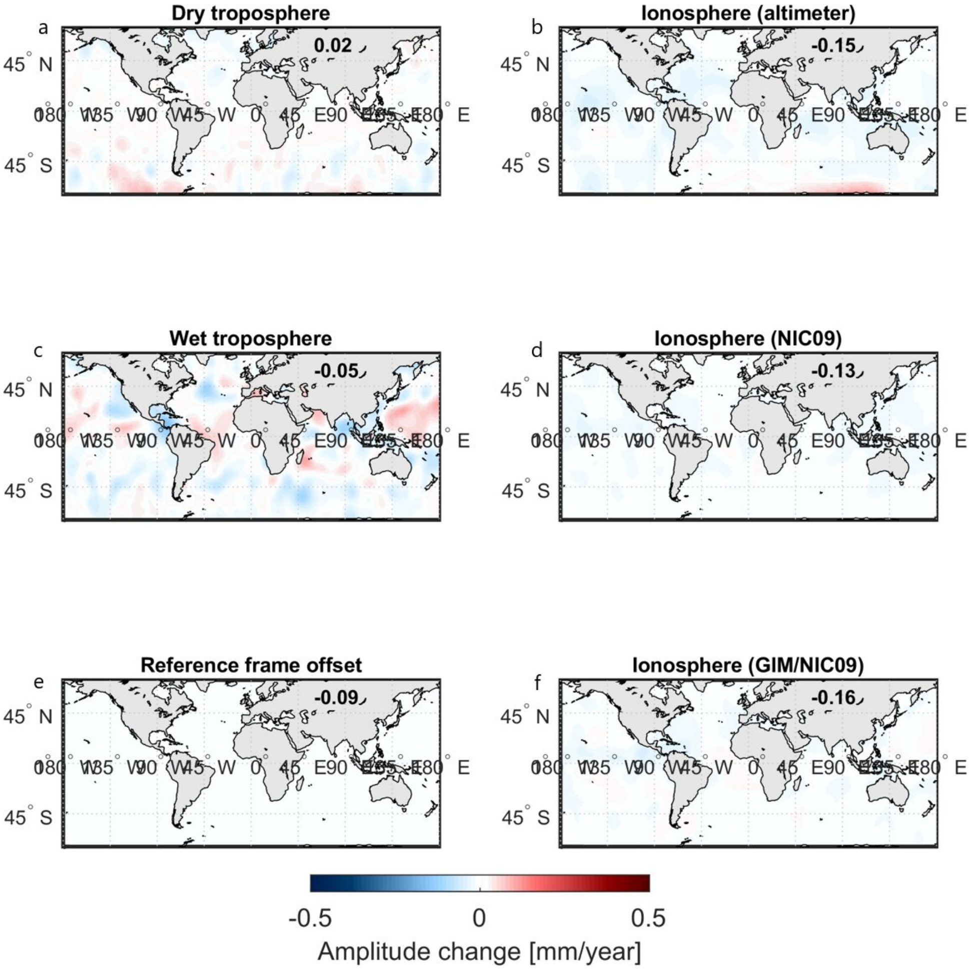


Figure C.9: Same as Figure C.7 but for O_1.

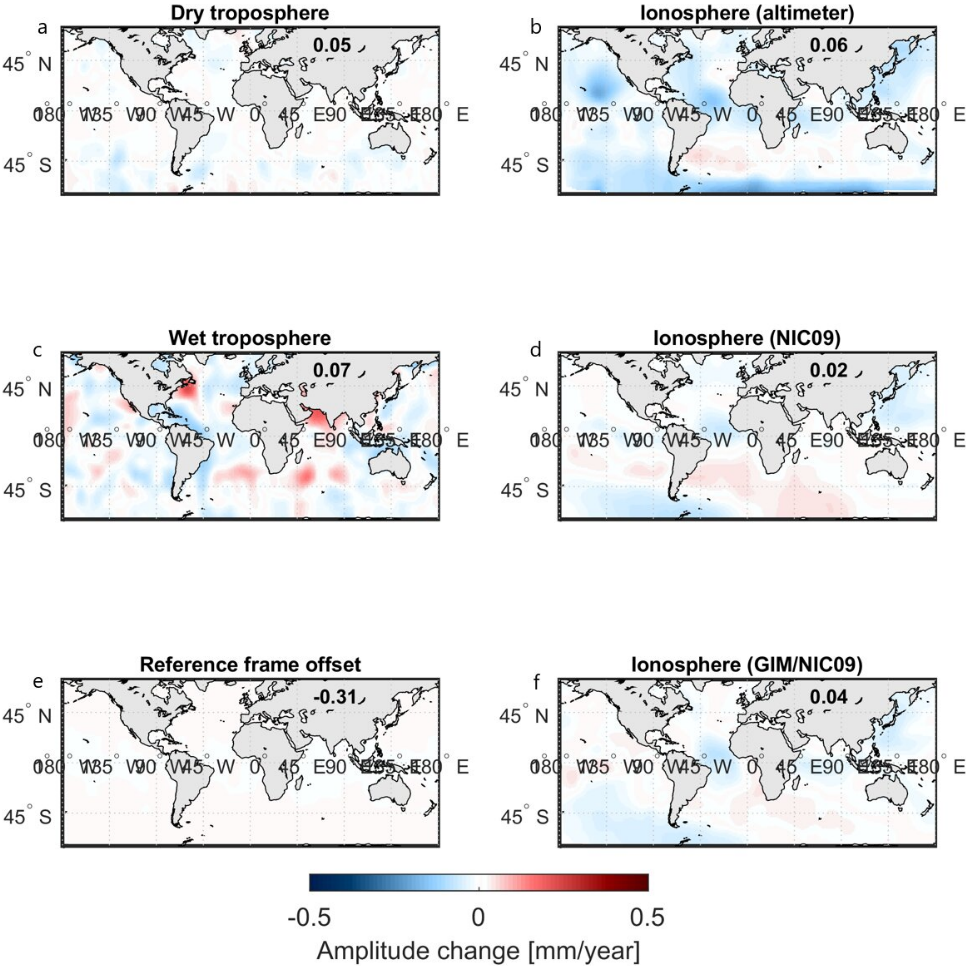


Figure C.10: Same as Figure C.7 but for K_1.

D

TIDE VALIDATION GLOBAL TIDES AND SURGE MODEL

The accuracy of the latest version of the Global Tides and Surge Model (GTSM) has been assessed with respect to FES_2014 by Wang et al. (2021); there referred to as GTSM with fine grid, Approach 2. Where their study included few locations in the Arctic, we performed a similar analysis with specific focus on this region.

For this purpose, GTSM was run for the period from July 1, 2014 to July 31, 2014 with a spin-up time of 1 week, forced by tides only (no ice). For 459 locations that are evenly distributed over the area north of 50° N, sea level timeseries were created with a time-step of 10 minutes. In addition, the same timeseries were computed from FES_2014. FES_2014 was produced by Noveltis, Legos and CLS and distributed by Aviso+, with support from CNES (https://www.aviso.altimetry.fr/). Subsequently, the two signals were compared and the RMSE was calculated for each location respectively (Figure D.1). With the RMSE varying from 0 to 25 cm across the domain, an average RMSE of 7.16 cm was computed. This is larger than the RMSE of 4.48 cm computed by Wang et al. (2021), which is a natural consequence of the larger number of data points, particularly in shallow seas and bays. In fact, our analysis resulted in an average RMSE of 4.01 cm for the deep ocean (> 1000 m deep) and 10.55 cm for the shelves (< 1000 m deep).

Based on a global set of 1973 locations, Wang et al. (2021) computed an average RMSE of 3.67 cm for the same period and version of the model. A related assessment was done by Stammer et al. (2014) for other purely hydrodynamic global models. The performance of these models was analyzed by comparing the simulated M_2 elevations to TPX08. This resulted in RMSEs ranging from 4.41 to 7.76 cm for the deep ocean and from 17.4 to 27.9 cm for shelf seas. Keeping in mind the different methods of both analyses, the performance of GTSM appears comparable to these other hydrodynamic models.

In addition, RMSEs between simulated tidal amplitudes (derived from GTSM for July-September 2014, no ice) are compared to tide gauge constants for 208 tide gauge stations. The tide gauge constants of four major Arctic tides (M_2, S_2, K_1 and O2) were obtained from Kowalik and Proshutinsky (1994). For each tide gauge, the difference be-

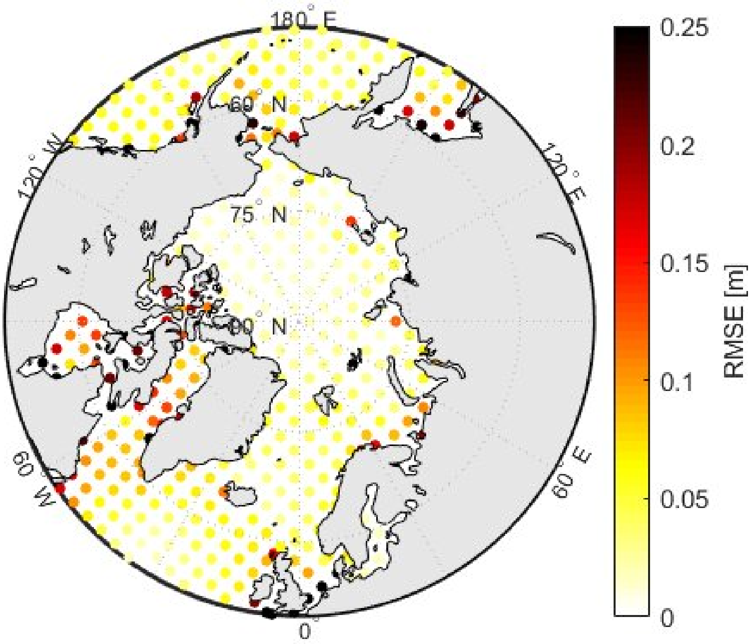


Figure D.1: RMSE [m] between the reconstructed tidal signal from FES2014 and simulated tidal signal by GTSM from July 1 to July 31, 2014.

tween the simulated M_2-amplitude and that from Kowalik and Proshutinsky (1994) is displayed in Figure D.2. As can be seen in this figure, the modelled M_2 amplitudes exceed the tide gauge constants in Baffin Bay and the Canadian Archipelago, while the opposite is the case along the coast of Scandinavia and Russia. On average a RMSE of 20.38 cm was computed for M_2, while the RMSEs for S_2, K_1 and O_1 are 7.47, 6.38 and 3.20 cm respectively. These values are significantly larger (2- to 4-fold) than those presented by Stammer et al. (2014) based on a similar analysis only involving data assimilating models. The latter models assimilated satellite altimeter data and are therefore expected to produce tidal constants that are much closer to reality than purely hydrodynamic models. However, for a sensitivity study like we present in the accompanying paper, the usage of a hydrodynamical model is required. In addition, the difference between tide gauge constants and those derived from GTSM emphasizes the relevance of improving hydrodynamic models by incorporating seasonal effects.

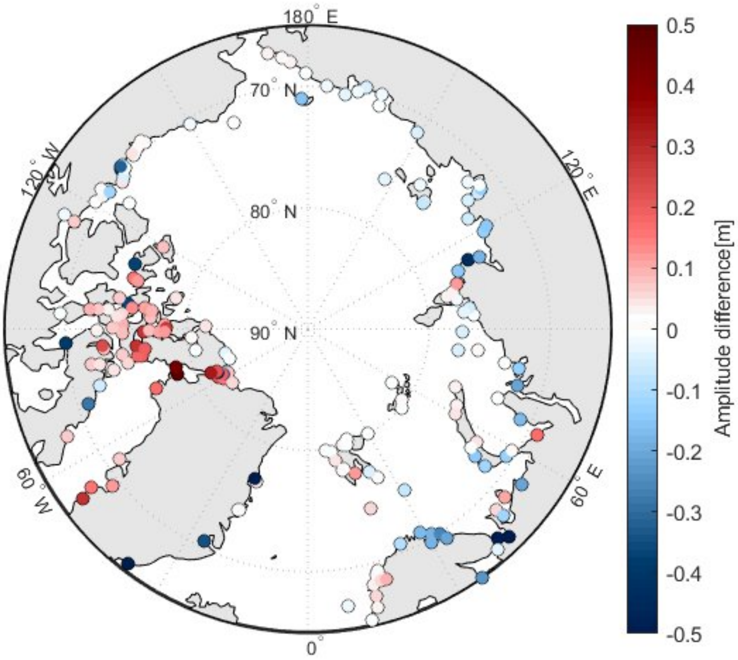


Figure D.2: Differences in M_2-amplitudes derived from simulations (GTSM, from 1 July to 30 September 2014) and tide gauge constants from Kowalik and Proshutinsky (1994). Positive values indicate simulated amplitudes exceed tide gauge-derived amplitudes.

E

FAST ICE INDUCED MONTHLY CHANGE IN M_2 AMPLITUDE

An additional experiment was performed where the effect of monthly change in fast ice cover on the Arctic M_2 amplitude was studied. This experiment was done with the same model settings as described in the main text, using fast ice covers for March 11, April 8, May 6, June 17, July 15 and August 12, 2013 obtained from U.S. National Ice Center (2009) (Figure E.1). In Figure E.2, the resulting monthly M_2 amplitudes with respect to the September amplitude are shown.

This experiment supports the assumptions mentioned in Section 4.2.1. Fast ice induced modification of the M_2 amplitude is largest in case of the maximum fast ice cover (March here). Additionally, in most of the domain the magnitude of the relative amplitude monotonically reduces to zero. This applies to the Hudson Bay area, where the relative amplitude gradually reduces from April onwards until it reaches 0 in July. It also applies to the Canadian Archipelago and Russian coast, where the loss of fast ice during July and August (Figure E.1) directly translates to an amplitude increase (Figure E.2e,f). Only in Ungava Bay the amplitude difference changed sign for one month (positive in June). However, based on altimeter-derived seasonal modulation (Figure 4.2) this positive effect was expected in March already. As could be seen in Figure D.1, GTSM performs less well in the area surrounding Ungava Bay/Hudson Strait and possibly the late emergence of the positive effect in Ungava Bay is an artifact of this.

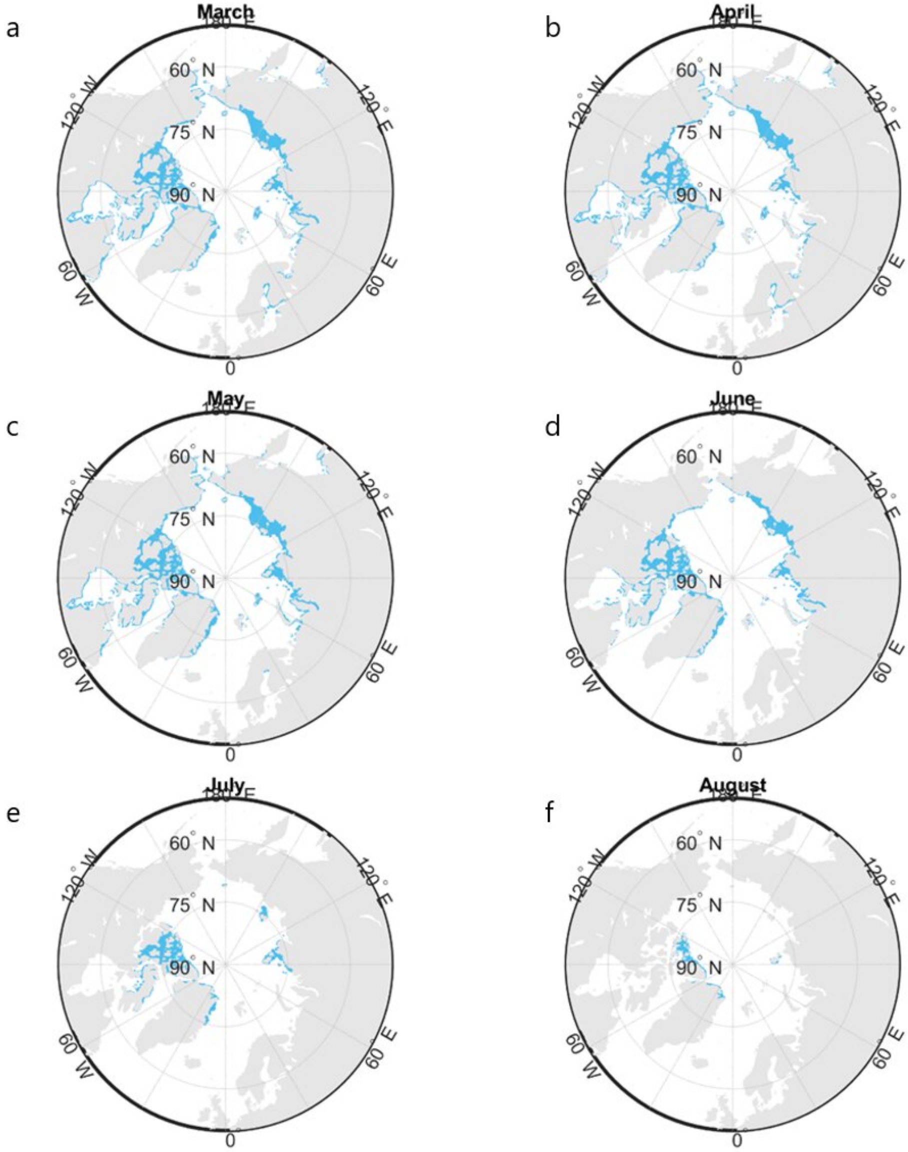


Figure E.1: Fast ice cover extents of March (a), April (b), May (c), June (d), July (e) and August (f), 2013. Data obtained from updated dataset of U.S. National Ice Center (2009).

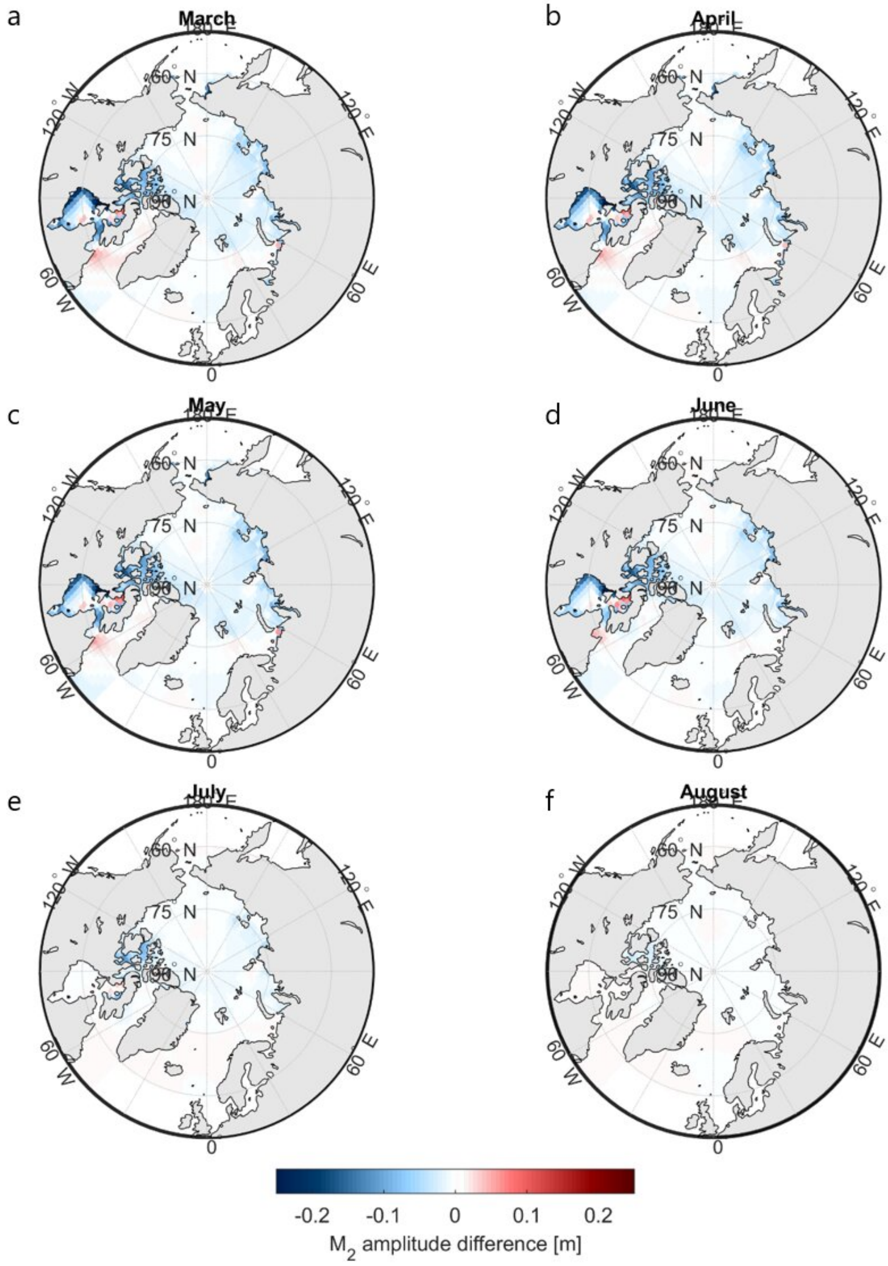


Figure E.2: Model-derived differences between monthly M_2 amplitude of respectively: March (a), April (b), May (c), June (d), July (e), August (f) and September. Where the model is forced by monthly average fast ice covers (Figure E.1) and the amplitude differences are calculated by subtracting the September amplitudes from that of the other months.

F

SUPPORTING FIGURES FOR CHAPTER 4

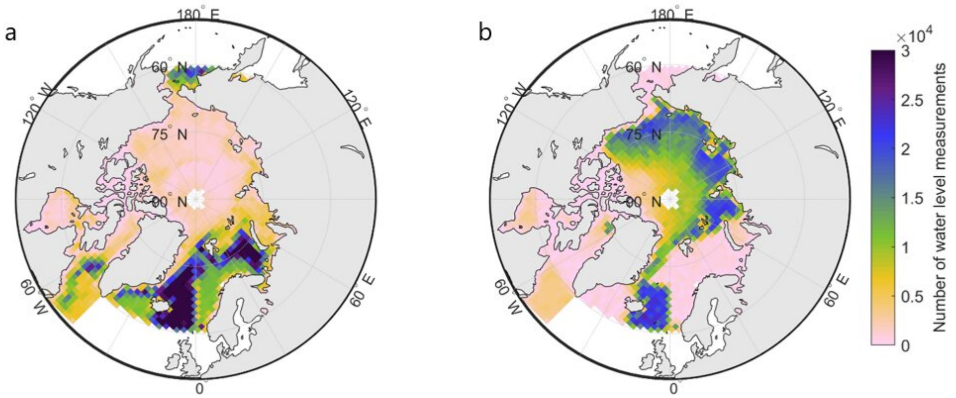


Figure F.1: Number of SAR measurements for March (a) and September (b) from June 2010 till December 2019. Data availability largely depends on the SAR coverage of both Sentinel-3 and CryoSat-2 (the latter varies per month) and the sea ice coverage.

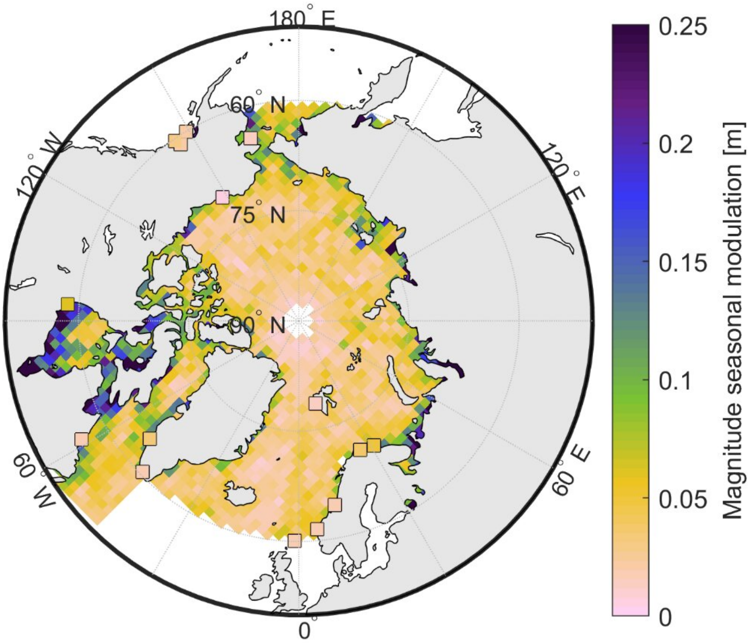


Figure F.2: Magnitude of seasonal modulation in the Arctic derived from altimetry (as in Figure 4.2c) over the period 2011-2019 (background) and the average magnitude of seasonal modulation derived from tide gauges over the same period (squares). Seasonal modulation was calculated following the approach specified in the main text (Section 4.2.1). Tide gauge data were obtained from the University of Hawaii Sea Level Center (Caldwell et al., 2015). The RMSE between the two measures is 0.099 m.

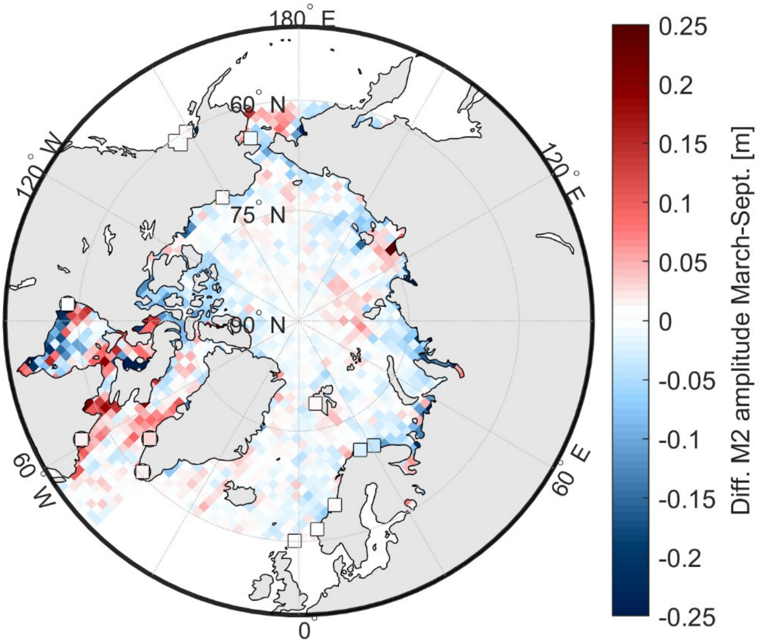


Figure F.3: Difference between March and September M_2 amplitude in the Arctic derived from altimetry (as in Figure 4.5a) over the period 2011-2019 (background) and the average March-September amplitude difference derived from tide gauges over the same period (squares). Seasonal modulation was calculated following the approach specified in the main text (Section 4.2.1). Tide gauge data were obtained from the University of Hawaii Sea Level Center (Caldwell et al., 2015). The RMSE between the two measures is 0.080 m.

G

DEFINITION OF THE THRESHOLD CLASSIFIER

For the Threshold classifier, the threshold values were determined by solving an optimization problem. For this, the distribution of the waveform features was studied on a class-by-class basis. For each feature, a range of thresholds was empirically deter-mined that would separate most leads from sea ice waveforms. This resulted in the following ranges; MAX: 3000-7000, skew: 7-8, PP: 0.15-0.35, PPloc: 0.5-0.7, and ww: 25-50. Note that this only involves the five features with the best predictive capacity as was determined in Appendix H. Subsequently, 200 random combinations of thresholds within these ranges were created and applied to the data (Wernecke and Kaleschke, 2015). The result of this random search is analyzed based on the produced TLR and FLR (Figure G.1). The final set of thresholds was chosen such that the total number of misclassifications was minimized. Waveforms are classified as leads when: MAX $>$ 3000 counts, PPloc $>$ 0.55, ww $<$ 45 bins, PP $>$ 0.24, and skew $>$ 7. For the D-05 experiment, the thresholds for ocean classes were determined as fol-lows: 500 $>$ MAX $>$ 1500 counts, 0.2 $>$ PPloc $>$ 0.35, 85 $>$ ww $>$ 110 bins, PP $<$ 0.1, and 1.5 $>$ skew $>$ 3.5. The remaining data were classified as sea ice.

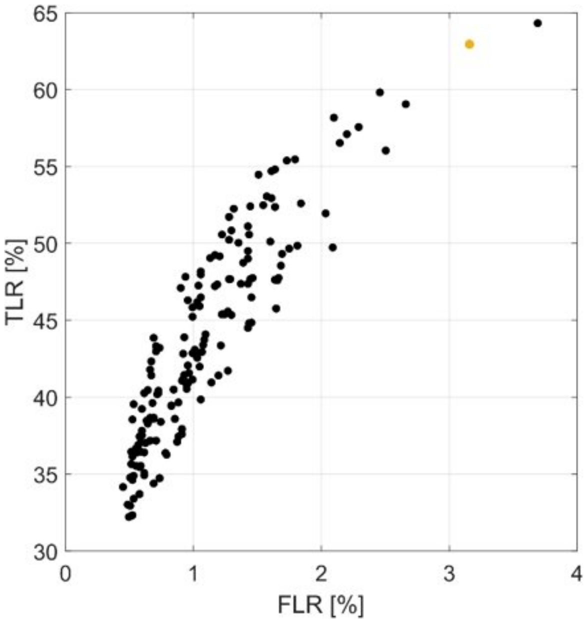


Figure G.1: ROC graph showing the results of the random search of thresholding values for lead classification. The orange point depicts the final choice of the thresholding values used in the paper.

H

TUNING OF CLASSIFIERS

As described in Section 5.3.2, the classification potential of individual waveform features was studied. For each considered classifier, the resulting accuracies are given in Table H.1. The produced accuracies are generally high, with most of the waveform features producing more than 80% accuracy for most of the classifiers. However, the unsupervised learning classifiers (Kmedoids and HC) achieved substantially lower accuracies when using LeW, PPL, PPR, and NrPeaks. The HC classifier also did not perform well when using kurtosis. It was found that using WW, PP, PPloc, skewness, and MAX, produces high accuracy for all classifiers. Though TeW has achieved a high average accuracy, it has not been selected due to the relatively low accuracy produced by the Kmedoids classifier. Additionally, most of the supervised learning classifiers produced very similar results when using NrPeaks (79.8%). However, all classifiers that produced this accuracy predicted all of the wave-forms to be sea ice, i.e. obtained a TLR of 0%. Because NrPeaks can only have discrete integer values, it is not a suitable feature for machine learning algorithms. Previous studies which used NrPeaks as a predictor, were thresholding-based classifications (e.g., Bij de Vaate et al., 2021; Schulz and Naeije, 2018).

The sensitivity to the addition of more predictors to the initial set of five (MAX, skew, PP, ww, PPloc) was assessed by a comparison of the produced ROC graphs (Figure **??**). Six possible combinations of the features were tested (Table H.2). Here the NrPeaks was excluded completely.

From Figure H.1, it appears that for most classifiers, the addition of predictors does not have a significant effect on classifier performance. The largest effects are observed for DT and SVM (Figure H.1a/i), but this concerns reduced performance when more predictors are included. Adaboost and Bagged (Figure H.1b/c) may benefit to some extent from the larger set of predictors, although this is mostly restricted to the far low FLR range (< 4%). In addition, both unsupervised classifiers appear to benefit from the addition of kurt, TeW and or PPR, although the effect is small.

Table H.1: Training accuracies [%] of classifiers trained with a single predictor. The shaded rows indicate the parameters that were selected for this paper.

	AdaBoost	Bagged	KNN	SVM	DT	NB	LD	ANN	RUSBoost	Kmedoids	HC	Average
MAX	88.76	88.76	88.84	88.96	88.18	88.92	88.93	88.92	88.76	83.41	82.86	87.85
Kurt	87.84	87.78	87.83	88.06	87.06	88.07	88.05	88.07	87.85	86.57	49.50	84.55
Skew	89.04	89.05	88.95	89.08	88.40	89.05	89.06	89.06	89.00	87.59	88.39	88.81
PP	90.79	90.82	90.90	91.01	90.25	90.97	90.98	90.98	90.82	90.68	90.67	90.82
WW	91.02	91.00	91.05	91.10	90.99	91.10	91.10	91.10	91.03	90.93	90.84	91.03
LeW	85.82	85.82	85.82	85.82	85.82	79.88	85.82	85.82	85.82	20.91	28.46	75.14
TeW	88.38	88.38	88.38	88.38	88.38	88.38	88.38	88.38	88.38	71.28	86.45	86.79
sigma0	87.27	99.43	87.15	87.19	87.82	86.02	86.06	87.21	84.15	69.23	64.41	84.18
PPL	86.35	86.45	86.50	86.52	85.65	86.52	86.55	86.47	86.38	43.60	20.18	77.30
PPR	87.53	87.56	87.72	87.79	86.96	87.76	87.80	87.82	87.61	70.48	20.17	80.58
PPloc	90.32	90.48	90.41	90.49	89.89	90.49	90.49	90.46	90.50	89.79	90.26	90.34
NrPeaks	79.80	79.80	79.80	79.80	79.80	79.80	79.80	79.80	44.30	59.10	70.22	73.80

Table H.2: Combinations of waveform features which were used in this analysis.

Combination	Waveform Features
Combination 1	MAX, skew, PP, ww, and PPloc
Combination 2	MAX, skew, PP, ww, PPloc, and kurt
Combination 3	MAX, skew, PP, ww, PPloc, and TeW
Combination 4	MAX, skew, PP, ww, PPloc, and PPR
Combination 5	MAX, skew, PP, ww, PPloc, kurt, TeW, and PPR
Combination 6	MAX, skew, PP, ww, PPloc, kurt, TeW, PPR, LeW, sigma0, and PPL

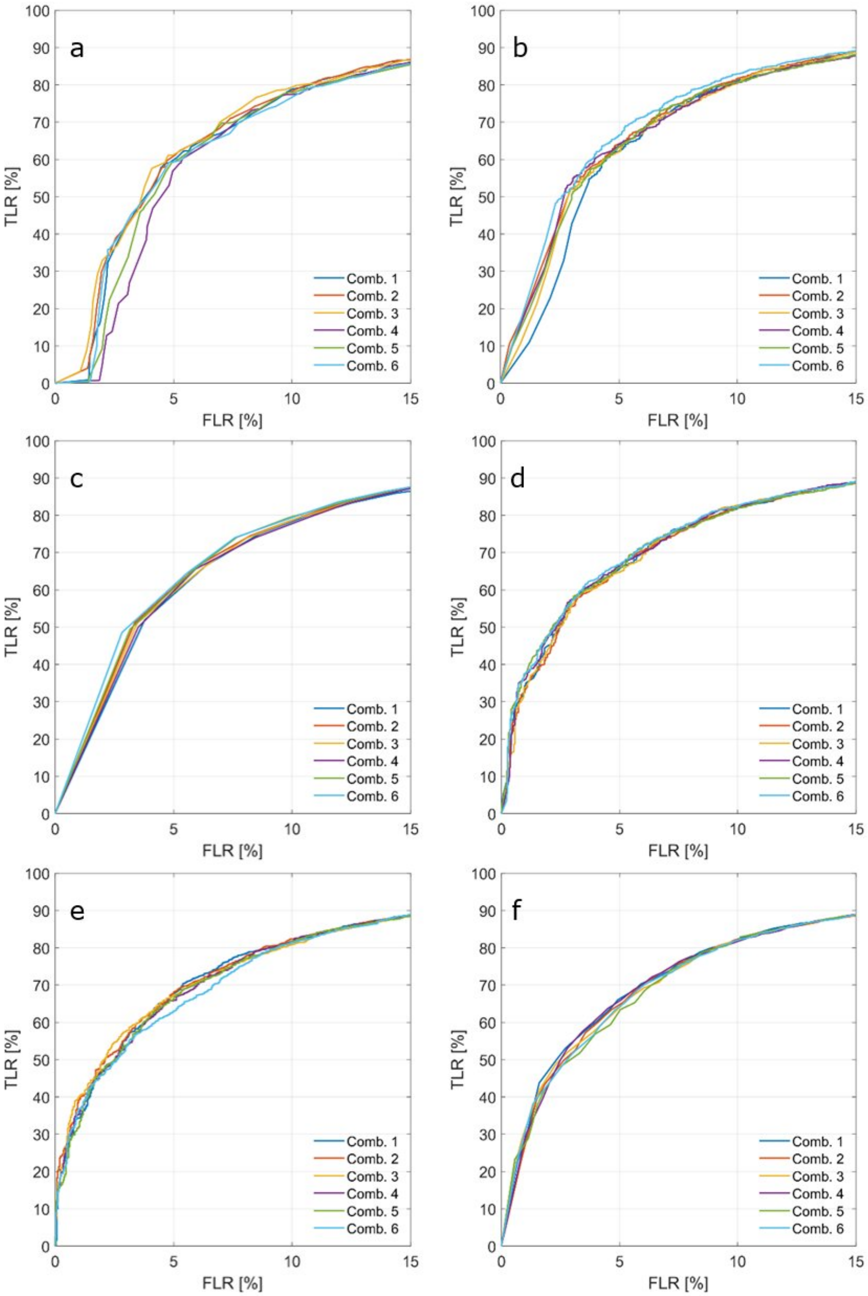


Figure H.1: ROC graphs of supervised learning algorithms for different waveform feature combinations (see Table A2.2): DT (a), AdaBoost (b), Bagged (c), RUSBoost (d), ANN (e), KNN (f), LD (g), NB (h) and SVM (i), Kmedoids (j) and HC (k).

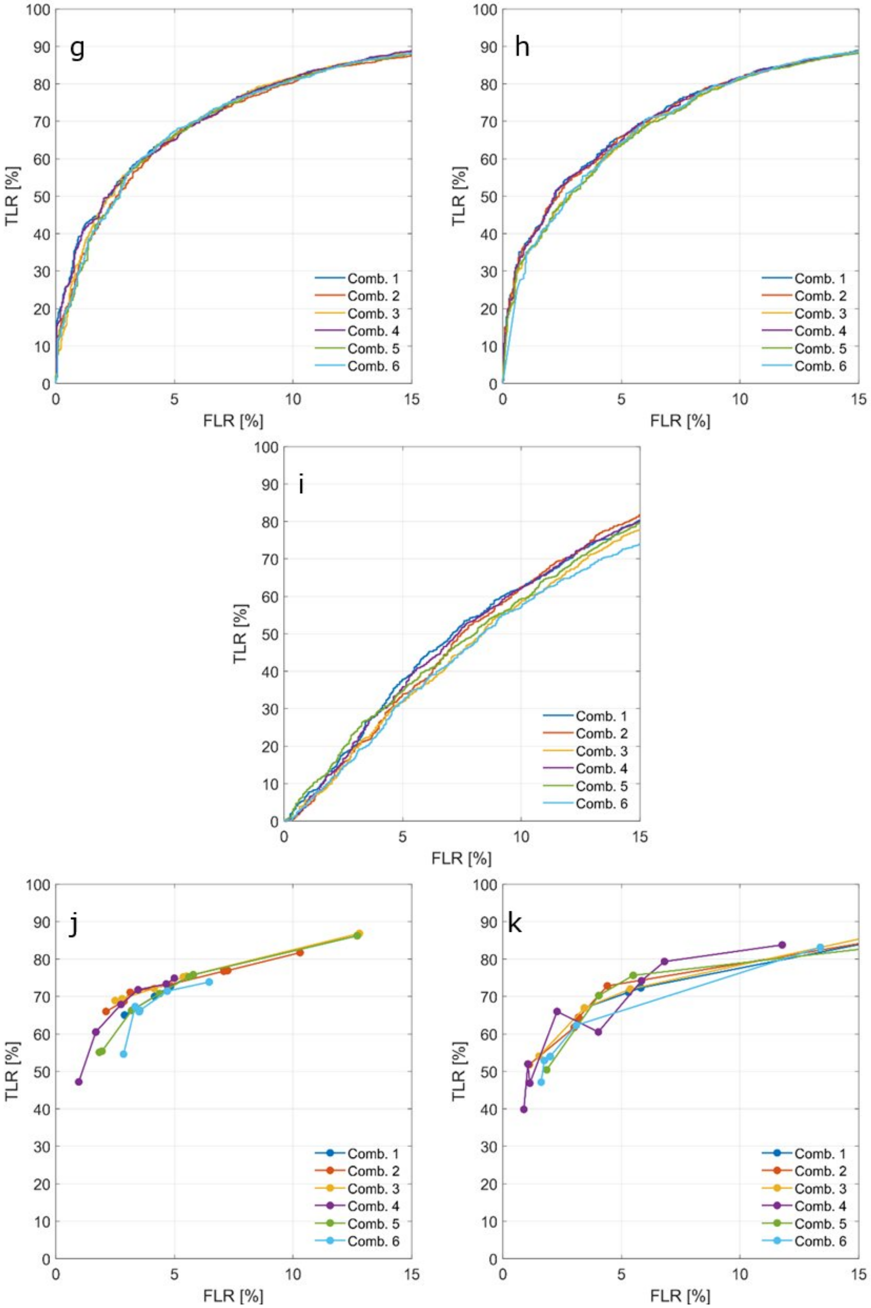


Figure H.1: Continued

I

IMPROVING THE ADDITIONAL OPEN OCEAN CLASSIFICATION

Since the main goal of this paper is to improve the detection of leads, classifier settings and selection of waveform features have been optimized with the focus on the distinction between leads and sea ice. However, when one is interested in processing satellite tracks that may include SAR returns from the open ocean, the classification may be further optimized. As can be seen in Figure I.1a, the misclassification of sea ice returns as open ocean (isolated black dots between $70-77°$ N) can result in large SSH errors. To prevent these misclassifications, one could opt for the addition of a certain along-track history parameter that combines values from neighboring samples (referred to as test case D-05h).

Here, this history parameter was defined as the moving standard deviation of the PP over 25 neighboring samples (see Figure I.1b). In the Threshold classifier, the following condition was added for the ocean class: $\text{movstd}(\text{PP}, 25) < 0.01$.

The addition of the history feature significantly improved the classification of both ocean and lead waveforms (Figure I.2). All performance measures improved, but the effect on the TwR and FwR was the strongest, in particular for the Threshold and Kmedoids classifier. Note that in this study, the ocean data was added as a single separate track. The addition of a history feature would likely not improve the classification of consecutive ocean and sea ice data points. However, most ocean and sea ice data are well separated.

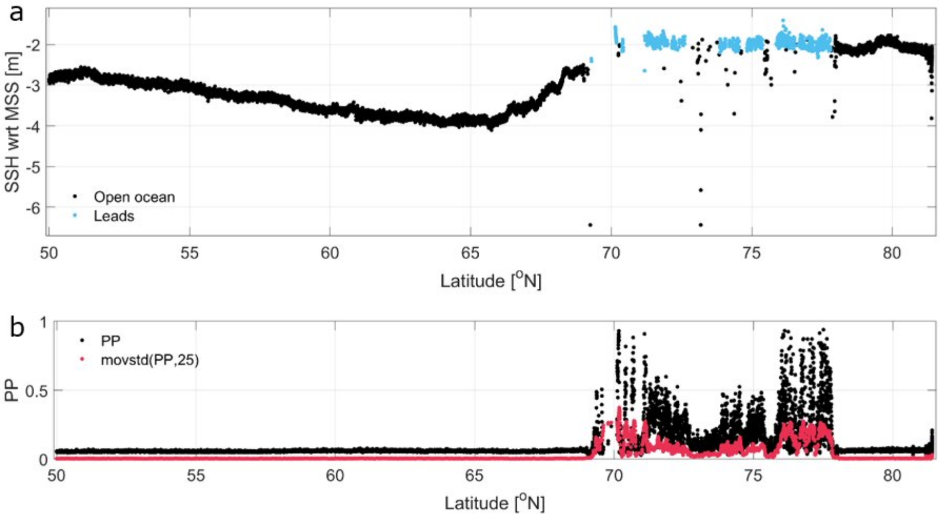
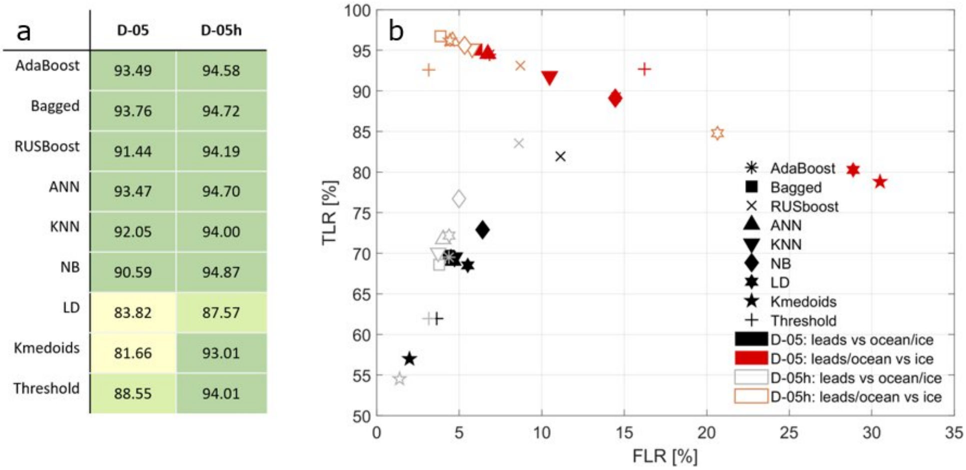


Figure I.1: An example of an along-track SSH series referenced to DTU18-MSS (Sentinel-3B track from November 5, 2021) obtained by the AdaBoost classifier and retracked by fitting the SAMOSA model (Dinardo et al., 2018; Ray et al., 2015) (a). Along-track PP and moving standard deviation of the along-track PP (b).



a	D-05	D-05h
AdaBoost	93.49	94.58
Bagged	93.76	94.72
RUSBoost	91.44	94.19
ANN	93.47	94.70
KNN	92.05	94.00
NB	90.59	94.87
LD	83.82	87.57
Kmedoids	81.66	93.01
Threshold	88.55	94.01

Figure I.2: Overall classification accuracies [%] for D-05 and D-05h (a), comple-mented by ROC graphs showing classification performances (b). The D-05(h) results are plotted in two ways: the actual TLR/FLR where lead classes are compared to ocean and ice combined (black/grey) and the TwR/FwR (red/orange) that combines the water classes (ocean and leads). Note that the red/orange markers for AdaBoost and ANN overlap.

J

PREDICTOR CORRELATION

A common issue to occur in multivariate analyses is multicollinearity. This phenomenon – where one or more predictors are linearly related - can lead to biased estimation and may cause the trained model to be unstable (Yoo et al., 2014. Ideally, the predictor variables would be chosen in such a way that the correlation between predictors (waveform features) is minimized, yet the correlation with the response variable (class in this case) is large. To test the impact of multicollinearity in the study case described in this paper, Pearson correlation coefficients are calculated for each predictor pair and Kendall correlation coefficients for each predictor-response pair. This is done for all data divisions as described in the paper (Figure J.1).

Four of five predictors used in the study are highly correlated: skew, PP, ww, and PPloc (Figure J.1a). However, these features are all directly dependent on the shape of the waveform and therefore the correlation is in this case deemed inevitable. Nevertheless, if the correlation among predictor variables is consistent across the data divisions that the trained model is applied to, multicollinearity does not necessarily degrade classifier performance (Kutner et al., 2005. This mainly concerns the D-02 and D-03 test cases (Table 5.2). From Figure J.1b/c and d/e, it appears that the correla-tion between the considered predictors remains consistent, regardless of differences in the study area or sensing period. This does not apply to all predictors (e.g., sigma0). The distinct difference between Figure J.1f and other subfigures suggests that a classifier trained with winter data should not be applied to classify summer data. In addition, the low correlation between all predictors and the response variable (Figure J.1f) emphasizes the poor potential in summer surface classifications based solely on altimetry data.

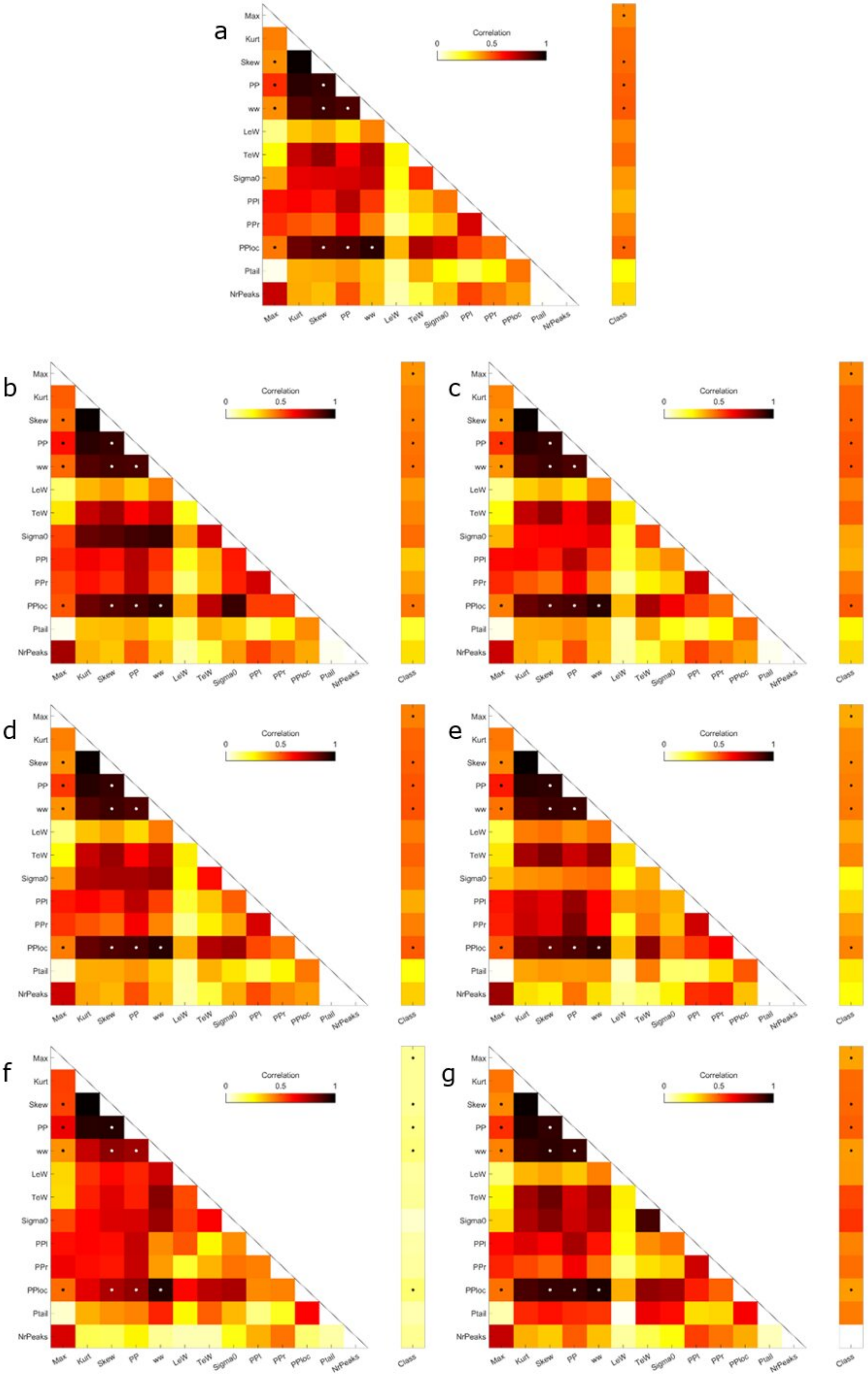


Figure J.1: Absolute correlation coefficients of all predictor pairs and the correlation coefficient between each predictor and the response variable (Class): for the D-01 dataset (a), the D-02 training (b) and testing dataset (c), the D-03 training (d) and testing dataset (e), the D-04 summer data (f) and the D-05 set including ocean data (g). The dots indicate the features that were used in this study.

ACKNOWLEDGEMENTS

Now it is time for me to reflect on this four-year journey across the wonders of the ocean. I will use this page to thank the people that were involved in the creation of this work or helped me to stay on the right track.

First, I want to thank my daily supervisor and copromotor Cornelis Slobbe, who has been particularly involved in my work, for his very constructive and meticulous input. He has challenged me to pay great attention to details and always think critically. Furthermore, I appreciate my second daily supervisor and promotor Martin Verlaan for his vast knowledge and experience. His fresh view has often given rise to new ideas. I also want to express my gratitude to my promotor Roland Klees, for his faith and mental support throughout the PhD.

I would like to thank Ivan Haigh, Michael Schindelegger, July Pietrzak, CK Shum, and Bert Vermeersen for being my committee members and taking the time to review my thesis and provide valuable suggestions.

Waves come and waves go, and so did my fellow PhDs. Despite of that, I valued the consistent feeling of cohesion amongst the PhDs, staff and students at the Geoscience and Remote Sensing department. Many people have contributed to my PhD life being a pleasant experience, but I specifically want to thank Yosra, Amey, Henrique and Frithjof for the many discussions related to our research as well as the get-togethers outside of work.

In that light, I also treasure the time spent with my friends and family. The opportunities to shift my thoughts and change my perspective have been very constructive in the making of this work. Hence, a big thankyou to the people that accompanied me in relaxing and enriching activities; from climbing, hiking and, woodworking to building a ridiculously unstable sailing raft. In particular, I want to thank Dagmar, Igor, Anke, and Jan Geralt for being the foremost source of inspiration and fuel to my creativity. And above all, many thanks to Menno, the best ever coworker I could have wished for during the Covid lockdowns. I am grateful to him for always being willing to join me on whatever train of thought I had boarded and ensuring that I got off at the right stop.

CURRICULUM VITÆ

Inger BIJ DE VAATE

16-09-1995	Born in Hengelo, The Netherlands.
2006–2013	**Gymnasium** Stad & Esch, Meppel
2010–2011	**Study Abroad** Applecross Senior High, Perth (Australia)
2013–2016	**Bachelor of Science in Future Planet Studies** University of Amsterdam Specialization: Earth Sciences Thesis subject: *The relation between glacial habitat factors and the present day distribution and migration of brent geese.*
2016–2018	**Master of Science in Marine Sciences** Utrecht University Thesis subject: *Effect of different vegetation species traits on channel development in tidal marshes.* Internship at Insitute for Marine Research, Bergen (Norway)
2018-2022	**Doctoral Candidate** Geosciences and Remote Sensing Civil Engineering and Geosciences Delft University of Technology

LIST OF PUBLICATIONS

Bij de Vaate, I., Slobbe, D. C., & Verlaan, M. (2023). Mapping the spatiotemporal variability in global storm surge water levels using satellite radar altimetry. *Under review at Ocean Dynamics*

Bij de Vaate, I., Slobbe, D. C., & Verlaan, M. (2022). Secular trends in global tides derived from satellite radar altimetry. *Journal of Geophysical Research: Oceans*, *127*(10), e2022JC018845. https://doi.org/https://doi.org/10.1029/2022JC018845

Bij de Vaate, I., Martin, E., Slobbe, D. C, Naeije, M., & Verlaan, M. (2022). Lead Detection in the Arctic Ocean from Sentinel-3 Satellite Data: A Comprehensive Assessment of Thresholding and Machine Learning Classification Methods. *Marine Geodesy*, *45*(5), 462–495. https://doi.org/https://doi.org/10.1080/01490419.2022.2089412

Bij de Vaate, I., Vasulkar, A. N., Slobbe, D. C., & Verlaan, M. (2021). The Influence of Arctic Landfast Ice on Seasonal Modulation of the M2 Tide. *Journal of Geophysical Research: Oceans*, *126*(5), e2020JC016630. https://doi.org/https://doi.org/10.1029/2020JC016630

LIST OF PRESENTATIONS

Oral presentations

Nederlands Centrum voor Geodesie en Geo-informatica (NCG) Symposium, April 2022, Wageningen, The Netherlands.

American Geophysical Union (AGU) Fall Meeting, December 2021, New Orleans, United States of America.

European Geosciences Union (EGU) General Assembly, April 2021, Online.

TU Delft Sea Level Talks, December 2020, Online.

European Geosciences Union (EGU) General Assembly, May 2020, Online.

51st International Liege Colloquium on Ocean Dynamics, May 2019, Liege, Belgium.

Poster presentations

Living Planet Symposium, May 2022, Bonn, Germany.

Annual Polar Symposium, May 2022, The Hague, The Netherlands.

Annual Polar Symposium, November 2019, The Hague, The Netherlands.

GSE/GRS Research Day, April 2019, Delft, The Netherlands.