Nr.	Review comment	Author's answer	Change in manuscript
1	The paper is mostly descriptive and practically no physical analysis of the observations is performed. It looks mostly like a report on some routine observations, like "a pile of data"…	As indicated by the title, we acknowledge that the aim of the paper is to describe the phenomenon of turbulent wakes, and thus being "*mostly descriptive* ". However, we humbly disagree that the paper is "*mostly like a report on some routine observations"* and that *"practically no physical analysis of the observations is performed"* . Even though the aim of the paper is to describe the extent of wake influence, we have made analysis of the turbulence in the wakes (dissipation rate of turbulent kinetic energy). Furthermore, we have presented an example of what the described spatiotemporal extent of the wake would implicate in terms of temporal and spatial wake-impact in a highly frequented ship lane. Nevertheless, we fully agree with and acknowledge the potential of further analysis of this phenomenon. However, resolving all the parameters determining the characteristics and impact of ship-induced turbulent wakes, will require years of further studies. Therefore, there is a need for a first, more descriptive study, which will provide an understanding of the relevant scales and parameters to consider in future studies, as well as a well described methodology. The work of measuring large quantities of turbulent wakes in ship lanes, is far from routine, and not many observations have so far been published. In addition, we would like to highlight the interdisciplinary nature of this work. The process of identifying and linking the wakes with the ship inducing the wake is time-consuming, but necessary to relate the wakes to the physical properties of the vessels. However, we agree and acknowledge that the vessel-related analysis can be developed further, and we have suggested how to do so in detail in the answer to that specific comment below.	Changed title, revised paragraph to clarify the aim and further motivate the need for the study.
2	...only methodological aspects of the work are discribed, although not clearly enough (see, some specific comments below)	The specific comments are addressed individually below. Furthermore, we agree and acknowledge that there is a focus on describing the methodology of the work. That is because there are currently no published descriptions of best practice or standard methodologies to study ship wakes and passages in heavily trafficked ship lanes during an extended period of time without interfering with the traffic itself. It is an activity with many technical and practical challenges, and therefore, the methodology has been explained in detail. Moreover, most previous studies of the turbulent ship wake, have been performed using echo sounders/multibeam mounted on the ship hull of a small ship, which have travelled across the ship wake in serpentine movements behind the ship, measuring the wake from above. Here we propose a method based on upward facing instruments placed on the sea floor under the ship lane. We therefore consider it important to describe and discuss the methodology in detail, and we propose to add additional figures to illustrate the experimental setup, to clarify the questions asked.	For changes in manuscript see answer to the specific comments below.

3	"One cannot find in the text any new physical effects" and "The paper in its present form does not look interesting and informative from a scientific point of view"	Regarding the general comment "*One cannot find in the text any new physical effects* " and "*The paper in its present form does not look interesting and informative from a scientific point of view* ", we respectfully disagree. To our knowledge, there is no published dataset showing this consistent pattern of turbulent and temporal wakes, and no studies including more than a handful of different ships. The observed maximum wake depth presented in this study, exceed previously reported wake depths, which sheds new light on a novel perspective of the environmental impact from shipping. Moreover, the longevity and persistence of the thermal wakes shown in the satellite data highlights the importance of considering the impact of ship wakes in highly trafficked areas (with up 55 000 passages per year). This effect is important to raise awareness of, especially within the FerryBox community, as the temperature measurements made within ship lanes can be biased if they are made in the wake of another ship. The longevity of the temperature difference indicates that the ship wake water stays a separate entity (is not diluted/mixed) for a substantial period of time (+ 1 hour). Even though only temperature has been measured in this study, the thermal signal can be considered as proxy for the ship wake water, and potential changes in other chemical/physical parameters such as salinity, nutrient concentration etc. should also be sustained in the ship wake water as long as it is not mixed with the surrounding water mass.	This comment is addressed by the revision of the manuscript which clarifies the aim and motivation of the paper. See comment 9 and 1.
4	The aim/motivation of the paper is not clear	Firstly, we suggest changing the title to: "*In situ* observations of turbulent ship wakes and their spatiotemporal extent", to clarify that the aim is to describe the characteristics of the turbulent ship wake. We will also rephrase the last paragraph (lines 105–115) in the introduction as below, to further motivate and describe the aim of the paper:	The title has been changed to clarify that the aim of the paper is to describe the spatiotemporal extent of turbulent ship wakes.
5	An error in formula (1)	We will correct the error.	The error has been amer

6	A scheme of the ADCP deployment and recording of ship wakes has to be presented to understand how the ship wakes are recorded by the ADCP. For instance, the bubble wake manifestations similar to one in Fig.2 appear when the ADCP is towed across the ship wake, or if the wake is moving in the cross wake direction due to currents passing by the ADCP beams. How thus the record of a ship wake in Fig.2 could be obtained for a stationary looking upward ADCP? Was that due to a current moving a wake through a zone illuminated by ADCP?	We appreciate the suggestion of adding as scheme to describe the instrument deployment and recording. We suggest adding a figure describing the instrumental setup in the material and methods section and a more detailed description of what is measured and how. Moreover, we suggest adding an additional sketch in Figure 2 in the manuscript, to further illustrate what the ADCP is recording. In addition, we have also made some complimentary illustrations that could be added to the manuscript or a supplementary information section, if requested.	A sentence has been added to clarify that the ADCP measures the wake development thorough time, with a reference to figure 3. A figure of the experimental setup has been added. Previous figure 2 (now 3) has been revised to include an indication of the ship passage to clarify what the ADCP instrument is measuring. The possible reasons to why the wake signal is detected from ships passing up to 184 m from the ADCP instrument is discussed in detail in the answer to question 7.

7	Why the wakes appeared in the ADCP records for ships passed by at some distances from the ADCP? Because of the wake turbulent diffusion? If so, why not to analyze, e.g. the characteristic times of the turbulent diffusion, the diffusion spatial/temporal decay, etc.?	We consider two main reasons for the wakes appearing in the ADCP records for ships passing at some distance from the wake. Firstly, currents can move the wake towards the instrument. However, when looking at the impact of currents in our data, we could not see that that current speed and/or direction was affecting the existence of a wake in our data. Moreover, during most of the measurement period the current speed at the instrument position was very low. These two circumstances indicate that wake drift due to currents is probably not the main reason that wakes are detected from ships passing at long distances from the instrument. The second possible reason, as suggested, is that turbulent diffusion widens the wake. As shown in the satellite data, the median thermal wake width was 157 mindicationg, that a ship passage at 150–200 m away from the instrument could induce a wake, where the edge of the wake eventually reaches the instrument. In addition, the wake will be widened by buoyancy effects if the water is stratified. There is also a reason to why we have not analysed the characteristic times of the turbulent diffusion. As the ADCP only measure in one point, it is not clear if the entire length of the wake is captured or which part of the wake is being measured. Since it is not sure that the entire wake from start to end is measured, making calculations on diffusion rate based on it would not be fully representative. For the aim of the current paper – estimating the temporals extent of the turbulent wake, we considered the large quantity of measurements sufficient to give an estimate of the overall/general temporal extent, without analysing the characteristic times of the turbulent diffusion.	A sentence has been added to explain why wakes from ships passing at some distances from the ADCP can be detected.
8	Line 301. I cannot understand how this can happen : "…when two ships passed the instrument at the same time"	The reason why two ships can "pass at the same time" are two. Firstly, large ships may require pilot assistance and/or tugboats to enter the harbour. In these cases, the ships pass right next to each other and it is impossible to separate the wakes form the different ships (see figure xx in supplementary info). Most of the double passages in the dataset where these types of occasions. The other occasion of double passages is related to the width of the wake and the fact that the wake from ships can be detected even when ships are passing at a distance from the instrument. There were a few occasions where two ships passed the instrument at the same distance from the instrument, but on different sides (figure xx in supplementary info). In these cases, it was not possible to tell which of the ships that induced the wake detected by the ADCP, as our analysis of the data showed now clear correlation between current direction and wake detectability. Hence wakes from these occasions were treated as a double passage, as the detected wake could come from either ship or be a combination of both passages. We will add a sentece clarifying in which occasions two ships can "pass at the same time".	A clarification to how to ships can pass the instrument at the same time has been added.

9	Categorization of the ships in the context of their turbulent wakes does not look physically justified. More reasonable would be to relate the wakes to the ship weight, draught, speed, possibly to the size/number of propellers	We acknowledge the comment to relate the wake depth and longevity to another parameter than ship type, and therefore suggest a revision of the result section. It is beyond the scope of this paper to investigate the dependence between various non-dimensional parameters, which would be the most physically justified thing to do. However, we expect that the wake size to a large degree depends on the force or power put into the water by the propeller. We therefore propose replacing the current figures 5 and 6, to figures relating the wake depth and longevity to force (F), calculated as ρ*ship width*ship draught*ship speed^2 [kg m s^{-2}], with seawater density (ρ) equal to 1025 kg m^{-3}. This parameter is proportional to ship drag and will relate the wake depth and longevity to vessel size and speed, which we agree are parameters expected to have an impact on the formation of the turbulent wake. In addition to the change in figures, we also propose a change to table 2 and 3 in the manuscript. We will remove the ship type category statistics and instead include statistics for the close passage category (0-3 ship widths). The change of figures and tables will naturally be accompanied with a revised description and analysis of the result.	The methods section has been revised to describe the new presentation of the data, removing the ship type categories and adding the calculation of ship Force. The result and discussion section has been revised to describe the new presentation of the data, removing the ship type categories and adding the calculation of ship Force. This provides a basis for discussing possible physical explanations to the variation in wake depth and longevity in the detected wakes, as well as the importance of ship size and speed.
10	line 331 "As the fraction of detected induced wakes at similar distances differ between ship types, it is an indication that the ship type impacts the characteristic of the turbulent wake" . I disagree with the statement and I think that the difference is determined mostly by the ship weigh and ship speed.	This comment is addressed in detail in the answer to comment 5. As we now suggest replacing the figures which presents the data based on ship type, this sentence will be removed.	This comment is addressed in detail in the answer to comment 9. As we now suggest replacing the figures which presents the data based on ship type, this sentence will be removed.

11	The paper is full of obvious, trivial statements, e.g. “in general the deepest wakes were caused by ships passing closer to the instrument, whereas ships passing at larger distances from the instrument (100–199 m) mainly caused shallower wakes : : :” (lines 369-370) “the maximum dissipation rates : : : in the core of the wake : : :.are : : :..much larger than what is usually observed in the core of, or below, the surface mixed layer” (lines 403-405), etc. etc.	We acknowledge and understand there are statements in the manuscript that can be perceived as trivial and obvious, depending on the researcher’s specialisation. To balance the content to suit a diverse audience, from different highly specialised disciplines, is a general challenge in interdisciplinary research. Therefore, the second example in this comment was included for a reason. Our aim is to reach an interdisciplinary audience within ocean science, in line with the scope of this journal. This specific comment was included because the non-oceanographic co-authors of the paper explicitly asked for a comparison between our measured values and values that would occur naturally in the system. We believe that these types of statements fill an important function in making the content of the paper more accessible to an interdisciplinary audience. However, with the new suggested figures, much of the result section will be rewritten, and will make sure not to pay extra attention to this aspect.	With the revised result section, the specified sentences have been removed.
12	The main finding of this study is that turbulent ship wakes can reach deeper than the previously observed values, no physical explanation or discussion is provided.	In relation to the new figures in the revised result section, we will develop the discussion regarding possible physical explanations to the variation in wake depth and longevity. See answer to commment 1 and 9.	See answers and revisions to Comment 1 and 9.
13	It is expected that ship specifications and speeds determine ship wake depths, so the authors should be able to discuss further based on the available ship information.	We acknowledge the comment to relate the wake depth and longevity to another parameter than ship type, and therefore suggest a revision of the result section. Se answer to reviewer comment 9.	

14	It is not clear why the authors chose two separate locations for the in-situ and remote sensing study. Is it possible to find satellite imagery for the in-situ measurement period?	We appreciate this very relevant comment. There is of course satellite imagery available for the in-situ study area as well, and we have retrieved it for the in-situ study period. However, the satellite passes only every 16th day, and during the study period there were no cloud-free images. Regarding why we chose two different study areas all together, it was due to logistical reasons. The Bornholm study area was chosen as the ideal spot for the satellite study, as it one of the most intensely trafficked areas in the Baltic Sea. Initially, instruments were placed in the satellite study area, but due to unfortunate events they were lost. To have a better possibility to monitor the instrument, but still have an area with a lot of ship passages, we chose the Gothenburg harbour area for the second attempt of an in-situ study. The reason we still chose to keep the Bornholm are for the satellite study, was due to the more favourable weather conditions. A cloud free sky is needed for the satellite images to be usable, and since it only passes two times per month, the rainy Gothenburg area was ill suited for the satellite study. We did not find this information suitable to include in the manuscript, but if requested we will motivate the choice of study area in more detail.	A motivation to why the Bornholm study area was chosen for the satellite analysis, instead of the Gothenburg study site, has been added in the method and discussion section.
15	Also, vertical profiles should have been measured more frequently to see the effect of wakes on stratification and mixing.	We agree and acknowledge this as the main potential improvement of our observations, which we have also discussed in this section of the manuscript. Furthermore, the long-term aim of our research is to be able to study and discuss the effect of wakes on stratification and mixing. However, the aim of the current study was to describe the spatiotemporal scales of the turbulent wake, and not resolving all the parameters determining the effect of wakes on stratification and mixing. As mention in a previous answer, we realise that the current title and aim, could give the impression that the paper is of a more explanatory nature. However, we consider the current amount of vertical profiles enough for the aim of this paper and suggest leaving the discussion regarding the effect of wakes on stratification and mixing for the next paper (where high-resolution profiling will be included).	To address this comment, we will clarify the aim of the paper as being mostly descriptive.
16	Line 191: Capitalize Python	To address this comment, we will Capitalize Python on line 191.	Python has been capitalised.
17	Line 224: Why 15%? Please justify.	As stated, the wake area was manually defined using imagery of the echo amplitude and dissipation rate of turbulent kinetic energy. The exact value used for delimiting the wake region, was manually adjusted for each wake, using trial and error until the defined area sufficiently overlapped with the wake region in the image. Hence, the limit was chosen as the approximate value were noise was excluded but the wake region was included. However, this value was not decided upon in advance, but rather after the analysis it was clear that most of the values were approximately 15% higher compared to the daily/nightly mean.	It has been clarified that the value of 15 % was not chosen in advance but was manually adjusted based on visual scrutiny of plotted figures, but that most values were ~15% higher than the daily/nightly mean.

18	Line 355-359: Why do the bubble wakes look different from turbulence? Please discuss further.	We acknowledge that this can be discussed further. The bubbles are an indication that surface water is mixed down at depth and has been mixed with the ambient water. They will remain there, or rise or collapse with time depending on size etc. The dissipation, on the other hand, is a measure of the turbulent motions in the water that mixes the water down. When the turbulence decays (due to dissipation) the dissipation also decays and dies out. But the bubbles may remain after that has happened. The dissipation estimate is also influenced by neighbouring cells (equation 1), so the estimate may be deeper just due to the method used. I.e. if there is strong turbulence in one cell and none in the next, the method may still show some turbulence in the calm cell.	The discussion regarding the differences between the bubble and turbulent wake has been expanded.
19	Line 550-552: This seems to be a negative result: the stratification was not affected by the wake. Remove this part. I suggest that the authors measure more vertical profiles in the study area and/or provide literature for more data to characterize general and unusual environmental conditions. 4 casts x 2 days are not enough.	See answer to comment 15 regarding measuring more vertical profiles. However, regarding the specfic section mentioned here, we humbly disagree that it is a negative result. Turbulence will not be able to reach below the mixed layer, to 17.5 m depth, without mixing the water. Thus, we find it highly unlikely that the thermal stratification at 5 m was present within the wake. This means that the stratification was influenced by the wake and that waters above and below the thermocline were mixed with each other. However, three hours later the water has re-stratified and the mixed water has spread out laterally. We do not claim that there was a long-lasting effect on the stratification. However, during the longevity of the wake the stratification was most likely affected. Hence, in a scenario with very frequent ship passages, there will be less time for the re-stratification to occur, and a more long-lasting effect on the stratification could be possible.	The example of the ship passage affecting the thermal stratification has been clarified. For the comment regarding measuring more vertical profiles in the study area, see Comment 15.
20	Line 574-581: As mentioned above, please try to find satellite imagery that covers the in-situ measurement area.	See answer to comment 14.	A section has been added in the discussion and method, motivating the choice of the two different sites. A sentence has been added in the discussion suggesting future improvements of the method. See answer to Comment 14.
21	Line 587: Note that winds are also important.	We agree that winds are important, both as they affect waves and currents. As we have measurements of the water speed and waves, we consider the wind effect on currents and waves to be captured by those measurements. However, we acknowledge that the wind should still be mentioned.	Wind is now also mentioned. As we have measured the water speed and waves, we consider the wind effect to be captured by those measurements.
22	Line 618: What parameters?	To address this comment, we will name the parameters that have been discussed in the paper in relation to the ship wake depth, to make it clear which parameters we mean.	The parameters have been specified.

In situ observations of turbulent ship wakes and their ~~potential implications for vertical mixing~~spatiotemporal extent

Amanda T. Nylund[1], Lars Arneborg[2], Anders Tengberg[1], Ulf Mallast[3], Ida-Maja Hassellöv[1]

[1]Department of Mechanics and Maritime Sciences, Chalmers University of Technology, Gothenburg, 412 96 Gothenburg, Sweden.
[2]Swedish Meteorological and Hydrological Institute (SMHI), Gothenburg, 426 71 Västra Frölunda, Sweden.
[3]Department Monitoring and Exploration Technologies, Helmholtz Centre for Environmental Research, Leipzig, 04318 Leipzig, Germany.

Correspondence to: Ida-Maja Hassellöv (ida-maja@chalmers.se)

Commented [AN1]: **Comment 4** (R1, general comment): The aim/motivation of the paper is not clear.

Commented [AN2R1]: Changes:
- The title has been changed to clarify that the aim of the paper to describe the spatiotemporal extent of turbulent ship wakes.

Abstract. In areas of intensive ship traffic, ships pass every ten minutes. Considering the amount of ship traffic and the ~~fact that global maritime trade is~~ predicted ~~to~~ increase in global maritime trade, there is a need to consider all ~~effects~~ type of impacts shipping has on the marine environment~~; both~~. While there is increasing awareness about, and efforts to reduce, chemical pollution from ships, less in known about ~~and~~ physical disturbances and ship-induced turbulence has so far been completely neglected. To address the potential importance of ship-induced turbulence on e.g. gas exchange, dispersion of pollutants, and biogeochemical processes, there is an immediate need for characterisation of the temporal and spatial scales of the turbulent wake. ~~This paper studies a previously disregarded physical disturbance, namely ship-induced vertical mixing in the turbulent wake. A characterization of the temporal and spatial scales of the turbulent wake is needed to estimate its effect on gas exchange, dispersion of pollutants, and to identify in which areas ship-induced vertical mixing could have an impact on local biogeochemical cycles.~~ There is a lack of field measurements of turbulent wakes of real-size ships, and this study addresses that gap by *in situ* and *ex situ* measurements of the depth, width, length, intensity and longevity of the turbulent wake for ~240 ship passages of differently sized ships. A bottom-mounted Acoustic Doppler Current Profiler (ADCP) was placed at 32 m depth below the ship lane outside Gothenburg harbour, and used to measure wake depth and temporal longevity. Thermal satellite images of the Thermal Infrared Sensor (TIRS) onboard Landsat 8 were used to measure thermal wake width and spatial longevity, using satellite scenes from the major ship lane North of Bornholm, Baltic Sea. Automatic Information System (AIS) records from both the investigated areas were used to identify the ships inducing the wakes. The results from the ADCP measurements show median wake depths of ~10 m, and several occasions of wakes reaching depths > 18 m. The temporal longevity of the wakes had a median of around 8 min and several passages of > 20 min. The satellite analysis showed a median thermal wake length of 13.7 km, and the longest wake extended over 60 km, which would correspond to a temporal longevity of 1 h 42 min (for a ship speed of 20 knots). The median thermal wake width was 157.5 m. The measurements of the spatial and temporal scales are in line with previous studies, but the deep mixing and extensive longevity presented in this study, has not previously been documented. The results from this study ~~have shown~~show that ship-induced ~~vertical mixing~~turbulence occurs at temporal and spatial scales large enough to imply

that this process should be considered when estimating environmental impact from shipping in areas with intense ship traffic. Moreover, the possibility that deep vertical mixing could occur in a highly frequent manner highlights the need of further studies to better characterize the spatial and temporal development of the turbulent wake.

1 Introduction

The shipping industry holds a key role in today's society, as 80–90 % of all global trade is transported via ship (Balcombe et al., 2019). In areas of intensive ship traffic, e.g. in the Baltic Sea, there can be more than 50.000 ship passages annually, which in turn is approximately one ship passage every ten minutes (HELCOM, 2010). Yet, maritime trade is predicted to increase by 3.4 % annually until 2024 (UNCTAD, 2019). Transport by ship is also advocated as the most energy efficient as it in general has low carbon footprint per tonne and distance of transported goods (Balcombe et al., 2019). However, the carbon footprint is only one of many environmental impacts from shipping, and to fully estimate the impact of this growing industry, a holistic assessment is needed (Moldanová et al., 2018). To make a reliable holistic assessment, all types of impacts on the marine environment need to be considered, both from polluting and physical disturbances. This paper will focus on a previously disregarded physical disturbance from shipping, namely ship-induced turbulent wakes and their spatiotemporal extent~~vertical mixing~~.

When a ship moves through water, the hull and propeller create turbulence, which forms a turbulent wake behind the ship, characterised by an increased turbulence and a dense bubble cloud (NDRC, 1946; Soloviev et al., 2010; Voropayev et al., 2012; Francisco et al., 2017). ~~In a natural marine system, the water column is often stratified due to surface heating and/or freshwater influence. The wake turbulence interacts with this stratification by mixing the water and entraining deeper waters into the wake. The stratification may, in turn, reduce the vertical extent of the wake relative to what it would have been in a homogeneous water column (e.g. Voropayev et al. (2012)).~~ Knowing and being able to properly characterise temporal and spatial scales of the turbulent wake has several reasons. ~~A good characterization of the temporal and spatial scales of the turbulent wake is necessary~~It can be used to estimate the distribution~~, dispersion and dilution~~ of contaminants and pollutants ~~that are~~ discharged from ships (Katz et al., 2003; Loehr et al., 2006). Furthermore, the bubbles created in the turbulent wake can affect the gas exchange between ocean and atmosphere, in addition to the increased gas exchange due to the turbulence itself (Trevorrow et al., 1994; Weber et al., 2005; Emerson and Bushinsky, 2016). Moreover, in areas with intense ship traffic, the ship-induced vertical mixing could possibly affect nutrient availability and natural biogeochemical cycles in seasonally stratified waters.

~~During periods of seasonal stratification, nutrients in the surface layer are depleted, and the supply of nutrients from below is limited due to damping of the vertical mixing by the stratification (Reissmann et al., 2009; Snoeijs-Leijonmalm and Andrén, 2017). In coastal regions, nutrients can be brought up to the upper mixed layer by coastal upwelling, but in the open ocean,~~

the nutrient supply is dependent on vertical mixing (Reissmann et al., 2009). If the vertical mixing is intense and deep enough, it will bring up nutrient rich water from below the stratification to the upper surface layer, which can increase primary production and sustain algal blooms. In ocean systems unaffected by human activities, vertical mixing in the surface layer is induced by wind, and the depth of the mixing depends on the wind strength and duration, as well as the input of buoyancy from heating and fresh water (Thorpe, 2007). In temperate oceans, the seasonal stratification occurs during the summer season, which is also the period with the least wind (Reissmann et al., 2009). Thus, in unaffected seasonally stratified waters, there is little vertical mixing during the summer months. However, in areas with intense ship traffic there is an input of ship-induced vertical mixing. Consequently, if the depth of the ship's vertical mixing is similar to the stratification depth, intense ship traffic has the potential to regionally affect natural biogeochemical cycles.

Up to now, the environmental impact of ship-induced vertical mixing has been overlooked. There are few studies about ship-induced turbulence in general and none investigating the possible environmental impact of ship-induced vertical mixing. Remote sensing approaches focused on detecting wakes from a surveillance perspective (Fujimura et al., 2016) or the theoretical possibility of doing so (Issa and Daya, 2014). These approaches mainly rely on Synthetic Aperture Radar (SAR) to identify sea surface roughness. Other studies focused on the vertical distribution of the turbulent wake for military purposes, with the interest of detecting the wake and minimizing the wake signal (Smirnov et al., 2005; Liefvendahl and Wikström, 2016). Moreover, the formation and distribution of the bubble cloud in the turbulent wake has been in focus, rather than the turbulence and mixing. Besides the different foci, most of the available studies are numerical modelling studies of ship wakes. Measurements are on model-scale ships for validation (Carrica et al., 1999; Parmhed and Svennberg, 2006; Fu and Wan, 2011; Liefvendahl and Wikström, 2016), which generally only resolve the wake for distances up to a ship length after the ship. In real world, temporal and spatial scales of the turbulent wakes are significantly larger. Turbulent processes are difficult to investigate at laboratory scale, since the Reynolds number is much too small in the laboratory and the results can therefore not be expected to represent turbulence in nature. Thus, there is a lack of field measurements of the turbulent wake of real-size ships that allow to evaluate to what extent ships induce vertical mixing and which scales they really cover (Carrica et al., 1999; Parmhed and Svennberg, 2006; Ermakov and Kapustin, 2010).

The few studies that are based on field measurements or focus on the spatial and temporal scales of the turbulent wake, report measured wake depths between 6–12 m (Table 1). Measured The reported wake widths are more varied, with a range of 10–250 m (Table 1). This large variation could partly be due to the different methods used to define the wake region, as well as the difference in size and type of the investigated vessel. The longevity of the wake has been measured both as a temporal duration and as a length. Already in 1946, the United States National Defense Research Committee (US NDRC) reported detectable bubbles and temperature differences in the turbulent wake 30–60 min after ship passage. Trevorrow et al. (1994) made measurements of the temporal scale of the turbulent wake and reported strong acoustic scatters from the bubbles in the wake for 7.5 min after passage. Furthermore, Voropayev et al. (2012) conducted experiments with a model

~~ship in a thermally stratified tank, to investigate the temporal longevity of the turbulent wake. According to their results, the turbulent mixing subsided at about 30 ship lengths behind the ship, which would correspond to 4.5 km for a 150 m long ship. This longevity estimate is supported by the results of~~ Soloviev et al. (2010) even~~, who~~ reported that ~~the~~ bubbles from the turbulent wake were visible from 10–30 min after ship passage, corresponding to a distance of 4–10 km, for a ship with a speed of 12 knots. It becomes clear ~~Clearly, there are studies showing~~ that the turbulent wake can reach depths of 10–15 m and can have a longevity of up to 30 min and/or 10 km. However, except Trevorrow et al. (1994) and NDRC (1946), information of wake width, length, or duration were always a by-product of these studies. Therefore, they naturally lack ~~none of these studies had the aim to investigate the temporal and spatial scales of the turbulent wake, and they lack~~ simultaneous measurements of depth, width, and length of the turbulent wake, as well as a statistical sound and reliable data basis with higher number and variety of vessels (type, speed, size). Thus, there is currently too few ~~a lack of~~ field measurements of the turbulent wake of real-size ships, to reliably ~~that allow to evaluate~~estimate ~~to what extent ships induce vertical mixing and which scales they really cover~~the temporospatial scales of turbulent wakes (Carrica et al., 1999; Parmhed and Svennberg, 2006; Ermakov and Kapustin, 2010). ~~Moreover, these studies have only included measurements of 1–4 vessels, and only Katz et al. (2003), Ermakov and Kapustin (2010) and Weber et al. (2005) performed field measurements in stratified conditions. Hence, there is a current lack of data to reliably estimate the contribution from shipping to vertical mixing in stratified waters.~~

Field Code Changed

The aim of this study is therefore to obtain a reliable ~~n~~ overview of the magnitude of the spatiotemporal influence of ship-induced vertical mixing through the integration of ~240 observations of ship passages, and~~. Understanding the scales at which the wakes occur is the first step in estimating the environmental implications of ship-induced vertical mixing. To be able to estimate the environmental impact of the ship-induced vertical mixing, the spatial and temporal development of the turbulent wake needs to be characterised and the mixing quantified, for a large set of different ship types. In this study~~Here~~,~~ a combination of methods ~~has been used~~ to describe the depth, width, length, ~~intensity~~ and longevity of the turbulent wake ~~for a large set of ship passages (~240)~~. As the study has been conducted *in situ* and *ex-situ*, on different temporal and spatial scales, and includes ships of different types and varying size, it constitutes a solid base for a first estimate of the order of magnitude of the spatiotemporal ~~scales~~ extent of ~~ship-induced vertical mixing~~ turbulent ship wakes. A better understanding of the spatial and temporal ~~scales~~ extent of the turbulent wake, ~~makes it possible~~is needed to ~~estimate the effect on gas exchange, dispersion of pollutants, and to~~ identify ~~in which areas~~where ship-induced vertical mixing could have a~~n~~ significant impact on local biogeochemical cycles, and thus should be studied further. Moreover, it provides a basis for estimating the summed wake area in a region, where an effect on gas exchange could be expected. In addition, it will provide valuable information for monitoring in areas with intense ship traffic, as well as for studies of the dispersion of pollutants from ships. It will be of particular importance for the FerryBox community, as FerryBoxes perform continuous measurements onboard ships ~~e~~on route, often in major ship lanes that may lead to biased results compared to surrounding water. In short, increased knowledge about the spatiotemporal extent of turbulent ship wakes, makes it possible to identify when and where ship-induced turbulence needs to be considered.

Commented [AN3]: **Comment 15** (R2, general comment): Also, vertical profiles should have been measured more frequentl see the effect of wakes on stratification and mixing.

Commented [AN4R3]: Changes:
- The aim of the paper is to describe the spatiotemporal extent the wake, and not to investigate the effect on stratification and mixing. We acknowledge that the previous title and introductio might have given the impression that we aimed to discuss the effect on stratification and mixing. Therefore, we have address this comment by clarifying the aim of the paper by revising the title and the last paragraph of the introduction.

Table 1. Previously reported field measurements of the spatial and temporal scales of the turbulent wake. The method used to estimate the turbulent wake is indicated in the "Method" column. For studies where only the temporal wake longevity was measured, an estimate of the wake length has been calculated using the wake duration and a ship speed of 12 knots.

Study	Method	Wake depth [m]	Wake length [km]	Wake duration [min]	Wake width [m]
NDRC (1946)	Acoustic/thermal	3–10	11–22	30–60	40–90
Trevorrow et al. (1994)	Acoustic	6–12	2.8*	7.5	66 (average)
Katz et al. (2003)	Dye concentration	8–10	3**		
Weber et al. (2005)	Acoustic	8	6	15	
Stanic et al. (2009)	Acoustic		1.5–2	20	10
Ermakov & Kapustin (2010)	Acoustic	4–8	3.7–5.5*	10–15	40–80
Soloviev et al. (2010)	Acoustic	10–15	4–10*	10–30	
Gilman et al. (2011)	Visible surface trace				100–250
Soloviev et al. (2012)	Acoustic	7			
Francisco et al. (2017)	Acoustic	6–12	0.5*	1.5	

*Calculated based on temporal longevity and a ship speed of 12 knots, **Distance at which the max width was documented

2 Materials and methods

The data collection was conducted in two parts: one field study in the large ship lane outside Gothenburg harbour, and a satellite image analysis of sea surface temperature in the large ship lane north of Bornholm, Baltic Sea (Fig. 1). The field study covered the vertical scale and the temporal longevity of the turbulent wake, and the satellite image analysis was used to estimate the thermal wake width and spatial longevity.

2.1 Gothenburg harbour study

The field study was conducted off the Swedish west coast, in the large ship lane outside Gothenburg harbour (Fig. 1). Gothenburg harbour is the largest harbour in Scandinavia, with 120 port calls per week, including large container ships, oil tankers, car carriers, and passenger ferries (The Port of Gothenburg, 2020). The size of the harbour, the frequency of port calls, and the variety of ship types, makes it a suitable study area for ship-induced vertical mixing. The site of instrument deployment was outside the port area, under the fairway where all incoming large ships need to pass (Swedish Maritime Administration, 2020). It was also inside the area where tugboats and pilots are required when applicable, but outside the speed restriction area, thus ships were traveling at normal speed. For the *in situ* measurements, the Gothenburg site was considered more suitable compared to the Bornholm study area, as it was more easily accessible and the risk of losing the instrument to other maritime activities was lower.

Commented [AN5]: **Comment 14** (R2, general comment): It is not clear why the authors chose two separate locations for th situ and remote sensing study. Is it possible to find satellite imag for the in-situ measurement period?

Commented [AN6R5]: Changes:
- A motivation to why the Gothenburg study site was chosen i favour of the Bornholm study area, has been added.

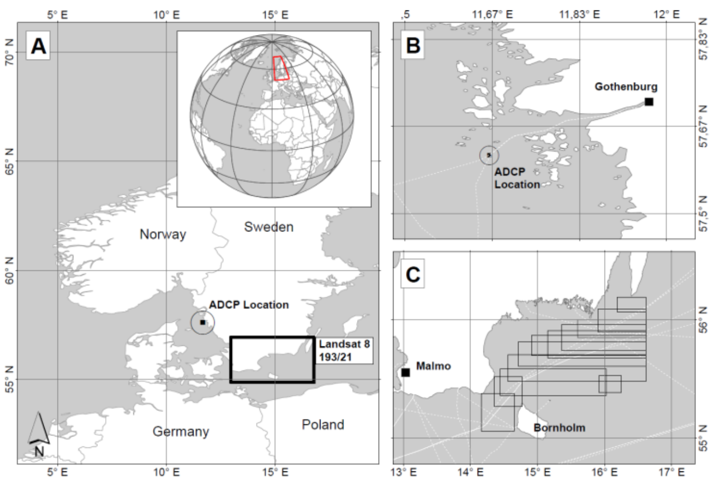


Figure 1: Overview of the two study areas (a) showing the location of the ADCP under the ship lane outside Gothenburg (b) and the area covered by the analysed satellite images (c). White dashed lines indicate ship routes of ferry lines and the boxes in (c) indicate the area defined as the ship lane area in the satellite image analysis. The ship lane and traffic separated zone north of Bornholm are shown in Figure 12.

2.1.1 Field measurements and data collection

A bottom-mounted Nortek Signature 500 kHz broadband Acoustic Doppler Current Profiler (ADCP) was deployed under the ship lane (57.61178 N, 11.66102 E), fixed in upward-looking position in a bottom frame (Figure 2). Similar setups have previously been used to study the bubble cloud of the turbulent wake, by Trevorrow et al. (1994) and Weber et al. (2005). The instrumental setup provides measurements of the overlaying water column trough time (Figure 3), hence, recording the wake development in a fixed point through time. Under the assumption of a stationary wake moving with the ship velocity, the observations can also be interpreted in terms of the spatial change of the wake with distance from the ship. The instrument was deployed at approximately 30 m depth, for a duration of 4 weeks (28 August to 25 September 2018). The ADCP measured along beam current velocities, using four slanted beams (25° angle) and one vertical beam (ping frequency 1 Hz, cell size 1 m on all beams). The echo amplitudes from the beams were also used to detect the wake bubbles. All single ping data on currents and echo amplitude was stored on-board the instruments and analysed, see sect. 2.1.2. The range of sonar frequencies that are suitable for detecting bubbles in the turbulent ship wake is 30 kHz to 1 MHz and depends on the size of the bubbles in the wake (Liefvendahl and Wikström, 2016). A SonTek CastAway®-CTD (Xylem, San Diego, California) was used to measure salinity and temperature profiles at the time of the instrument deployment (August 28, 2018, 4 casts) and retrieval (September 25, 2018, 4 casts).

Commented [AN7]: **Comment 6** (R1, specific comment 2): A scheme of the ADCP deployment and recording of ship wakes to be presented to understand how the ship wakes are recorded by ADCP. For instance, the bubble wake manifestations similar to o Fig.2 appear when the ADCP is towed across the ship wake, or if wake is moving in the cross wake direction due to currents passin the ADCP beams.

(this part of the question is addressed further down)
How thus the record of a ship wake in Fig.2 could be obtained fo stationary looking upward ADCP? Was that due to a current mov a wake through a zone illuminated by ADCP?

Commented [AN8R7]: Changes:
-A figure illustrating the deployment setup has been added.
- A sentence has been added to clarify that the ADCP measure wake development thorough time, with a reference to figure 3.

Commented [AN9]: **Comment 15** (R2, general comment): Also, vertical profiles should have been measured more frequentl see the effect of wakes on stratification and mixing.

Commented [AN10R9]: To address this comment, we will clarify the aim of the paper as being mostly descriptive. See introduction and discussion for changes.

Commented [AN11R9]: We agree and acknowledge this as main potential improvement of our observations, which we have discussed in section 3.4 Limitations and Future outlook. Furtherm the long-term aim of our research is to be able to study and discu the effect of wakes on stratification and mixing. However, the ai the current study was to describe the spatiotemporal extent of the turbulent wake, and not resolving all the parameters determining effect of wakes on stratification and mixing. We realise that the current title and aim, could give the impression that the paper is o more explanatory nature. However, we consider the current amou of vertical profiles enough for the aim of this paper and suggest leaving the discussion regarding the effect of wakes on stratificati and mixing for the next paper (where high-resolution profiling wi included).

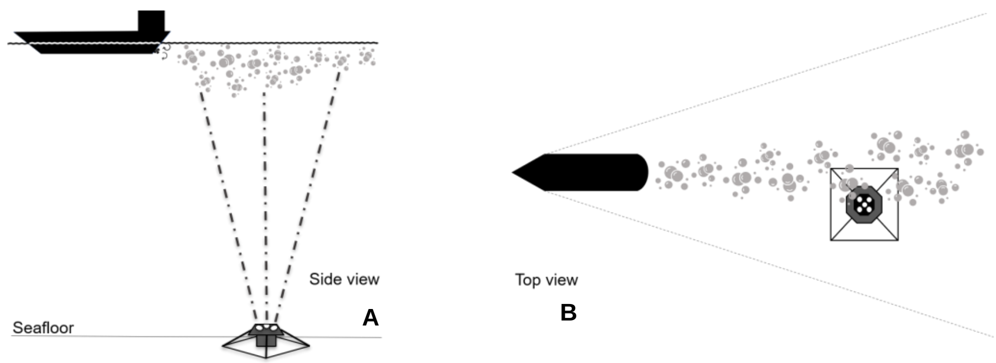


Figure 2: Scheme of instrument deployment, showing a) the sideview with perspective of the ADCP~~,~~ placed on the seafloor facing upward~~,~~ and recording the turbulent wakes during ~~a~~ ship passages, and b) a ~~bird~~top view perspective of the ADCP recording bubbles f~~o~~rom a turbulent wake, induced by a ship passing above, but slightly to the side of the instrument.

A dataset of the ships passing the study area during the field measurement period~~,~~ was purchased from the Swedish Maritime Administration. The dataset is from the Baltic Marine Environment Protection Commission (HELCOM) Automatic Information System (AIS) database, which is processed according to the procedure described in the annex of the HELCOM Assessment on maritime activities in the Baltic Sea 2018 (HELCOM, 2018). The Swedish Institute for the Marine Environment (SIME) provided additional files from the same HELCOM database, with AIS data for the analysed satellite scenes and the Gothenburg harbour study area. Vessel information from MarineTraffic – Global Ship Tracking Intelligence (www.marinetraffic.com) was used to retrieve detailed information about the width, length and draught of the ships in the dataset. ~~The SMA AIS dataset lacked information about the ship's current draught. Therefore, a dataset of the current draught of the ships visiting the Port of Gothenburg was retrieved from the Port of Gothenburg. As not all ships passing the instruments visited the Port of Gothenburg, the current draught was not available for all passing ships.~~

2.1.2 Data analysis

Compiling the ADCP wake dataset

All ship wakes in the dataset were identified manually using high resolution figures of the echo amplitude of the ADCP beams (see Fig. 3 for example). As the bubbles in the turbulent wake reflect~~s~~ the sound more efficiently than water, it results in an elevated echo amplitude in the turbulent wake region (NDRC, 1946; Marmorino and Trump, 1996; Trevorrow et al., 1994; Weber et al., 2005; Ermakov and Kapustin, 2010; Francisco et al., 2017). Generally, the wake signal could be clearly distinguished from bubbles induced by waves or signal noise from fish or zooplankton. However, ambiguous cases were noted, and the wake dataset was therefore divided into wake categories based on the quality of the wake signal. Each wake in the dataset was then linked to a ship passing in the vicinity of the ADCP, using the HELCOM AIS dataset and manual comparison. This introduced additional uncertainties, as not all wakes had clear match with a ship passage. After incorporating the matching uncertainties, the final wake categories used in the analysis were: "**wake**", only including clear wakes with one clear match or delayed match; "**double**", clear wakes where two or three ships passed the instrument at the same time; and "**no wake**", which included all passages within 184 m of the instrument that did not induce a visible wake, as well as all uncertain wakes and matches, which were mostly due to windy conditions which created noisy data. The distance at which a wake can be detected from a passing ship is affected by wake broadening, drifting, and ship width. In this study, ~~T~~the 184 m radius was chosen, as it was the furthest distance at which a clear wake and match was found in the dataset. There were two factors contributing to the existence of the "**double**" category. Firstly, the turbulent wakes in the dataset could be detected from ships passing at distances up to 184 m from the ADCP instrument. Hence, in cases when two ships passed at similar distances from the instrument at the time of a detected wake, it was not possible to distinguish which of the ships ~~that~~ induced the wake. Secondly, large ships may require pilot assistance and/or tugboats to enter the harbour. In these cases, the ships pass right next to each other and it is not possible to assign the wake to a single ship. ~~In addition~~Lastly, some wakes and passages were removed from the analysis altogether. These included ships with missing information in the AIS data (size information) and small sailing vessels, as they due to their small size and engine power were not deemed relevant for the investigated process.

Commented [AN12]: **Comment 8** (R1, specific comment 4): Line 301. I cannot understand how this can happen : "…when two ships passed the instrument at the same time"

Commented [AN13R12]: Change:
- The explanation of the passages in the double category has be expanded.

Commented [AN14]: **Comment 7** (R1, specific comment 3): Why the wakes appeared in the ADCP records for ships passed by some distances from the ADCP? Because of the wake turbulent diffusion? If so, why not to analyze, e.g. the characteristic times of the turbulent diffusion, the diffusion spatial/temporal decay, etc.?

Commented [AN15R14]: Changes:
- A sentence has been added to explain why wakes from ships passing at some distances from the ADCP can be detected.

(This comment is also addressed in the discussion section)

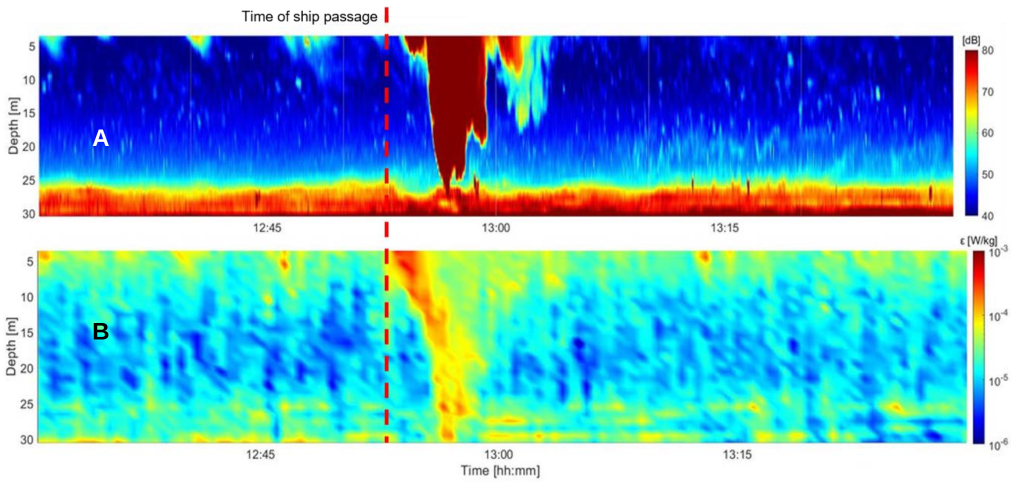


Figure 3: Example of the bubble wake signal in the echo amplitude dataset (a) and the calculated dissipation rate of turbulent kinetic energy, *ε*, (b) from one hour of ~~the~~ ADCP measurements. The upward facing ADCP was placed at approximately 30 m depth, repeatedly measuring the water column in one point. The dashed red line marks the time of ship passage. The high intensity red (a) and yellow (b) areas after the ship passage represent~~s~~ the wake region. The increase of *ε* down to the bottom is evidence of increased turbulence and ~~that there is~~a vertical mixing down to 30 m depth ~~in this case~~. The wake was induced by ~~the~~ a cargo ship ~~Ficaria Seaways~~ (width 25 m, length 229 m, draught 7 m), which passed the instrument at a distance of 34 m and a speed of 19 knots.

Commented [AN16]: **Comment 6** (R1, specific comment 2): Continuation…

How thus the record of a ship wake in Fig.2 could be obtained for stationary looking upward ADCP? Was that due to a current mov a wake through a zone illuminated by ADCP?

Commented [AN17R16]: Changes:
-A line indicating the time of the ship passage has been added the figure to further relate the measurements to the ship wake.
- The possible reasons to why the wake signal is detected from ships passing up to 184 m from the ADCP instrument is discus in detail in the answer to question 7.

Commented [AN18]: **Comment 16** (R2, specific comment): Line 191: Capitalize Python

Commented [AN19R18]: Changes:
-Python is capitalised.

Distance calculation, AIS and ADCP dataset

The AIS dataset included position reports for each ship every 2–10 seconds, which were used to calculate the ship's track. The closest distance between the ship-track and the vertical beam of the ADCP instrument was then calculated, using a local planar coordinate system, with the instrument at origo. The coordinates for the closest point on the track was also calculated, using the P~~p~~ython GeoPy package function distance.distance, and the points just before and after the closest point on the track were then identified.

Turbulence calculation, ADCP dataset

The dissipation rate of turbulent kinetic energy is a measure for the strength of the turbulence. Per definition it is the rate of energy conversion from kinetic energy to heat due to viscous friction in the smallest eddies, but in a stratified water column it is also proportional to the mixing between different water masses. There are various ways of determining dissipation rates. In the present work it is estimated from the ADCP data using the structure function method (e.g. Lucas et al. (2014)), which estimates the dissipation rate of turbulent kinetic energy from the second-order structure function following Eq. (1):

$$D_{11}(r,\Delta r) = \overline{\left(u_r{}'(r + \Delta r) - u_r{}'(r+\Delta r)\right)^2}, \qquad (1)$$

Commented [AN20]: Comment 5 (R1, specific comment 1): An error in formula (1)

Commented [AN21R20]: Changes:
- The error has been amended.

where u_r' is the fluctuating velocity in the r-direction (in this case the beam direction), Δr is the separation distance between two points along the beam, and overbar denotes time averaging. For separation distances shorter than the largest eddies the structure function relates to the dissipation rate and separation distance as in Eq. (2):

$$D_{11}(r,\Delta r) = C\varepsilon^{2/3}\Delta r^{2/3}, \qquad (2)$$

where C is a universal constant. Since the shortest distance (the ADCP bin size) was 1 m, the method is only expected to work for very strong turbulence with vertical eddy scales of magnitude larger than 2–3 m.

For each ship wake in the "**wake**" and "**double**" category, the along beam current velocity measurements from the ADCP were used for turbulence calculations in the wake region. One of the slanting beams was malfunctioning but the four remaining beams were analysed. A 1-hour dataset following each passage, identified by the start of the bubble cloud, was analysed. Spikes deviating more than four times the standard deviation from the mean in overlapping windows of 100 sec length were removed. Since the velocity signal of surface waves at different depths may be expected to be coherent whereas turbulent signals are not, the two Empirical Orthogonal Function (EOF) modes with largest variance were removed from the series to reduce the influence of surface waves. A fourth order Butterworth high-pass filter with cut-off period 600 sec was used to extract the turbulent velocity fluctuations. The dissipation rate of turbulent kinetic energy was estimated in 30 sec bins using the structure function method according to the method described in Lucas et al. (2014). One dissipation rate estimate was based on the average of the result for the three slanting beams (see Fig. 2 for an example), and another was based on the vertical beam.

Calculating wake depth, longevity, and maximum ε intensity, ADCP dataset

For each wake in the categories "**wake**" and "**double**", the wake region was defined for the parameters echo amplitude (bubble wake), dissipation rate of turbulent kinetic energy (ε), and the maximum velocity variance. To reduce noise in the dataset induced by turbidity at the sea floor, the data was normalised with respect to vertical distance from the instrument, assuming exponential decay of the signal strength. The wake region was defined by visual scrutiny of echo amplitude and ε figures (see Fig. 2 for an example) and manually annotated. comparing the wake region to the daily/nightly mean, and all values ~15% higher than the mean was considered part of the wake. As this procedure often identified noise as part of the wake, The elevation in echo amplitude/ε used for delimiting the wake region, as well as the depth and duration to consider, was manually adjusted for each wake to exclude noise. In general, the threshold was ~15% higher compared to the daily/nightly mean. both the percentage limit and the start time, stop time and maximum depth to include in the calculation, were manually defined for each wake to exclude noise. The deepest part of the wake region was used as a measure of the maximum wake depth and the maximum ε intensity in the wake region was used as a measure of the maximum turbulence.

Commented [AN22]: Comment 17 (R2, specific comment): Line 224: Why 15%? Please justify.

Commented [AN23R22]: Changes:
- It has been clarified that the value of 15 % was not chosen in advance but was manually adjusted based on visual scrutiny of plotted figures, but that most values were ~15% higher than the daily/nightly mean.

The duration of the wake (temporal longevity in min) was calculated using the start time and end time of the wake region. All calculations were pursued using an individually developed Python code.

Statistical analysis and graphical presentation of the ADCP wake dataset

For the statistical analysis ~~and graphical presentation,~~ the categories "**wake**", "**double**", "**close wake**", and "**no wake**" were used. The category "**close wake**", comprises all wakes induced by ships passing within 0–3 ship widths from the instrument, which roughly corresponds to 75 m. This cut-off was chosen as there was a substantial decrease in the percentage of induced wakes at passages > 3 ship widths from the instrument, indicating difficulties in detecting wakes at larger distances. As the wakes in the "**double**" category lack information about the ship inducing the wake, no "**double**" wakes are included in the "**close wake**" category~~.~~, ~~h~~Hence the double category is presented separately. ~~The dataset was then analysed by ship type, using the five categories *cargo*, *tanker*, *passenger*, and the double categories *cargo + pilot* and *tanker + pilot*.~~ For each ~~ship type~~category the median wake depth (m) and temporal wake longevity (min), was calculated for the bubble wake and the ε dissipation rate wake, together with standard deviation (std) and the 25th and 75th percentile. Furthermore, the percentage of ship passages that induced a visible wake in the ADCP beams was calculated along with the maximum ε intensity in the wake region. ~~A Welch Analysis of Variance (ANOVA) was also performed, comparing the maximum wake depth and longevity between the five ship type categories.~~

For the graphical presentation, the wake depth and longevity results are presented in relation to vessel force (F) [kg m s^{-2}]. F was calculated from the ship width (B) [m], draught (T) [m], and speed (s) [m s^{-1}], as in Eq. (3):

$$F = \rho * B * T * s^2 , \qquad (3)$$

with seawater density (ρ) equal to 1025 kg m^{-3}. The F parameter is proportional to ship drag and relates the wake depth and longevity to vessel size and speed, which are parameters affecting the formation of the turbulent wake.

2.2 Bornholm satellite study

The Bornholm study area was chosen, as it covers the most intensely trafficked ship lane in the Baltic Sea, with approximately 50,000 ship passages per year (HELCOM, 2010). All large ships heading for the Eastern and Northern ports of the Baltic Sea, must use the Bornholm ship lane (HELCOM, 2018), which makes it ideal for studying ship-induced vertical mixing from a variety of different ship types. Besides the purely traffic-related reason, a second reason for choosing the Bornholm area compared to the Gothenburg area, in which in.-situ data (ADCP, CTD) was retrieved was the availability of cloud-free satellite scenes, which is essential for detecting any surface object in the optical and thermal wavelength. The Bornholm area (path 193/ row 21) had 23 scenes with less than 23% cloud cover above the sea until August 2018, for the Gothenburg area (path 196/ row 20) only 9 scenes were available.

Commented [AN24]: **Comment 9** (R1, specific comment 5) Categorization of the ships in the context of their turbulent wakes does not look physically justified. More reasonable would be to r the wakes to the ship weight, draught, speed, possibly to the size/number of propellers.

Commented [AN25R24]: Changes:
- The section has been revised to describe the new presentatio the data, removing the ship type categories and adding the calculation of ship Force.

Commented [AN26]: **Comment 14** (R2, general comment): It is not clear why the authors chose two separate locations for th situ and remote sensing study. Is it possible to find satellite imag for the in-situ measurement period?

Commented [AN27R26]: Changes:
- A motivation to why the Bornholm study area was chosen fo satellite analysis, instead of the Gothenburg study site, has bee added.
- A comment regarding satellite images for the period of *in sit* measurements at the Gothenburg site has been added.

2.2.1 Data collection

All required optical and thermal infrared data from Landsat 8 were retrieved from https://s3-us-west-2.amazonaws.com. The study area for the Bornholm area in the Baltic Sea was covered by path/row 193/21 (see Fig. 1 for overview of study area).

2.2.2 Data analysis

Compiling the satellite dataset

To obtain average wake lengths and widths indicating vertical mixing on regional scales, optical, near-infrared and thermal-infrared bands from Landsat 8 were analysed. The dataset includes Landsat 8 data having a cloud cover < 23% (n=23). For optical and infrared data cloud coverage acts as opaque layer hindering to infer any information below it. The procedure includes a general and automatized data pre-processing scheme (Matlab), an automatic ship detection (Matlab) and a manual wake digitization (ArcMap). The pre-processing encompasses i) an automatic download of all available satellite scenes with less than 23% cloud coverage of the given path/row, ii) a masking of land areas using a combination of the modified normalized difference water index (MNDWI) after Xu (2006) and a Otsu-based threshold procedure (Otsu, 1979), iii) a masking of opaque and cirrus clouds classified as such based on the CFMask (Foga et al., 2017), and iv) finally a conversion from top-of-the-atmosphere (TOA) spectral radiances of band 10 to sea surface temperatures (SST) using transmission, downwelling and upwelling radiances modelled for each scene using a MODTRAN based online tool (Barsi et al., 2003).

Detecting ships was pursued semi-automatically following an optical approach similar to the one described by Heiselberg (2016). After masking, the remaining and analysable area is open water only. Spectrally, ships can be differentiated using the visual and short-wave-infrared part of the spectrum, even on the basis of coarser spatial resolution of 30 m as in the present case. As both parts of the spectrum are included in the MNDWI a global threshold of 0.09 was used on the MNDWI image for each scene to detect potential ships. To reduce the number of false positives due to unmasked cloud interference, a further selection criterion was added, using optical ship wake characteristics described in Gilman et al. (2011) and Heiselberg (2016), which is also visible in MDWNI space. Around all potential ships, a search window of 15x15 pixel (450x450m) was created. If MNDWI values > 0.13 representing ship wakes was detected, the potential ship was converted to a true ship, while remaining potential ships were neglected.

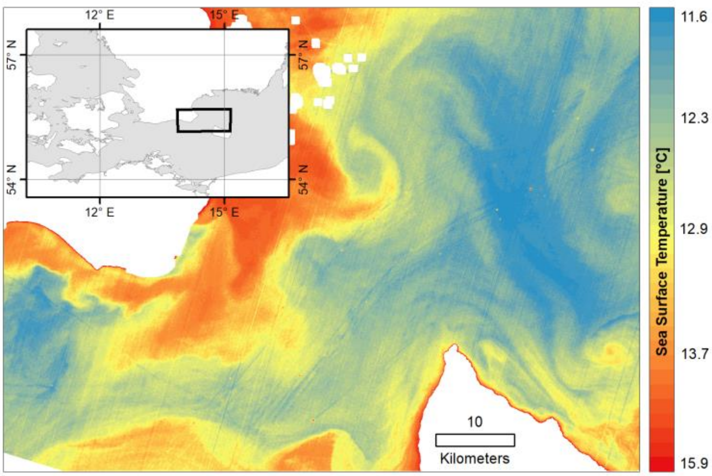


Figure 4: Example of satellite scene with visible thermal wakes in the Bornholm study area. Note the ships visible as warmer yellow dots and the thermal wakes visible as colder blue lines stretching behind the ships. The oblique regular stripes are not ship wakes, but a sensor induced radiometric artefact. Landsat-8 image courtesy of the U.S. Geological Survey.

Using the ships as spatial indication, all available 23 scenes were screened for thermally indicated ship wakes. In case of an occurrence, all thermal wakes for which a ship was detected, were digitalised. Using this approach, the wake lengths were obtained (see Fig. 4 for example of visible thermal wakes). To also retrieve wake widths, cross profiles were subsequently created in intervals of 250 m along the thermal ship wake, with a length of 400 m each. The cross-profile lengths were orientated at the maximum widths of <300 m presented in Gilman et al. (2011). Wake width was automatically determined analysing the local minima (thermal wake centre) and local maxima (surrounding uninfluenced water area) for each of the cross profiles.

Combining the satellite wakes with AIS data

Identified wakes and ships from satellite data were automatically matched against AIS data, to identify the ships inducing the wakes. All scenes were manually controlled to make sure the automatically matched ships were moving in the correct direction to have induced the wake. As the area of interest was the large ship lane north east of Bornholm, only the ships in the traffic separated part of the ship lane stretching from Bornholm to Öland's south tip, were included in the analysis (see boxed area in Fig. 1c). In addition to the matched satellite ships, all other ships present in the area at the time of each satellite scene were identified.

Statistical analysis of satellite wake dataset

For ~~the statistical analysis~~ the satellite dataset, ~~was analysed in its entirety and by ship type, using the categories *cargo, tanker, passenger*, and *other*. T~~the median spatial wake longevity (m) and wake width (m), was calculated, together with standard deviation (std) and the 25th and 75th percentile. The percentage of ship passages inducing visible thermal wakes, was also calculated.

3. Results and discussion

In the Gothenburg harbour study, there was a total of 96 detected turbulent wakes which could be successfully matched to a passing ship. In the Bornholm satellite image analysis, 144 thermal wakes were detected in the ship lane area, and successfully matched to a ship. Thus, a total of 240 ship wakes were included in the analysis, and the results from each study area will be presented separately below.

3.1 Gothenburg harbour study

During the measurement period, there ~~wa~~e~~res~~were a total of 413 ship passages within 184 m of the instrument. Of these passages, there were 65 occasions when two ships passed the instrument at the same time. As a double ship passage only induces one wake, these occasions were considered as one passage when looking at the percentage of passages inducing wakes. In addition, 15 other passages were removed due to data uncertainties ~~stemming~~ originating from entirely missing data (n=3), small size vessels irrelevant for the present study such as sailing/pleasure vessels (n=5), and multiple passages or wakes with unclear matches (n=7). This resulted in a total of 333 passages included in the analysis. 96 of those passages induced clearly visible wakes (29 %) due to single ship passages (n=69) and double passages (n=27). The close wake passages (< 3 ship widths from the instrument) comprised 57 % (n=39) of all single ship passages. ~~ship type categories with ≥ 9 wake-inducing passages (*cargo, tanker, passenger, cargo + pilot*, and *tanker + pilot*), comprised 87.5 % of the detected wakes, and wake depth and longevity for these categories are presented in section 3.1.4 and 3.1.5.~~ The statistical analysis of wake depth and longevity is presented for the entire dataset in section 3.1.4 and 3.1.5, together with a graphical presentation of the close wake passage dataset.~~All ship type categories not shown had few total passages (1–7), except the *pilot* category, which had 27 passages but only 2 induced wakes.~~

3.1.1 Environmental parameters

At the time of deployment, there was a clear stratification at 10 m depth, with an upper mixed layer salinity of 25.5, and a gradual increase of salinity below the stratification, reaching a maximum salinity of 32 at 32 m depth (Fig. 4). The temperature profile showed a rather uniform profile, with only a slight increase towards the surface, indicating that salinity was the main stratifying component (Fig. 4). The surface layer had a temperature of 18–18.6 °C, the middle layer ranged from 17.6°C at 10 m to 17.3 °C at 20 m, and the deepest layer went from 17.4 °C to 16.4 at the sea floor. At the time of

Commented [AN28]: **Comment 9** (R1, specific comment 5) Categorization of the ships in the context of their turbulent wakes does not look physically justified. More reasonable would be to r the wakes to the ship weight, draught, speed, possibly to the size/number of propellers.

Commented [AN29R28]: Changes:
- The wake depth and longevity are now presented in figures relating them to vessel Force, instead of ship type. This provid basis for discussing possible physical explanations to the varia in wake depth and longevity in the detected wakes, as well as t importance of ship size and speed.

instrument retrieval, there was only one clear ~~stratification~~ pycnocline at 5 m depth, with an upper mixed layer temperature around 14 °C and salinity around 27. The temperature below the ~~stratification~~ pycnocline was around 16 °C and the salinity was 33. This type of structure is usual in this area, as the Baltic Surface current which brings low saline water from the Baltic Sea~~, which~~ is on top of the more saline water from the Skagerrak (Snoeijs-Leijonmalm and Andrén, 2017). Note that the water column is unstable in temperature, so also here salinity is the stratifying component.

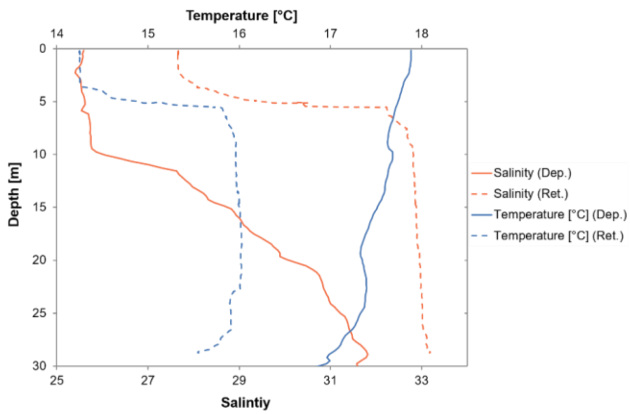


Figure 5: Salinity and temperature at the time of instrument deployment 28 August 2018 (solid lines) and retrieval 25 September 2018 (dashed lines).

3.1.3 Wake ~~occurrence~~detection rate

The wake ~~occurrence~~detection rate, ~~and maximum wake depth for th~~related to passing distances and vessel force (F)~~e ship type categories *cargo, tanker, passenger, cargo + pilot,* and *tanker + pilot*~~, ~~are~~ is show in Figure 6. For passages within 3 ship widths from the instrument (close wake category), the detection rate ranged between 36–100 %, with an average of 57 %. At distances > 3 ship widths, the wake detection was much lower (0–26 %). Due to the low detection rate in the two larger distance categories, only the close wake category will be used in the graphical presentation of how wake depth and longevity relate to vessel force. Surprisingly, the detection rate of wakes induced by ships passing at distances > 3 ship widths does not seem to be affected by the vessel force, as the percentage of detected wakes is similar for all force bins (Figure 6). Similarly, the close wake category does not show a clear correlation between vessel force and wake detection rate. However, more passages with large vessel force would be needed to be able to draw any conclusions, since the data is skewed towards lower vessel forces. Nevertheless, the results presented in Figure 6, indicates that passing distance affects the wake detection rate more than the vessel force.

~~Both the wake signal from the bubble wake and the turbulent kinetic energy dissipation rate (ε) are shown. The total number of passages and induced wakes for each ship category are shown in. For the *cargo* and the double ship type categories, the~~

percentage of induced wakes was high when the ship passed within 50 m of the instrument (75–100 %), and lower at distances > 50 m (0–30 %) (Fig. 5). Similarly, for the *tanker* category, with 70 % of the passages < 25 m inducing wakes, and 0–30 % for the rest of the distances. For the *passenger* category, on the other hand, only around 20 % of the passages induced wakes at a distance of 25–49 m from the instrument, but 40–50 % of the passages induced wakes at distances of 50–125 m from the instrument. As the fraction of detected induced wakes at similar distances differ between ship types, it is an indication that the ship type impacts the characteristic of the turbulent wake.

Commented [AN30]: **Comment 10** (R1, specific comment 6 line 331 "As the fraction of detected induced wakes at similar distances differ between ship types, it is an indication that the ship type impacts the characteristic of the turbulent wake" . I disagree with the statement and I think that the difference is determined mostly by the ship weigh and ship speed.

Commented [AN31R30]: This comment is addressed in deta in the answer to comment 9. As we now suggest replacing the fig which presents the data based on ship type, this sentence will be removed.

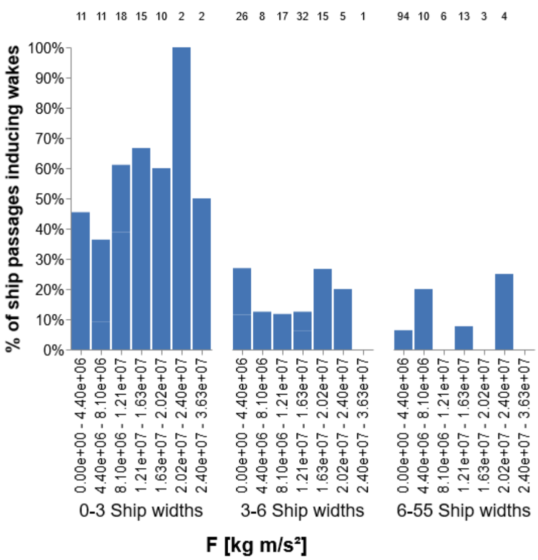


Figure 6. **Wake occurrence for three different categories of passing distances: 0–3, 3–6, and 6–55 ship widths from the instrument. For each distance category, the x-axis shows the force (F) of the passing vessel in Newton. The number above each bar indicate the total number of passages for that category. Note the cut-off in percentage detected wakes at passing distances > 3 ship widths.**

3.1.4 Maximum wake depth

The median maximum wake depth for all wakes was 9.5 m (std 4.2 m) for the bubble wake and 11.5 m (std 3.9 m) for the ε wake (**Error! Reference source not found.**). The close wake category had larger median values for both the bubble wake and ε wake, at 11.5 m (std 4.3 m) for the bubble and 13.5 m (std 3.7 m), respectively. It is worth pointing out that these ε wake depths were not the lower weak rim of the wake, as the threshold values defining the wake region mostly ranged between 10^{-4}–$10^{-3.5}$ W kg^{-1}. These threshold values are really large (e.g. Thorpe (2007)), indicating vigorously turbulent wakes, which probably were homogeneous down to the maximum depths of the wake region. Previous studies have mainly reported wake depths of 8–12 m (Table 1). In contrast, the results from this study show~~s~~ that 25 % of the detected bubble wakes were deeper than 12.5 m and 25 % of the ε wakes were deeper than 14.5 m (Table 2). The deepest detected wakes reached values of 27.5 m for the bubble wakes and 30.5 m for the ε wake. These values ~~were~~ are >10 m deeper than previously reported maximum depths.~~.~~

In Figure 7, the maximum wake depth is presented for the bubble wake and ε wake, in relation to vessel force (F). For the bubble wake, the percentage of induced wakes deeper than 12 m increases with increased vessel force, which can be seen in both the close wake category (Fig. 7a) and all single wake passages (Fig. 7b). There is a similar tendency for the ε wake, although it is more prominent in the figure of all single wake passages (Fig. 7d). However, there was no statistically significant correlation between F and maximum wake depth for either category. The lack of correlation could partly be explained by the skewed data distribution, as there were few passages with a large F (Figure 6). ~~Comparing the median maximum wake depth for the bubble wake and the ε wake, for the entire dataset, the ε wake was slightly deeper for all categories (~1 m), and the double wakes were deeper than the single wakes Fig. 5). Of the different ship type categories, the *cargo* ships had the deepest wakes, with a median of 10.5 m (std 4.6 m), and the *tankers* had the shallowest median wakes with 7.5 m (std 2.7 m).~~

Table 2. Mean, median, maximum value, first quartile (Q25), third quartile (Q75), and standard deviation (std), for wake depth and longevity, for the wake categories; close wakes (0–3 ship widths), single wakes, double wakes, and all wakes in the dataset.

Wake category	Bubble wake depth [m]						Bubble wake longevity [min]					
	Mean	Median	Max	Q25	Q75	Std	Mean	Median	Max	Q25	Q75	Std
Close	11.8	11.5	27.5	9.5	13.5	4.3	00:11:00	00:09:59	00:28:59	00:06:29	00:13:15	00:06:34
Single	10.3	9.5	27.5	7.5	12.5	4.1	00:10:14	00:08:00	00:28:59	00:05:29	00:13:29	00:06:29
Double	11.2	10.5	22.5	8.5	13.5	4.4	00:12:21	00:11:29	00:23:29	00:07:00	00:19:00	00:06:23
All	10.6	9.5	27.5	7.5	12.5	4.2	00:10:50	00:08:44	00:28:59	00:05:53	00:15:45	00:06:29

	ε wake depth [m]						ε wake longevity [min]					
	Mean	Median	Max	Q25	Q75	Std	Mean	Median	Max	Q25	Q75	Std
Close	13.4	13.5	30.5	11.5	14.5	3.7	00:06:17	00:05:59	00:13:30	00:04:45	00:07:44	00:02:33
Single	11.8	11.5	30.5	9.5	13.5	3.9	00:06:22	00:05:59	00:13:59	00:04:59	00:07:59	00:02:41
Double	12.9	11.5	19.5	9.5	17.0	3.8	00:09:07	00:08:00	00:20:00	00:06:44	00:10:14	00:03:53
All	12.1	11.5	30.5	9.5	14.5	3.9	00:07:08	00:06:30	00:20:00	00:05:00	00:08:30	00:03:18

	Distance to instrument [m]						
	Mean	Median	Max	Q25	Q75	Std	n
Close	32	29	82	16	42	21	39
Single	64	46	184	26	101	51	69
Double	31	18	120	9	46	32	27
All	55	38	184	16	82	49	96

Commented [AN32]: **Comment 12** (R2, general comment): The main finding of this study is that turbulent ship wakes can rea deeper than the previously observed values, no physical explanati or discussion is provided.

Commented [AN33R32]: See answers and revisions to **Comment 9**.

Mean, median, first quartile (Q25), third quartile (Q75), standard deviation (std), minimum value, and maximum value for wake depth and longevity for the five main ship type categories, the single wakes, the double wakes and for all wakes.

	Bubble wake depth [m]							Bubble wake longevity [mm:ss]						
	Mean	Median	Q25	Q75	Std	Min	Max	Mean	Median	Q25	Q75	Std	Min	Max
Cargo	11.3	10.5	8.5	13.5	4.6	5.5	27.5	10:29	08:30	06:29	13:00	05:28	02:00	24:29
Tanker	8.3	7.5	6.5	9.5	2.7	4.5	13.5	07:42	05:00	03:30	07:07	07:48	00:59	28:30
Passenger	10.0	9.5	7.5	12.0	3.7	5.5	17.5	12:12	10:29	06:30	16:30	07:17	02:29	28:59
Pilot + Cargo	9.8	10.5	9.5	10.5	3.1	3.5	15.5	11:26	07:30	06:30	17:59	06:48	03:59	20:59
Pilot + Tanker	11.1	8.5	6.8	14.8	5.5	6.5	22.5	09:56	08:30	05:30	10:52	05:51	04:29	20:29
Wakes	10.3	9.5	7.5	12.5	4.1	4.5	27.5	10:14	08:00	05:29	13:29	06:29	00:59	28:59
Double	11.2	10.5	8.5	13.5	4.4	3.5	22.5	12:21	11:29	07:00	19:00	06:23	03:59	23:29
All	10.6	9.5	7.5	12.5	4.2	3.5	27.5	10:50	08:44	05:53	15:45	06:29	00:59	28:59

	ε wake depth [m]							ε wake longevity [min]						
	Mean	Median	Q25	Q75	Std	Min	Max	Mean	Median	Q25	Q75	Std	Min	Max
Cargo	12.5	12.5	9.5	14.5	4.4	5.5	30.5	06:12	05:30	05:00	07:29	02:16	02:30	13:30
Tanker	9.8	9.5	8.3	12.5	2.5	5.5	13.5	05:32	05:29	04:07	06:14	03:06	00:29	12:00
Passenger	11.9	11.5	9.0	14.5	3.6	7.5	19.5	07:39	07:30	05:59	09:15	02:57	02:30	13:59
Pilot + Cargo	12.1	11.5	11.5	12.5	2.7	7.5	17.5	07:56	06:30	05:59	08:00	03:14	05:30	15:30
Pilot + Tanker	12.0	9.5	9.5	16.0	4.4	7.5	18.5	09:50	08:00	07:14	13:06	05:32	03:29	20:00
Wakes	11.8	11.5	9.5	13.5	3.9	5.5	30.5	06:22	05:59	04:59	07:59	02:41	00:29	13:59
Double	12.9	11.5	9.5	17.0	3.8	7.5	19.5	09:07	08:00	06:44	10:14	03:53	03:29	20:00
All	12.1	11.5	9.5	14.5	3.9	5.5	30.5	07:08	06:30	05:00	08:30	03:18	00:29	20:00

Comparing the median maximum wake depth for the bubble wake and the ε wake, for the entire dataset, the ε wake was slightly deeper for all categories (~1 m) (Table 2, Fig. 7). The bubbles in the wake are an indication that surface water has been mixed down at depth and that it has been mixed with the ambient water. The bubbles will remain in the water column, or they can rise or collapse with time, depending on the bubble size. Bubbles with positive buoyancy will have an upward motion counteracting the downward mixing, which could be one explanation to why the bubble wakes are slightly shallower than the ε wakes. The dissipation rate of turbulent kinetic energy on the other hand, is a measure of the turbulent motions in the water that mixes the water down. When the turbulence decays, the dissipation also decays and dies out. As the bubbles may remain after the turbulence has died out, it can explain why the bubble wake lasts longer compared to the ε wake. Another possible explanation to why the ε wakes are deeper, is the calculation method used. The dissipation estimate is influenced by neighbouring cells (Eq. 1), and, if there is strong turbulence in one cell and none in the next, the method may still show some turbulence in the calm cell.

Commented [AN34]: **Comment 18** (R2, specific comment): Line 355-359: Why do the bubble wakes look different from turbulence? Please discuss further.

Commented [AN35R34]: Changes:
-The discussion regarding the differences between the bubble a[...] turbulent wake has been expanded.

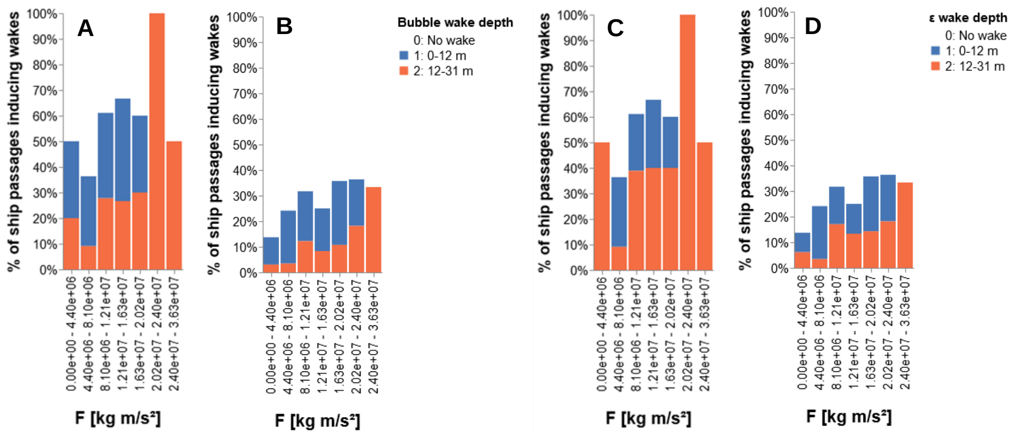


Figure 7. Maximum wake depth for the bubble wake (a, b) and dissipation rate of turbulent kinetic energy (ε) wake (c, d). (a) and (c) are show the close wakes, induced by ships passing at 0–3 ship widths from the instrument, and (b) and (d) show all wakes induced by single ship passages. The x-axis shows the force (F) of the vessel in Newton. Wake depths within the range presented in previous studies are shown in blue and wakes deeper than previously reported are shown in orange.

There is no statistically significant correlation between passage distance from the instrument and wake depth, per category or overall. A possible explanation to this lack of correlation could be the skewed data distribution, as there were few passages within 25 m of the instrument. The lack of correlation could also be an indication that the maximum wake depth depends on more variables than just proximity and that further studies are needed to resolve what influences the development of the turbulent wake. Nevertheless, in general the deepest wakes were caused by ships passing closer to the instrument, whereas ships passing at larger distances from the instrument (100–199 m) mainly caused shallower wakes (Fig. 5). Yet, the deepest wakes were caused by ships passing 25–75 m away, which demonstrates that even at distances up to 75 m from the ship, mixing down to 20–30 m depth can be induced (Fig. 5). Even though there is a lack of significant correlation, there are strong indications (Fig. 5) for a distance dependent detection of maximum wake depths. This in turn indicates that the median wake depths presented in this study could be an underestimation, as it includes wakes from all distance between 0 to 184 m.

Commented [AN36]: **Comment 11** (R1, specific comment 7 The paper is full of obvious, trivial statements, e.g. "in general th deepest wakes were caused by ships passing closer to the instrum whereas ships passing at larger distances from the instrument (10 199 m) mainly caused shallower wakes : : :" (lines 369-370)

Commented [AN37R36]: Change:
-With the revised result section, this sentence is removed.

3.1.5 Temporal wake longevity

Figure 8 shows the wake occurrence and wake temporal longevity related to vessel force, for the same ship typewake categories and parameters as in Figure 7, for the bubble wakes and ε wakes. The median longevity for all wakes was 08:44 min (std 06:29) and 06:30 min (std 03:18) for the bubble and ε epsilon wake respectively.) (Table 2). The close wake category had the same longevity for the ε wake, but the bubble wake was longer at 09:59 min (std 06:34 min). Figure 8 shows no clear correlation between wake longevity and vessel force, for the bubble or ε wake. Hence, based on the results from this study, it seems like parameters related to the vessel speed and size do not explain the variation in wake longevity to a very high degree. However, the relatively low number of passages with a large vessel force makes it difficult to draw any definite conclusions without further studies.

In similarity with the maximum wake depth, the double category had a longer duration on average, compared to the single categories, for both the bubble and ε wakes (Table 2). A majority of the longest wakes (20–30 min) were induced by ships passing within 50 m3 ship widths of the instrument (Fig. 8). If As this indicates that proximity plays a role in the ability to detect the entire wake temporal longevity of the wake, the close wake category median would be a better estimate of wake longevity, compared to the mean median longevitiesy presented calculated from all detected wakesin this study would be underestimat.ed, as wakes from all distances are included in the mean calculation. Compared with previous studies, Aa detectable signal of the bubble wake from 10 and up to 30 min, is in agreement (Table 1). with the results from previous studies (Furthermore, the timescale of the wake longevity indicates that in highly trafficked areas, where large ships passes every 10–15 min, there is a high potential of a constant influence of ship-induced vertical mixing

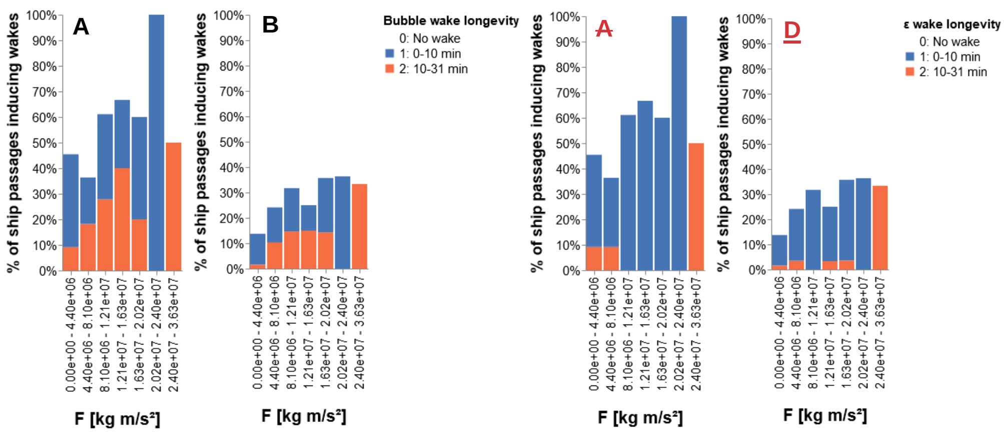
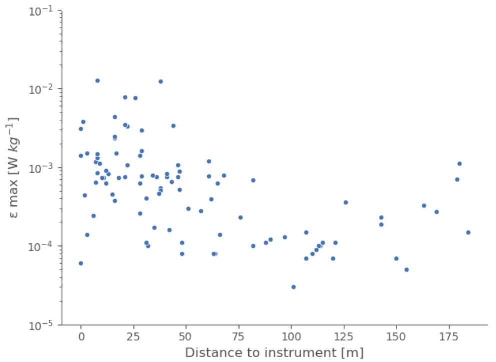


Figure 8. **Wake longevity for the bubble wake (a, b) and dissipation rate of turbulent kinetic energy (ε) wake (c, d). (a) and (c) are show the close wakes, induced by ships passing at 0–3 ship widths from the instrument, and (b) and (d) show all wakes induced by single ship passages. The x-axis shows the force (F) of the vessel in Newton. Wake temporal longevities < 10 min are shown in blue and wake longevities 10–31 min are shown in orange.**

hows the relation between the maximum turbulent kinetic energy dissipation rate (ε max) [W kg^{-1}], and the passing distance for the wake-inducing ships. All ε max intensity values are above the noise level of the measurements. The results show that the maximum dissipation rates are in the order of 10^{-4} to 10^{-2} W kg^{-1} in the core of the wake and decrease with distance from the wake core with a decay length scale of about 70 m. These values are comparable to what one would expect in breaking surface waves, and much larger than what is usually observed in the core of, or below, the surface mixed layer.



Maximum turbulent kinetic energy dissipation rate (ϵ) intensity [W kg^{-1}] in the wake region vs. the passing distance of the ship inducing the wake.

3.2 Bornholm satellite image analysis

There was a total of 94 satellite scenes from the period April 2013 to December 2018. Of these scenes, 25 % had a cloud cover of < 23 %, and were analysed for thermal wakes. 48 % of these (n=2311) had visible thermal wakes. The monthly distribution of ship passages and occurrence of thermal wakes are shown in Figure 9. As the number of analysed satellite scenes differed between months, the total number of ship passages for each month was divided by the number of analysed scenes. For all months, the majority of the passages did not induce visible thermal wakes. In April-July, there were several induced thermal wakes per scenes (Fig. 9), most of them in May and June. Occasional thermal wakes were found in September and October, but none were found during the winter months (December–February). In the satellite scenes where thermal wakes were visible, and the environmental conditions were right for thermal wakes to be visible, 21 % of the ship passages induced thermal wakes (Table 3). Looking at all the satellite scenes, including those without environmental conditions appropriate for inducing visible thermal wakes, 10 % of the ship passages induced thermal wakes.

Commented [AN38]: **Comment 11** (R1, specific comment 7 The paper is full of obvious, trivial statements, e.g. "the maximum dissipation rates : : : in the core of the wake : : :.are : : :..much lar than what is usually observed in the core of, or below, the surface mixed layer" (lines 403-405), etc. etc.

Commented [AN39R38]: We acknowledge and understand there are statements in the manuscript that can be perceived as tri and obvious, depending on the researcher's specialisation. To bal the content to suit a diverse audience, from different highly specialised disciplines, is a general challenge in interdisciplinary research. Therefore, the statement marked here was included for a reason. Our aim is to reach an interdisciplinary audience within o science, in line with the scope of this journal. This specific comm was included because the non-oceanographic co-authors of the pa explicitly asked for a comparison between our measured values a values that would occur naturally in the system. We believe that t types of statements fill an important function in making the conte of the paper more accessible to an interdisciplinary audience.

However, in the revised result section, this sentence was removed

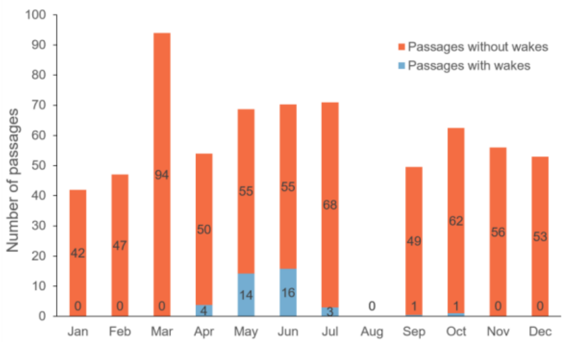


Figure 9. Seasonal distribution of ship passages for the satellite scenes with < 23 % cloud cover, for the period April 2013 to December 2018. The data labels in the stacked bar indicate the number of passages in each category. As some month has more than one analysed scene, the total number of ship passages for each month was divided by the number of analysed scenes, to get an average number of passages per scene for each month. August had no scenes with < 23 % cloud cover and therefore has no data.

~~. The main ship types inducing thermal wakes in the satellite dataset were *Cargo*, *Passenger* and *Tanker*, which all had > 40 passages and constituted 67 % of the total passages. Ship type categories with > 40 passages, but no thermal wakes were sailing (50 passages), pleasure (42 passages), and fishing (83 passages). All other ship types present in the dataset were combined within the *Other* passages category.~~

Commented [AN40]: **Comment 9** (R1, specific comment 5): Categorization of the ships in the context of their turbulent wakes does not look physically justified. More reasonable would be to r the wakes to the ship weight, draught, speed, possibly to the size/number of propellers.

Commented [AN41R40]: Change:
- The section related to the ship type categories has been remo

Table 3. Number of ship passages in the analysed satellite scenes and the percentage of passages inducing thermal wakes.

	Number of passages	% induced thermal wakes
Total passages	1430	10%
Total passages in scenes with thermal wakes	684	21%
Matched thermal wakes	144	
Unmatched thermal wakes	9	

3.2.1 Spatial wake longevity

The median length of the matched thermal wakes in the ship lane area was 13.7 km (std 11.8 km), and 25 % were ≥ 20.9 km (Fig. 10). Assuming that the median speed of the wake-inducing ships in the dataset (13.0 knots) is representative for the ship speed in the area, the calculated temporal wake longevity for the median wake length of 13.7 km was 34 min. The longest thermal wake was 62.5 km, which considering the speed of the wake-inducing ship (20 knots), corresponds to a longevity of 1 h 42 min. In model experiments by Voropayev et al. (2012), the thermal wake signature was still increasing at a distance of 30 ship lengths behind the ship, which would correspond to 6 km for a 200 m long ship. Thus, the thermal wake length reported in the current study, are up to one order of magnitude larger than previously reported experimental results, indicating an underestimation of thermal wake longevity in previous studies.

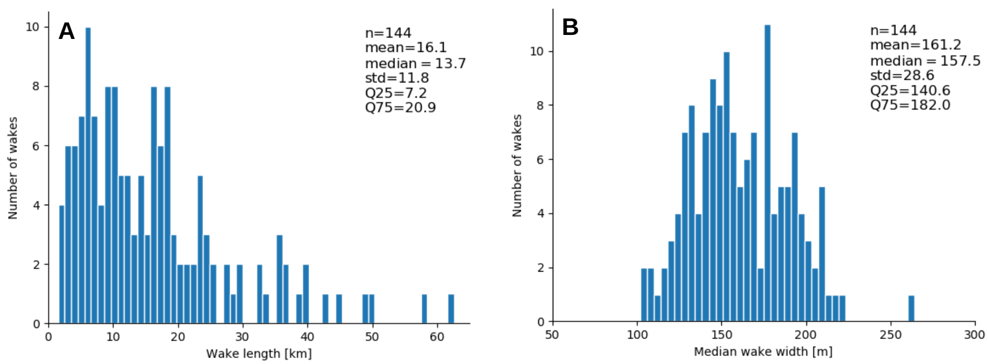


Figure 10. Wake length (left) and mean width of thermal wakes in the ship lane area, in satellite scenes with visible thermal wakes and < 23 % cloud cover, for the period April 2013 to December 2018 (n=144).

~~The thermal wake length occurrence for the main ship types is presented in Comparing the four categories *Cargo*, *Passenger*, *Tanker*, and *Other*, the first three have similar percentage of induced wakes (26 39 %), whereas the *other* category has a very low proportion of induced wakes (8.5 %).~~

~~**Number of ship passages and percentage induced thermal wakes per ship type.**~~

	~~Cargo~~	~~Other~~	~~Passenger~~	~~Tanker~~
~~Total passages~~	~~347~~	~~48~~	~~50~~	~~96~~
~~% wakes~~	~~26 %~~	~~8.5 %~~	~~39 %~~	~~39 %~~

3.2.2 Spatial wake width

The thermal wake width distribution is presented in Figure 10 and Figure 11. The median wake width for the entire dataset was 157.5 m (std 28.6), which is within the range 10–250 m range presented in previous studies (Table 1). The width in this study corresponds to the values presented in Gilman et al. (2011), who also used a ship-based remote sensing approach to estimate width from the visible wake on the sea surface. In contrast, Trevorrow et al. (1994) and Ermakov and Kapustin (2010) reported typical widths of 40–80 m, which is narrower than any widths detected in the current study. However, the last two studies used acoustic measurements of bubbles to estimate the wake width, which could explain the diverging results. The distribution of the median wake width for the different satellite scenes can be seen in Figure 11. Variations in stratification conditions~~,~~ could ~~also~~ be one of the explanations to why the thermal wake width varied between scenes. Another reason ~~Other environmental conditions,~~could be local and regional ~~such as~~ wind conditions ~~, can also affect the visible surface wake, as shown~~as pointed out ~~-~~in Gilman et al. (2011), or simply the varying temperature gradient between entrained cooler temperatures and warmer temperatures of the upper layer and the resulting exponential adaption process given Newton's law of cooling ~~(~~(Mallast and Siebert, 2019; Vollmer, 2009).

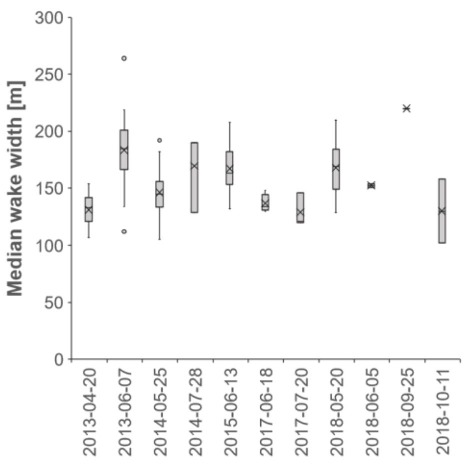


Figure 11. Median wake width distribution for the thermal wakes in the 11 satellite scenes with visible thermal wakes and < 23 % cloud cover, for the period April 2013 to December 2018. The median values are indicated with an X and outliers with rings (o).

3.3 Implications ~~for~~ of the spatial and temporal scales of ~~the~~ turbulent wakes

The environmental implications of the spatiotemporal scales of the turbulent wake presented in this study~~,~~ can be illustrated by an example. Using the longevity and width of the "median" wake turbulent wake, it is possible to estimate the area of the ship lane being affected by the turbulent wake at any given time. The traffic separated ship lane in the sound north of Bornholm is intensely trafficked, with 50, ~~-~~000 ship passages every year (HELCOM, 2010). A typical example of the number of ships present in the area at any given time can be seen in Figure 12, which shows the ships with AIS transmitters present in the sound at the time of the satellite scene from 2014-05-25 10:01. The ship lane area and the traffic separation zone are indicated, with each ship lane being 5 km wide and approximately 30 km long, which correspond to an area of 150 km^2. During the time of the satellite passage, there were four ships present in both the south-east ship lane in the traffic separated part in the Bornholm sound, and in the north-west ship lane. The median thermal wake length (13.7 km) and width (157.5 m) (Fig. 9) gives an average thermal wake area of 2.16 km^2. Consider a scenario where all wakes are uniformly distributed without overlap. With no overlapping wakes, four median ship wakes would cover an area of 8.6 km^2. With a ship lane area of 150 km^2, the area covered by thermal wakes would correspond to 5.8 % of the ship lane. It is not unusual that eight ships are present in the ship lane at the same time, in which case the covered area would be 17.3 km^2, which corresponds to 11.5 %. In Figure 12, some of the wake tracks are overlapping, thus it does not fully correspond to the scenario of uniformly distributed wakes. However, the figure still gives a conceptual visualisation of how large the part of the ship lane area is that can be influenced by thermal wakes.

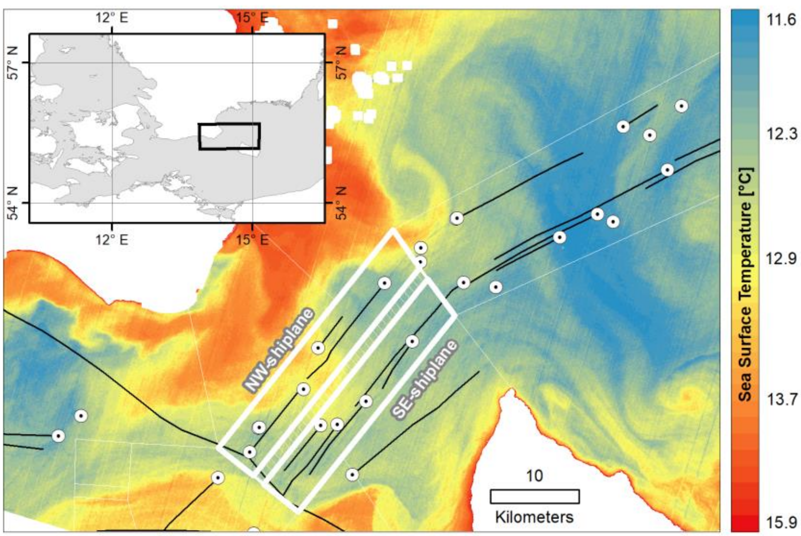


Figure 12. Ships and visible thermal wakes in the Bornholm sound from the analysed satellite scene from 2014-05-25 10:01. White circles with black dots indicate ships with an AIS transmitter present in the area at the time of the satellite passage, and the black lines are the digitalised visible thermal wakes from the satellite scene. White lines indicate the ship lane area, the bold lines mark the North-West and South-East ship lanes, and the dashed area is the traffic separation zone. Landsat-8 image courtesy of the U.S. Geological Survey.

In addition to an estimate of the area affected by the turbulent wake, it is also possible to consider the frequency at which the water mass in a certain point would be influenced by a turbulent wake. An average of 50,000 ship passages in the Bornholm sound, corresponds to 25,000 passages in each direction, which divided over a year would correspond to approximately one passage every 21 min (~ 3 per hour). Consider a scenario, where instead of a uniform distribution of ships in the entire ship lane, all ships travel along the exact same path. The calculated median temporal thermal wake longevity for the satellite data was 34:00 min. As the thermal wake longevity is longer than the average time between ship passages, the assumption that all ships travel the exact same route would mean that the water mass along the travelled route would be under constant influence of a ship-induced thermal wake. Now consider the same scenario, but using the median temporal longevity for all the ADCP wake measurements from the close wake passages, 09~~8~~:59~~44~~ min for the bubble wake and 05~~6~~:59~~30~~ min for the ε wake (Table 2). As the bubbles in the turbulent wake are visible for 09~~8~~:59~~44~~ min, the assumption that there is a ship passage every 21 min means that there ~~are~~ is 1~~2 13~~ 1min between each ship passages when there are no bubbles. If using the median temporal longevity for the ε wake instead, the time would be 15~~4.5~~ min. Hence, the bubble wake would influence the water mass in a certain point every ~~12 13~~11 min and the ε wake every 1~~4.~~5 min.

The difference in temporal longevity, between the ADCP measurements and satellite observations, can partly be explained by the fact that the two methods measures different aspects of the turbulent wake. The ADCP measurements show the very turbulent core of the wake. The dissipation rate of turbulent kinetic energy (ε) gives an estimate of the intensity of the mixing, and both the ε and bubble wake gives an estimate of the spatial scales of the turbulent wake. The satellite observations, on the other hand, show the mixed water that has been produced by the turbulent mixing. The mixed water from the turbulent wake will remain even after the turbulence has died away, but it is still a measure of water that has been influenced by mixing. Hence, both methods can be used to estimate the spatial and temporal scales of ship-induced mixing, but the ADCP measurements give an estimate of the turbulent wake, and the satellite analysis shows the scales of the water influenced by the turbulent wake.

The above calculated area coverage of thermal wakes, and the frequency at which the water mass in a certain point would be influenced by ship-induced mixing, represents two extremes. The first scenario assumes a uniform distribution of all ship wakes, and the second scenario assumes that all ships travel along the same route. However, in reality some of the wake regions would be overlapping, and most ships would travel similar, but slightly different routes in the ship lane. Nevertheless, based on the results presented in this study, areas like the Bornholm ship lane in the Baltic Sea could be considered under a near constant influence from ship-induced turbulent mixing. ~~Hence, the results of this study indicate that in areas with highly trafficked ship lanes, the local mixing dynamics can be affected by ship-induced turbulent mixing.~~ Even if the water column regains its stratification quite quickly, the mixing of the wake water with the surrounding water would take much longer. In a natural marine system, the water column is often stratified due to surface heating and/or freshwater influence. The wake turbulence interacts with this stratification by mixing the water and entraining deeper waters into the wake. The stratification may, in turn, reduce the vertical extent of the wake relative to what it would have been in a homogeneous water column (e.g. Voropayev et al. (2012)). During periods of seasonal stratification, nutrients in the surface layer are depleted, and the supply of nutrients from below is limited due to damping of the vertical mixing by the stratification (Reissmann et al., 2009; Snoeijs-Leijonmalm and Andrén, 2017). In coastal regions, nutrients can be brought up to the upper mixed layer by coastal upwelling, but in the open ocean, the nutrient supply is dependent on vertical mixing (Reissmann et al., 2009). If the vertical mixing is intense and deep enough, it will bring up nutrient rich water from below the stratification to the upper surface layer, which can increase primary production and sustain algal blooms. In ocean systems unaffected by human activities, vertical mixing in the surface layer is induced by wind, and the depth of the mixing depends on the wind strength and duration, as well as the input of buoyancy from heating and fresh water (Thorpe, 2007). In temperate oceans like the Baltic Sea, the seasonal stratification occurs during the summer season, which is also the period with the least wind (Reissmann et al., 2009). Thus, in unaffected seasonally stratified waters, there is little vertical mixing during the summer months. However, in areas with intense ship traffic there is an input of ship-induced vertical mixing. Consequently, if the depth of the ship's vertical mixing is similar to the stratification depth, intense ship traffic has the potential to regionally affect natural biogeochemical cycles. Consequently, during summer stratification, ship-induced turbulent mixing has the potential to alter gas exchange and nutrient availability on a local/regional scale, which should be considered when evaluating environmental impact from shipping.

The results presented in this study, also have implications for monitoring and data collection in areas with ship traffic. An especially interesting example are the so called FerryBox systems, which are placed on ships and ~~do~~ conduct continuous~~es~~ measurements of parameters such as O_2 concentration, salinity, temperature, and sometimes also pCO_2, Chlorophyll a, and pigments (Petersen, 2014). There are currently seven passenger ferries equipped with FerryBox systems in the Baltic Sea, which are traveling along the major shipping lanes all or part of the journey (https://www.ferrybox.com/routes_data/routes/baltic_sea/index.php.en). The intake of water is from an inlet in the ship hull (Petersen, 2014), which would correspond to somewhere between 2–10 m depth. Considering the wake longevity of the thermal and turbulent wakes presented in this study, there is a high likelihood that a ship traveling in a major ship lane, could be in the wake of another ship. In that case, the water being measured by the FerryBox is the water of the turbulent wake, and thus not representative for the conditions outside the ship lane. The validations made for FerryBox measurements are being made using the same water source as the FerryBox (Karlson et al., 2016), which would still be part of the ship lane area, and not the unaffected waters outside the ship lane. As the measured temperature differences between inside and outside the thermal wakes, was up to 1°C in some of the scenes (see Fig. 4 for example), and as the bubbly wake affects gas exchange and saturation, it is important to know if the measurements are affected by ship-induced turbulence. Hence, the effect of ship-induced vertical mixing should be considered when using data collected from FerryBox systems.

Among the ADCP measurements, there were a few wakes which reach depths of >18 m (Table 2). The deepest wake in this dataset was induced by ~~the~~ a cargo ship ~~Ficaria Seaways, a ship~~ with a beam of 25 m, length of 229 m, and draught of 7 m. It passed the instrument at a distance of 34 m and~~at~~ a speed of 19 knots. ~~Ficaria Seaways~~The cargo ship ~~has~~ had a Gross Tonnage similar to the average of container and Ro-Ro cargo ships in the Baltic Sea (HELCOM, 2018), indicating that very deep ship-induced mixing could be a common, but undetected occurrence. The hypothesis that deep vertical mixing could be more frequent than expected from previous studies is supported by the fact that similarly sized ships passing at the same distance as ~~Ficaria Seaways~~cargo ship inducing the deepest wake, also induce mixing to depths greater than 15 m. The lack of previous reports of deep vertical mixing of this magnitude can partly be explained by the fact that no previous study has targeted this specific research question. Moreover, measurements made using similar methods, but for other purposes, are seldom conducted in ship lanes and particularly not from below. On the other hand, the difference in wake depth for ships of similar size and passing distance could also be due to differences in stratification, as a strong stratification can dampen the vertical development of the wake (Kato and Phillips, 1969). During the ADCP measurement campaign, water column stratification was measured at deployment and retrieval of the instrument (Fig. 4). Three hours before the instrument retrieval, a cargo ship ~~the ship Magnolia Seaways~~ passed at a distance of 21 m and induced a bubble wake of 13.5 m depth and a ε wake 17.5 m depth. At the point of retrieval, the CTD measurement showed a ~~there was a~~ strong thermal stratification at 5 m depth. At the time of ship passage, intense vertical mixing induced the wake down to 17.5 m depth, and it is unlikely that the thermal stratification at 5 m depth was present during the longevity of the wake. ~~, but the turbulent wake still reached below 15 m depth.~~ This means that the stratification was influenced by the wake and that waters above and below the thermocline were mixed with each other. Three hours later the water had re-stratified and the mixed water has spread out laterally. However, during the longevity of the wake, the stratification was most likely strongly affected, and the mixing between water masses is irreversible even though it is difficult to observe after the re-stratification. It therefore affects the local and regional stratification, even though the contribution from each single ship is difficult to observe after the water has re-stratified. ~~As Magnolia Seaways has a draught of 7 m, parts of the hull itself would have been below the thermocline, which could explain the deep wake. Still, this example shows that deep vertical mixing is possible across a strong thermocline.~~ Nevertheless, further studies are needed to determine the impact of stratification on the vertical development of the turbulent wake, and how it varies with the ship's draught and speed. Thus, the results from this study shows that very deep vertical mixing occurs, and possibly at a high frequency. However, as the current knowledge about the wake distribution is poor (especially on a vertical scale), and further studies are needed to determine when, and at which frequency, deep vertical mixing occurs.

3.4 Limitations and Future outlook

The measurements in this study indicated resuspension and turbulence at the sea floor at 30 m depths, induced by the ~~wave~~ wake from passing ships. These effects were seen at quite large distances from the passing point, indicating the importance of including the effect of the wave wake when estimating the environmental impact on the marine environment in intensely trafficked ship lanes (if the water is not too deep). However, this effect was outside the scope of the current study and has been investigated by Soomere et al. (2009).

The lack of detectable thermal wakes in the satellite dataset during the winter months was expected. A thermal stratification is needed to entrain coole~~d~~r water from below, induced by the turbulent wake and cause a surface temperature gradient. The Bornholm region usually has a no thermal stratification during winter (Reissmann et al., 2009; van der Lee and Umlauf, 2011). Therefore, the method of estimating the spatiotemporal scales of the turbulent wake using satellite SST measurements~~,~~ is limited to seasons and regions where strong thermal stratifications occur. Moreover, the low percentage of

Commented [AN42]: **Comment 19** (R2, specific comment): Line 550-552: This seems to be a negative result: the stratificatio[n] was not affected by the wake. Remove this part. I suggest that the authors measure more vertical profiles in the study area and/or provide literature for more data to characterize general and unusu[al] environmental conditions. 4 casts x 2 days are not enough.

Commented [AN43R42]: As mentioned in a previous comm[ent,] we agree and acknowledge that the amount of vertical profiles is [too] low to have an in-depth discussion regarding the effects of ship-induced vertical mixing on stratification. However, for the aim of [this] paper, we consider the current amount of profiles enough to descr[ibe] the spatiotemporal scales of the turbulent wake.

However, regarding the comment related to the sentence above, w[e] humbly disagree that it is a negative result. Turbulence will not be able to reach below the mixed layer, to 17.5 m depth, without mix[ing] the water. Thus, we find it highly unlikely that the thermal stratification at 5 m was present within the wake. This means that [the] stratification was influenced by the wake and that waters above a[nd] below the thermocline were mixed with each other. However, thre[e] hours later the water has re-stratified and the mixed water has spr[ead] out laterally. We do not claim that there was a long-lasting effect [on] the stratification. However, during the longevity of the wake the stratification was most likely affected. Hence, in a scenario with v[ery] frequent ship passages, there will be less time for the re-stratifica[tion] to occur, and a more long-lasting effect on the stratification could [be] possible.

For the comment regarding measuring more vertical profiles in th[e] study area, see **Comment 15.**

Commented [AN44R42]: Change:
- The example of the ship passage affecting the thermal stratification has been clarified.

available satellite scenes with little enough cloud cover~~,~~ makes alternative remote sensing techniques, such as drones, a possible ~~better~~ alternative. Drones could also be used for a longer time period in the same area and in combination with under water measurements.

When comparing the observations from the satellite data and the ADCP measurements, it is important to remember that they were obtained in different ocean basins and during different stratification conditions. As stated in the methods section, the sites were chosen based on the needs for the two types of measurements. The frequent ship traffic and a separation zone in the Bornholm ship lane~~,~~ made it suitable for detecting the longevity and occurrence of thermal wakes. However, the intense maritime activity in the area and the larger depth, made it both riskier and logistically difficult to place ADCP instruments in the Bornholm ship lane. The authors have previously lost two instruments in the Bornholm area, due to maritime activity, which is why it was considered unsuitable for a longer measurement period. Instead, the Gothenburg area was chosen, as it has a varied and intense ship traffic, with several ferries and cargo vessels on route, ensuring daily passages. It also gave the possibility to access detailed draught information from the ships from the Port of Gothenburg, for the ships visiting the port. The Gothenburg site was not considered suitable for the satellite study, due to the lower amount of ship passages per day and cloudier weather conditions. These circumstances resulted in the decision to use different locations for the two studies. Nevertheless, satellite images were retrieved for the Gothenburg site for the *in situ* measurement period, but they were too cloudy to be usable for any analysis. For future studies focused on characterising the development of the turbulent wake, the ideal would be to make remote sensing and ADCP measurements simultaneously at the same site. However, it would probably be more suitable to use drones instead of satellite images, as a drone is more flexible and makes it possible to operate during cloudy conditions, to capture the development in time, and to use both static and dynamic approaches when documenting the wake.

In addition to the difference in geographical area, the satellite observations show a snapshot of the ocean surface, whereas the ADCP instrument does not measure the top 4 m of the water column. Hence, the two methods never capture the same part of the wake, which could lead to different results using the two methods. Moreover, the satellite observations show the effect of mixing, while ADCP observations show the actual turbulence that causes the mixing. After the mixing has occurred, the mixed water may move outwards – a movement not causing enough turbulence to be seen by the ADCP. This could be one explanation to why the thermal wake longevity is longer, compared to the ADCP wake longevity.

The satellite analysis showed a median wake width of 157.5 m (Fig. 10), from which it would be expected to frequently detect wakes from ships passing up to 75 m from the instrument. The ADCP results indicate a similar range for frequent detection, namely 50 m~~. Further, the decay scale for maximum dissipation rates (Fig. 7) is about 70 m, which also is in accordance with a wake width in the order of 150 m~~. When considering the detection range of the ADCP instrument, it is important to consider the influence of currents and wind. In this study, the water speed and waves were measured with the

Commented [AN45]: **Comment 14** (R2, general comment): It is not clear why the authors chose two separate locations for the situ and remote sensing study. Is it possible to find satellite image for the in-situ measurement period?

Commented [AN46R45]: Changes:
-A section has been added here motivating the choice of the tw different sites.

Commented [AN47]: **Comment 20** (R2, specific comment): Line 574-581: As mentioned above, please try to find satellite imagery that covers the in-situ measurement area.

Commented [AN48R47]: See answer to **Comment 14**.

Commented [AN49R47]: Changes:
-A section has been added in the discussion and method, motivating the choice of the two different sites.
-A sentence has been added in the discussion suggesting future improvements of the method.

Commented [AN50]: **Comment 7** (R1, specific comment 3) Why the wakes appeared in the ADCP records for ships passed b some distances from the ADCP? Because of the wake turbulent diffusion? If so, why not to analyze, e.g. the characteristic times o the turbulent diffusion, the diffusion spatial/temporal decay, etc.?

(This comment is also addressed in the methods section)

Commented [AN51R50]: We consider two main reasons fo wakes appearing in the ADCP records for ships passing at some distance from the wake. Firstly, currents can move the wake towa the instrument. However, when looking at the impact of currents our data, we could not see that that current speed and/or direction affecting the existence of a wake in our data. Moreover, during m of the measurement period the current speed at the instrument position was very low. These two circumstances indicate that wa drift due to currents is probably not the main reason that wakes a detected from ships passing at long distances from the instrument The second possible reason, as suggested, is that turbulent diffus widens the wake. As shown in the satellite data, the median therm wake width was 157 m. Even though the thermal wake and the turbulent wake are not exactly the same, it still indicates that a sh passing at a distance of up to 150–200 m away from the instrume could induce a wake, where the edge of the wake eventually reac the instrument. In addition, the wake will be widened by buoyanc effects if the water is stratified. We have indications from recent measurements that the turbulent wake can diffuse/spreads sidewa along the pycnocline and is also to some extent "retained" at that depth.
There is also a reason why we have chosen not to analyse the characteristic times of the turbulent diffusion. As the ADCP only measure in one point, it is not clear if the entire length of the wak captured or which part of the wake is being measured. Since it is sure that the entire wake from start to end is measured, making calculations on diffusion rate based on it would not be fully representative. We intend to perform these more detailed calculat

Commented [AN52]: **Comment 21** (R2, specific comment): Line 587: Note that winds are also important.

Commented [AN53R52]: Change:
- Wind is now also mentioned. As we have measured the wate speed and waves, we consider the wind effect to be captured b those measurements.

ADCP, and the wind effect on currents and waves were considered captured by those measurements. As a current can move the wake towards or away from the instrument, the current speed and direction must be taken into consideration when estimating at what distance from the ship a wake is likely to be detected. Trevorrow et al (1994) conducted measurements within 2–5 m of the turbulent wake and reported difficulties in catching the bubble signal from the wake using vertical sonars, as the wake often drifted out of the sonar range before it had completely dissipated. In the current study, In this study, aa majority of the passages (50–60 %) occurred when there was a weak or no current at the position of the ADCP instrument (data not shown). Moreover, a current speed towards the instrument did not increase the likelihood of detecting the wake, especially not when ships passed further away from the instrument (data not shown).

In addition to the currents, the width and size of the ship should also be taken into consideration when discussing detection related to the passage distance from the instrument. The distance between the ADCP instrument and ship, is calculated from the position of the AIS transmitter. As the transmitter is often located at the middle of the ship and, a wide ship might be passing right over the instrument even though the AIS stamp indicates that it is 25 m away. Thus, larger ships are likely closer to or further away from thepossibly closer to the instrument than what is registered by the AIS, which could potentially influence the wake detection. To adjust for this bias, the graphical presentation of the data has the distance to the instrument has presented in ship widths instead of meters. A large majority of the ships inducing wakes in the ADCP measurements were 20 m or wider, and the wider ships were overrepresented among the passages inducing wakes, comparing the wake width for the entire dataset. Moreover, the smallest ships (width < 10 m) rarely induced wakes, and then only when passing within 75 m of the ADCP. A similar pattern can also be seen when looking at the length of the ships inducing the wakes. For all ship type categories except *Passenger*, the longer ships occur more frequently among the ships inducing detectable wakes, compared to the ship length in the entire dataset.

In the current study, the water column stratification was only measured at deployment and retrieval of the instrument, hence the importance of stratification could not be included in the analysis of this study. However, the presence and strength of the stratification will influence how much turbulence that is required to mix water and substances across the thermocline (e.g. Kato and Phillips (1969)). In a stratified fluid, vertical mixing removes energy from the turbulence, reducing the vertical extent of the wake development. Stratification will also cause mixed fluid to spread out laterally, which causes an adjustment of the wake stratification to the surrounding stratification, resulting in a widening of the wake as well as an additional limitation of the vertical extent (Voropayev et al., 2012). As the aim of the current study was to present an order of magnitude estimation of the spatial and temporal scales of the turbulent wake, the lack of stratification measurements does not present at problem within the current scope. However, for future studies with the aim of characterising the development of the turbulent wake and quantifying the ship-induced vertical mixing, stratification measurements will be necessary in order to understand the interaction between the stratification and the turbulent wake. Moreover, as the stratification must be

Commented [AN54]: **Comment 15** (R2, general comment): Also, vertical profiles should have been measured more frequentl see the effect of wakes on stratification and mixing.

Commented [AN55R54]: We agree and acknowledge this a main potential improvement of our observations, which we have discussed in this section of the manuscript. Furthermore, the long term aim of our research is to be able to study and discuss the eff of wakes on stratification and mixing. However, the aim of the current study was to describe the spatiotemporal scales of the turbulent wake, and not resolving all the parameters determining effect of wakes on stratification and mixing. As mention in a prev answer, we realise that the current title and aim, could give the impression that the paper is of a more explanatory nature. Howev we consider the current amount of vertical profiles enough for the aim of this paper and suggest leaving the discussion regarding the effect of wakes on stratification and mixing for the next paper (w high-resolution profiling will be included).

Commented [AN56R54]: To address this comment, we will clarify the aim of the paper as being mostly descriptive.

expected to be an important factor for wake depth, it could be one explanation for the absence of statistically significant correlations between wake depth and vessel force ~~other parameters~~.

> **Commented [AN57]:** **Comment 22** (R2, specific comment): Line 618: What parameters?
>
> **Commented [AN58R57]:** Change:
> - The parameters have been specified.

In order to determine when deep vertical mixing occurs, and how common it is, future studies need to simultaneously measure the wake in more than one point, in order to get the cross section of the wake. One way of achieving this would be to conduct measurements with several ADCPs placed on a row perpendicular to the ship lane. This would give a cross-section of the wake, which could be used to describe both the width and depth of the turbulent wake. As the measurements in this study were made using one instrument, only the depth of the wake could be measured, and only at one point in the wake cross-section. Moreover, a line of instruments would also be able to capture a drifting wake and thus better estimate the true longevity. One of the limitations of the longevity estimation in this study, is that currents could potentially shift the wake away from the instrument. Using multiple instruments would increase the chance of capturing the entire wake development, as it would cover a larger area, thus increasing the reliability of the longevity estimation. As the results from this study indicate that proximity is of importance for detecting the turbulent wakes using ADCP measurements, multiple instruments would increase the area where ships can pass close to the instrument. In addition, if the maximum depth of the wake is located only in a certain region of the turbulent wake, the likelihood of measuring that part of the wake is small when only one instrument is used. This spatial limitation of the current study makes it difficult to determine if the small number of detected deep wakes was because of low occurrence, or because using only one instrument made it difficult to successfully measure the deepest part of the wake. Thus, multiple instruments would increase the ability to identify when and where the very deep mixing occurs and shed further light upon how frequently deep mixing is induced. It would also be beneficial to conduct concurrent measurements using ADCPs and remote sensing. In the current study, the satellite analysis and ADCP measurements have been conducted at different locations and time periods, but concurrent measurements would give a more complete picture of both the large horizontal temporal and spatial scales, as well as the vertical scales.

4 Conclusions

Based on a large sample of *in situ* measurements, the~~This study has~~ median spatiotemporal extent of turbulent ship wakes have been estimated to a depth of 13.5 m and longevity of 09:59 min, based on ADCP measurements. Thermal wake width and longevity have been estimated to a median of 157.5 m and 13.7 km respectively, based on SST satellite image analysis. The results show frequent detection of turbulent wakes deeper than 12 m, which is deeper than previously reported. Moreover, in areas with intense ship traffic, the presented ~~shown that ship-induced vertical mixing occurs at~~ temporal and spatial ~~scales~~ extent of the turbulent wakes are of a scale ~~that make this process important~~relevant to consider ~~in areas with intense ship traffic~~when assessing environmental impact from shipping.~~. Moreover, the possibility that very deep vertical mixing could be transpiring frequently, highlights the need of further studies to better characterise the spatial and temporal development of the turbulent wake, and the interaction between wake and stratification.~~

5 Data Availability

Acoustic measurement data available upon request for non-commercial purposes. AIS data available through HELCOM according to their data policy. Satellite images freely available at https://s3-us-west-2.amazonaws.com.

6 Author contribution

I-M. Hassellöv, A. T. Nylund, L. Arneborg and A. Tengberg conceptualised and conducted the *in situ* field measurements and consecutive analysis and visualisation. A. T. Nylund developed the code used in the analysis, with contribution from L. Arneborg. U. Mallast conducted the data curation and formal analysis of the satellite images, with contribution from A. T. Nylund. The manuscript was prepared by A. T. Nylund with contributions from all co-authors.

7 Competing interests

The authors declare that they have no conflict of interest.

8 Acknowledgements

Acknowledgment of funding for the OCEANSensor project by the Research Council of Norway (project number 284628) and co-funding by the European Union 2020 Research and Innovation Program, as part of the MarTERA Program. Acknowledgement to the Swedish Institute for the Marine Environment (SIME), for supplying the AIS dataset~~, and to the Port of Gothenburg for supplying data of the current draught of all ships visiting the port~~. This work has been partially supported by MarineTraffic, by the use of their database of vessel information.

9 Abbreviations

ADCP - Acoustic Doppler Current Profiler

AIS - Automatic Information System

ANOVA - Analysis of Variance

EOF - Empirical Orthogonal Function

HELCOM - Baltic Marine Environment Protection Commission

MNDWI - modified normalized difference water index

NDRC - National Defense Research Committee

SIME - Swedish Institute for the Marine Environment

SMHI - Swedish Meteorological and Hydrological Institute

TIRS - Thermal Infrared Sensor

TOA - top-of-the-atmosphere

References

Balcombe, P., Brierley, J., Lewis, C., Skatvedt, L., Speirs, J., Hawkes, A., and Staffell, I.: How to decarbonise international shipping: Options for fuels, technologies and policies, Energ. Convers. Manage., 182, 72-88, https://doi.org/10.1016/j.enconman.2018.12.080, 2019.

Barsi, J. A., Barker, J. L., and Schott, J. R.: An Atmospheric Correction Parameter Calculator for a single thermal band earth-sensing instrument, IGARSS 2003. 2003 IEEE International Geoscience and Remote Sensing Symposium. Proceedings (IEEE Cat. No.03CH37477), Centre de Congress Pierre Bandis, Toulouse, France 21-25 July, 2003.

Carrica, P. M., Drew, D., Bonetto, F., and Lahey, R. T.: A polydisperse model for bubbly two-phase flow around a surface ship, Int. J. Multiphase Flow, 25, 257-305, https://doi.org/10.1016/S0301-9322(98)00047-0, 1999.

Emerson, S., and Bushinsky, S.: The role of bubbles during air-sea gas exchange, J. Geophys. Res.-Oceans, 121, 4360-4376, https://doi.org/10.1002/2016jc011744 2016.

Ermakov, S. A., and Kapustin, I. A.: Experimental study of turbulent-wake expansion from a surface ship, Izv. Atmos. Ocean. Phy+, 46, 524-529, https://doi.org/10.1134/S0001433810040110, 2010.

Foga, S., Scaramuzza, P. L., Guo, S., Zhu, Z., Dilley, R. D., Beckmann, T., Schmidt, G. L., Dwyer, J. L., Joseph Hughes, M., and Laue, B.: Cloud detection algorithm comparison and validation for operational Landsat data products, Remote Sens. Environ., 194, 379-390, https://doi.org/10.1016/j.rse.2017.03.026, 2017.

Francisco, F., Carpman, N., Dolguntseva, I., and Sundberg, J.: Use of Multibeam and Dual-Beam Sonar Systems to Observe Cavitating Flow Produced by Ferryboats: In a Marine Renewable Energy Perspective, J. Mar. Sci. Eng., 5, 30, https://doi.org/10.3390/jmse5030030, 2017.

Fu, H., and Wan, P.: Numerical simulation on ship bubbly wake, Journal of Marine Science and Application, 10, 413-418, https://doi.org/10.1007/s11804-011-1086-x, 2011.

Fujimura, A., Soloviev, A., Rhee, S. H., and Romeiser, R.: Coupled Model Simulation of Wind Stress Effect on Far Wakes of Ships in SAR Images, IEEE T. Geosci. Remote, 54, 2543-2551, https://doi.org/10.1109/TGRS.2015.2502940, 2016.

Gilman, M., Soloviev, A., and Graber, H.: Study of the far wake of a large ship, J. Atmos. Ocean. Tech., 28, 720-733, https://doi.org/10.1175/2010JTECHO791.1, 2011.

Heiselberg, H.: A direct and fast methodology for ship recognition in Sentinel-2 multispectral imagery, Remote Sens., 8, 1033, https://doi.org/10.3390/rs8121033, 2016.

HELCOM: Maritime Activities in the Baltic Sea—An Integrated Thematic Assessment on Maritime Activities and Response to Pollution at Sea in the Baltic Sea Region, Balt. Sea Environ. Proc. No. 123, Helsinki Commission (HELCOM), Helsinki, Finland, 65, 2010.

HELCOM: HELCOM Assessment on maritime activities in the Baltic Sea 2018. Baltic Sea Environment Proceedings No.152, Helsinki Commission, Helsinki, Finland, 253, 2018.

Issa, V., and Daya, Z. A.: Modeling the turbulent trailing ship wake in the infrared, Appl. Opt., 53, 4282-4296, https://doi.org/10.1364/AO.53.004282, 2014.

Karlson, B., Andersson, L., Kaitala, S., Kronsell, J., Mohlin, M., Seppälä, J., and Wranne, A. W.: A comparison of Ferrybox data vs. monitoring data from research vessels for near surface waters of the Baltic Sea and the Kattegat, J. Mar. Syst., 162, 98-111, doi: 10.1016/j.jmarsys.2016.05.002, 2016.

Kato, H., and Phillips, O.: On the penetration of a turbulent layer into stratified fluid, J. Fluid Mech., 37, 643-655, doi: 10.1017/S0022112069000784, 1969.

Katz, C. N., Chadwick, D. B., Rohr, J., Hyman, M., and Ondercin, D.: Field measurements and modeling of dilution in the wake of a US navy frigate, Mar. Pollut. Bull., 46, 991-1005, https://doi.org/10.1016/S0025-326X(03)00117-6, 2003.

Liefvendahl, M., and Wikström, N.: Modelling and simulation of surface ship wake signatures, Report FOI-R--4344--SE, Swedish Defence Research Agency (FOI), Stockholm, 36, 2016.

Loehr, L. C., Beegle-Krause, C.-J., George, K., McGee, C. D., Mearns, A. J., and Atkinson, M. J.: The significance of dilution in evaluating possible impacts of wastewater discharges from large cruise ships, Mar. Pollut. Bull., 52, 681-688, https://doi.org/10.1016/j.marpolbul.2005.10.021, 2006.

Lucas, N., Simpson, J., Rippeth, T., and Old, C.: Measuring turbulent dissipation using a tethered ADCP, J. Atmos. Ocean. Tech., 31, 1826-1837, https://doi.org/10.1175/JTECH-D-13-00198.1, 2014.

Mallast, U., and Siebert, C.: Combining continuous spatial and temporal scales for SGD investigations using UAV-based thermal infrared measurements, Hydrology & Earth System Sciences, 23, 2019.

Marmorino, G., and Trump, C.: Preliminary side-scan ADCP measurements across a ship's wake, J. Atmos. Ocean. Tech., 13, 507-513, https://doi.org/10.1175/1520-0426(1996)013<0507:PSSAMA>2.0.CO;2, 1996.

Moldanová, J., Fridell, E., Matthias, V., Hassellöv, I.-M., Eriksson, M., Jalkanen, J.-P., Tröeltzsch, J., Quante, M., Johansson, L., and Majutenko, I.: Information on completed BONUS SHEBA project, Baltic Marine Environment Protection Commission, Maritime Working Group, MARITIME 18-2018, Hamburg, Germany, 18, 2018.

NDRC: The Physics of Sound in the Sea, United States Office of Scientific Research and Development, National Defense Research Committee, Division 6, Washington, D.C., 1946.

Otsu, N.: A Threshold Selection Method from Gray-Level Histograms, IEEE T. Syst. Man Cyb., 9, 62-66, https://doi.org/10.1109/TSMC.1979.4310076, 1979.

Parmhed, O., and Svennberg, U.: Simulering av luftbubblor och ytvågor runt ytfartyg, Repot FOI-R-2217-SE, FOI - Swedish Defence Research Agency, Tumba, 2006.

Petersen, W.: FerryBox systems: State-of-the-art in Europe and future development, J. Mar. Syst., 140, 4-12, doi: 10.1016/j.jmarsys.2014.07.003, 2014.

Reissmann, J. H., Burchard, H., Feistel, R., Hagen, E., Lass, H. U., Mohrholz, V., Nausch, G., Umlauf, L., and Wieczorek, G.: Vertical mixing in the Baltic Sea and consequences for eutrophication–A review, Prog. Oceanogr., 82, 47-80, https://doi.org/10.1016/j.pocean.2007.10.004, 2009.

Smirnov, A., Celik, I., and Shi, S.: LES of bubble dynamics in wake flows, Comput. Fluids, 34, 351-373, https://doi.org/10.1016/j.compfluid.2004.05.004, 2005.

Snoeijs-Leijonmalm, P., and Andrén, E.: Biological oceanography of the Baltic Sea, Springer Science & Business Media, Dordrecht, Netherlands, 683 pp., 2017.

Soloviev, A., Gilman, M., Young, K., Brusch, S., and Lehner, S.: Sonar measurements in ship wakes simultaneous with TerraSAR-X overpasses, IEEE T. Geosci. Remote, 48, 841-851, https://doi.org/10.1109/TGRS.2009.2032053, 2010.

Soloviev, A., Maingot, C., Agor, M., Nash, L., and Dixon, K.: 3D sonar measurements in wakes of ships of opportunity, J. Atmos. Ocean. Tech., 29, 880-886, https://doi.org/10.1175/JTECH-D-11-00120.1, 2012.

Soomere, T., Parnell, K., and Didenkulova, I.: Implications of fast-ferry wakes for semi-sheltered beaches: a case study at Aegna Island, Baltic Sea, J. Coast. Res., 128-132, 2009.

Stanic, S., Caruthers, J. W., Goodman, R. R., Kennedy, E., and Brown, R. A.: Attenuation measurements across surface-ship wakes and computed bubble distributions and void fractions, IEEE J. Oceanic Eng., 34, 83-92, https://doi.org/10.1109/JOE.2008.2008411, 2009.

Swedish Maritime Administration, Nautical Information - Sjofartsverket: https://www.sjofartsverket.se/en/Maritime-services/Pilotage/Pilot-Areas/Goteborg-Pilot-Area/Harbour-Information-Goteborg/, access: 18 May 2020, 2020.

The Port of Gothenburg: https://www.portofgothenburg.com/FileDownload/?contentReferenceID=12900, access: May 18, 2020.

Thorpe, S. A.: An introduction to ocean turbulence, Cambridge University Press, New York, 240 pp., 2007.

Trevorrow, M. V., Vagle, S., and Farmer, D. M.: Acoustical measurements of microbubbles within ship wakes, J. Acoust. Soc. Am., 95, 1922-1930, https://doi.org/10.1121/1.408706, 1994.

UNCTAD: Review of Maritime Transport 2019, UNCTAD/RMT/2019, United Nations Publications, New York, 2019.

van der Lee, E. M., and Umlauf, L.: Internal wave mixing in the Baltic Sea: Near-inertial waves in the absence of tides, JGRC, 116, doi: 10.1029/2011jc007072, 2011.

Vollmer, M.: Newton's law of cooling revisited, EJPh, 30, 1063, 2009.

Voropayev, S., Nath, C., and Fernando, H.: Thermal surface signatures of ship propeller wakes in stratified waters, Phys. Fluids, 24, 116603, https://doi.org/10.1063/1.4767130, 2012.

Weber, T. C., Lyons, A. P., and Bradley, D. L.: An estimate of the gas transfer rate from oceanic bubbles derived from multibeam sonar observations of a ship wake, JGRC, 110, https://doi.org/10.1029/2004JC002666, 2005.

Xu, H.: Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery, Int. J. Remote Sens., 27, 3025-3033, https://doi.org/10.1080/01431160600589179, 2006.