Tuning Detection of Ship Wakes by Detectability Modelling

Björn Tings
SAR signal processing
German Aerospace Center (DLR)
Bremen, Germany
0000-0002-1945-6433

Yi-Jie Yang
SAR signal processing
German Aerospace Center (DLR)
Bremen, Germany

Christoph Schnupfhagn
SAR signal processing
German Aerospace Center (DLR)
Bremen, Germany

Sven Jacobsen
SAR signal processing
German Aerospace Center (DLR)
Bremen, Germany

***Abstract*— The appearance and detectability of ship wakes in SAR imagery has been investigated for decades. Various wake components have been identified in the complex appearance of wake signatures. Wake signatures in SAR imagery can also be exploited for indirect ship detection if ship signatures are weak or absent. A systematic investigation on dependencies between wake detectability and physical parameters affecting the detectability of wake components, i.e. influencing parameters, has recently been published. Recently also machine learning has been applied for wake detection, making the wake detection performance more robust. This study presents a novel approach to increase detection accuracy during post-processing. A combination of the systematic wake detectability modelling with a robust wake detector based on machine learning is proposed. The results show that the dynamic filtering on the basis of wake detectability models can increase the precision while maintaining the recall.**

Keywords—Synthetic Aperture Radar (SAR), Machine Learning, Object Detection, Object Detectability, Ship Wakes

I. Introduction

Synthetic Aperture Radar (SAR) sensors mounted onboard satellite platforms offer the opportunity to observe the maritime domain, even during nighttime and under foggy or cloudy weather conditions. Further, maritime traffic can be tracked externally without requiring cooperation from the traffic under observation. However, SAR acquisitions form one SAR sensor are roughly available once per day in an area of interest. In order to increase the revisit time of an area, for maritime situation awareness typically an ensemble of SAR sensors and earth observation (EO) missions is used and combined with data from Automatic Identification System (AIS [1]) [2]. Thus, SAR sensors are an acknowledged choice for supporting the pursuit of illegal maritime activities using an operational observation environment [2, 3].

SAR acquisitions can cover hundreds of square kilometers of the sea surface. An automatic analysis of those acquisitions is necessary, as a manual analysis is time consuming and in an operational environment near real-time (NRT) capabilities are crucial [2, 3]. Since decades, methods for automatic ship detection are developed and improved [4]. While a robust direct detection of ship signatures is possible using analytical solutions based on first-order statistics [4, 5], the robust indirect detection of ships by their wakes requires methods from the field of machine learning (ML) [6].

The pursuit of dark vessels [7] requires dynamic information on course over ground or ship speed. Dynamic information cannot be derived by detecting ship signatures directly. Indirect detection of ships using their wake signatures enables the extraction of such dynamic ship information. Figure 1 provides an exemplary overview on dynamic information about ships, which can be derived indirectly from wake signatures, but not directly from ship signatures.

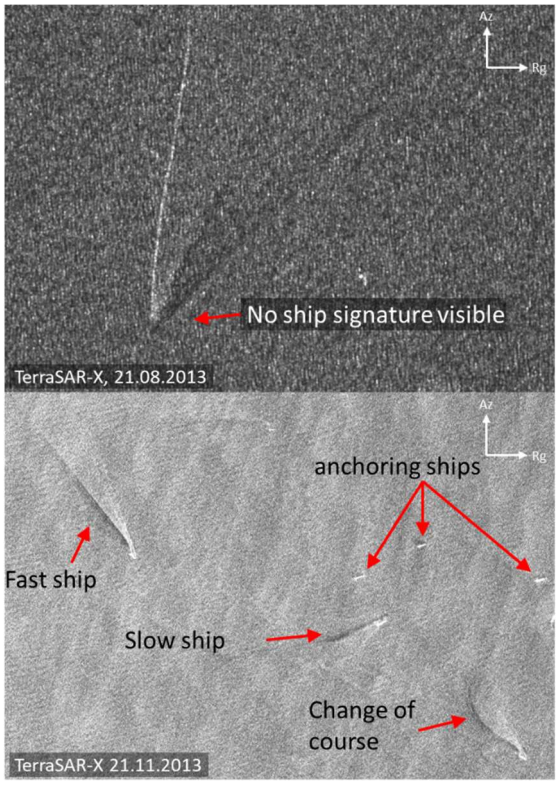

Figure 1: Examples of dynamic information derivable from ship wakes, e.g. detection of ships without imaged ship signatures (top) or derivation of ship's behaviors (bottom)

Wake signatures in SAR consist of multiple wake components. In this study, those wake components are considered, which are most frequently detectable in SAR imagery. The so-called Near-hull turbulences (NT), turbulent wakes (TW), Kelvin wake arms (KW) and V-narrow wakes (VW) are detectable for more than 20% of cases of moving ships [8, 9]. Those four wake components are schematically presented in Figure 2. Detailed information on the imaging of ship wakes in SAR data and the detectability of individual wake components is summarized in [8, 9].

The novelty of this study is the combination of wake detection with wake detectability models, in order to increase the accuracy of wake detection. An object detection framework called "You Only look Once" (YOLO) is adapted for the task of wake detection (specifically YOLOv4 is applied

[10]). After wake detection using the retrained YOLO-detector, the resulting detections are filtered on the basis of detectability models.

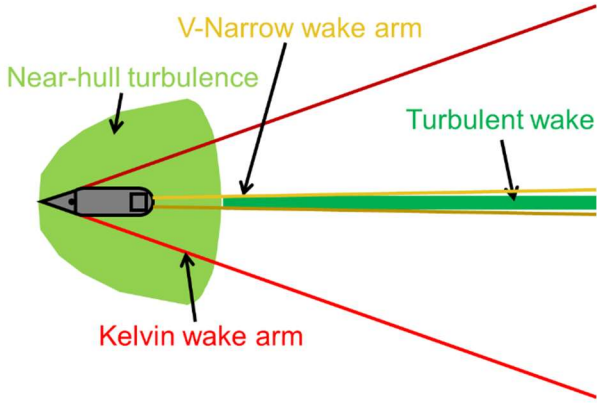

FIGURE 2: SCHEMATIC REPRESENTATION OF FOUR WAKE COMPONENTS MOST FREQUENTLY DETECTABLE IN SAR IMAGERY

II. DATA AND METHODS

In the following subsection A, the generation of SAR datasets for development and verification of the wake detector is described. Subsection B summarizes the ML-based modelling of wake detectability. The training of a YOLO-based wake detector is depicted in subsection C. Finally, subsection D introduces the method of static filtering of detection and explains the new method of dynamic filtering, both aiming for an increase in detection accuracy.

A. SAR datasets

SAR products from four different SAR missions, i.e. TerraSAR-X (TSX), CosmoSkymed (CSK), Sentinel-1 (S1) and RADARSAT-2 (RS2), are applied in this study. In Table I a summary of the datasets is provided. The proportion of TSX data is the largest in this study. Data from the other sensors are only added to increase the robustness of the developed wake detector. The used acquisition modes provide data with similar pixel spacing for TSX, CSK and RS2 of around 1 to 4 meters per pixel. The S1 data has uniform pixel spacing of 10 meters per pixel. TSX and CSK data is acquired using a X-band SAR sensor, while S1 and RS2 data is acquired by a C-band SAR sensor.

TABLE I. SUMMARY OF SAR DATASETS INCLUDING

	TSX	CSK	S1	RS2
Frequency band	X	X	C	C
Acquisitions modes	SL, SM	HI	IW	MF, F, S
Number of products	1097	11	31	53
Number of wake samples	2881	94	618	407
Proportion of wake samples without detected wake components [%]	21	17	29	41
Proportion of wake samples containing near-hull turbulences [%]	59	55	65	52
Proportion of wake samples containing turbulent wakes [%]	60	63	44	32
Proportion of wake samples containing Kelvin wake arms [%]	31	56	23	11
Proportion of wake samples containing V-narrow wake arms [%]	40	48	2	14

Manual inspection and retracing of wake components are executed on all SAR products to extract a collection of wake samples including measures for wake component length for each wake component. Each of the wake samples is collocated with AIS data and models for estimation of environmental parameters (i.e. XMOD-2 [11, 12], CMOD-5n [13], SAR-SeaStaR [14] and WRF [15]). The percentages in Table I show the resulting distributions of numbers of wake components in the wake samples. Some wake samples exist, where no wake component is detectable. Wake samples only partially collocated are discarded, so the final dataset contains in total 4000 wake samples with full context information.

An overview of datasets' preparation is provided in Figure 3. The training of detectability models is based on these 4000 wake samples. In order to train the wake detector, image patches are extracted for each wake sample containing sigma naught calibrated SAR backscatter information of the wakes and the surrounding ocean surface. The sigma naught intensity is converted to amplitude values, which are normalized between 0 and 2 and logarithmically rescaled. Land areas and ship signatures are masked by replacing the corresponding pixel values by the average of the surrounding ocean surface. The masking is applied to force the developed detector to learn the actual wake features and not to focus on the more prominent ship signatures.

Each image patch has pixel dimensions of 512×512 and a pixel size of ~10 m per pixel. The dataset has been augmented by the operations "vertical flipping", "horizontal flipping" and "vertical plus horizontal flipping". The operations simulate different settings of orbit looking direction (ascending and descending) and sensor looking direction (left and right), multiplying the dataset size by four. The dataset is split into training and evaluation sets with a relative proportion of 9:1; the training set is used for training and the evaluation set is used to compute the model performance during the training process.

For measuring the performance of the wake detector, a verification dataset based on 16 manually inspected SAR images is applied, which is not used by either of the ML models during training. The full scenes are split into image patches of same pixel spacing and dimensions as the image patches in the training data. An overlap of 256 pixels is applied between the split patches, resulting in 2136 samples in the verification set. The individual detections in the split patches are merged afterwards so that the detections can be displayed in the context of the full scenes.

B. Wake detectability modelling

In order to model the detectability of wake components, a figure of merit for quantification of their detectability is required. In this study, the detectable length of the wake components is used as an indicator for detectability. The length is normalized within a minimum and maximum length boundary to obtain a measure, which is in the following called Detectable Length Metric (DLM). The DLM in [8, 9] is physically dependent on the prevailing local situation defined by a set of nine influencing parameters.

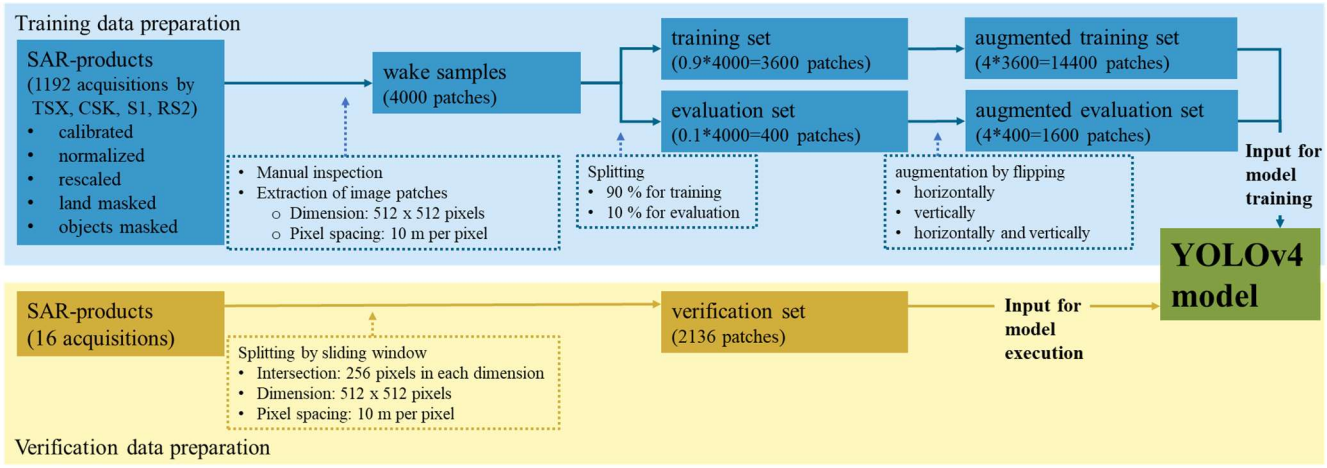

FIGURE 3: OVERVIEW OF DATA PREPARATION STEPS FROM PREPROCESSING OF SAR PRODUCTS OVER DEFINITION OF TRAINING, EVALUATION AND VERIFICATION DATASETS UNTIL DATASET'S INGESTION BY THE YOLO-BASED WAKE DETECTOR

However, the detectability models in this study are only based on a set of seven influencing parameters $X \in \{x_3; x_4; x_5; x_6; x_7; x_8; x_9\}$, as specified in Table II. The two removed parameters are ship's speed and ship's length, as those two parameters are only reliably available when the ship is already detected, which renders wake detection irrelevant. The ship's course over ground (x_3), can reliably be derived from the detected wake signature [16]. The YOLO-detector provides bonding boxes encompassing each detection. In this study, the ship's course over ground (x_3) is calculated from the bounding boxes. As x_3 is projected onto a value range of 0° to 90° [8, 9], the course over ground corresponds to the polar angle of the two pairs of facing corner points of a bounding box.

TABLE II: LIST OF SEVEN INFLUENCING PARAMETERS AS MOTIVATED IN [8]

ID	Influencing Parameter Name	Description
x_3	AIS-CoG	Course over ground (CoG) derived from AIS relative to the radar looking direction
x_4	Incidence-Angle	Incidence angle of the radar
x_5	SAR-Wind-Speed	Wind speed estimated from SAR backscatter of ocean background
x_6	SAR-Significant-Wave-Height	Significant wave height estimated from SAR backscatter of ocean background
x_7	SAR-Wave-Length	Wavelength estimated from from SAR backscatter of ocean background
x_8	AIS-CoG-SAR-Wave-Direction	Absolute angular difference between AIS-CoG and wave direction estimated from SAR backscatter of ocean background
x_9	AIS-CoG-WRF-Wind-Direction	Absolute angular difference between AIS-CoG and wind direction estimated by the Weather Research and Forecasting Model (WRF)

Detectability models are built by training "Support Vector Regression" (SVR) models for each wake component, which use the influencing parameters as input parameters (i.e. features) and DLM as output parameter (i.e. prediction parameter). This means, the detectability models for each wake component $w \in \{NT; TW; KW; VW\}$ can be defined by:

$$DLM_w = f_w(X), \tag{1}$$

where $DLM_w \in [0,1]$ defines the detectable length metric.

Details about the method for modelling wake component detectability can be found in [8] and details about uncertainty quantification for those models can be found in [9].

C. Wake detection

As a result of the manual retracing of wake components, the bounding boxes of SAR signatures of each wake component are known. Intersecting bounding boxes are considered as belonging to the same wake signature. The intersecting bounding boxes are merged to determine a bounding box comprising all wake components belonging to one wake signature. By doing so the YOLO detection model can be trained taking the wake components into account individually and as a composite belonging to one wake signature. YOLOv4 [10] is used in this study, so that multiple detections of the same image region under varying class labels are possible.

The main hyperparameter settings for the YOLOv4 training are specified in the following Table III.

TABLE III: YOLOV4 MAIN HYPERPARAMETER SETTINGS USED FOR TRAINING

YOLOv4 hyperparameter name	value
batch size / subdivisions	64 / 64
width x height	512 x 512
channels	3
learning_rate	0.001
mosaic augmentation	True

YOLOv4 includes a pretraining with RGB images. As the SAR images only contain one image channel with backscatter information, YOLOv4 converts those single channel images into RGB representation by duplication of the one channel.

D. Filtering of false alarms

The wake detector assigns confidence to each detected wake component w, which is in the following referred to by probability of detection (PoD_w). By setting a static threshold T_s, detections with low probability of detection can be filtered out. Such a static filtering can be described by the following function g_s, which is 1 for confirmed detections and 0 otherwise:

$$g_s(PoD_w) = \begin{cases} 1 & \text{if } PoD_w \geq T_s \\ 0 & \text{if } PoD_w < T_s \end{cases}, \tag{2}$$

where in this study a static filtering threshold of 25% is applied, i.e. $T_s = 0.25$. YOLO provides the option to apply an internal static filtering, which also applies a static filtering threshold of 25%, when activated. However, this option was deactivated, so that the dynamic filtering, which is described in the following paragraph, can also consider detections with lower PoD_w.

The DLM of the wake detectability models expresses how probable a detected wake under the prevailing local situation is. A detection is confirmed using the detectability models, when the sum of PoD_w and DLM_w reaches a dynamic filtering threshold of 100%, i.e. $T_d = 1$. All other non-confirmed detections are considered false positives (FP). By discarding those non-confirmed detections, the number of false positives can be decreased. The approach is in the following called dynamic filtering. Analogically to equation (2), dynamic filtering is described by the function g_d:

$$g_d(PoD_w, DLM_w) = \begin{cases} 1 & \text{if } \alpha \cdot PoD_w + \beta \cdot DLM_w \geq T_d \\ 0 & \text{if } \alpha \cdot PoD_w + \beta \cdot DLM_w < T_d \end{cases}, (3)$$

where DLM_w defines the detectable length metric of w. The factors α and β are introduced as PoD_w and DLM_w, originate from two independent statistical models. For this study, the factors are set to $\alpha = 1$ and $\beta = 2$.

The factors α and β as well as the thresholds T_s and T_d require tuning for optimal performance. Such tuning of the filtering hyperparameters has not been executed in the scope of this study, as the focus of this study is to demonstrate that any dynamic filtering outperforms static filtering.

III. Results

While for training of ML models only image patches are used, the verification of the detection performance is executed using the complete scenes in the verification dataset. By doing so, the verification process reflects the actual application of the wake detection system.

The verification dataset contains 86 confirmed wakes, which are either true positive (TP) when identified by the detector, or false negative (FN) when not identified by the detector. Additionally, 14 false positives (FP) have been identified by the detector, which are ship signatures without wakes e.g. in anchoring state or artefacts on the ocean surface like eddies. After static and dynamic filtering some FPs are reclassified as true negatives (TN). Also, some TPs become FNs after the static filtering. TNs in this study are cases, where in the vicinity of ship positions known from AIS the absence of wake signatures is confirmed.

An overview of the verification procedure for the wake detector's performance including the filtering of false alarms during post-processing is shown in Figure 4. Some new detections are not present in the verification set, as no AIS-collocation is possible for some TPs or the cases represent unexpected FPs. Thus, the verification procedure involves a manual confirmation step, where to each new detection verification information is added. In case of TPs, adding verification information involves a manual retracing of wake components, in case of FPs, their nature of origin is flagged (e.g. eddy or wind shadow). Based on this verification information, all wake samples are automatically categorized into TP, FP, TN or FN.

In the Table IV and the Table V performance of wake detection is compared between the wake detection results without filtering applied and with static or dynamic filtering applied during post-processing step. In Table IV wake samples are categorized into TP, TN, FP and FN, respectively. Table V provides measures of performance based Table IV.

Table IV: Confusion matrix for verification of wake detection

Number of samples: 107		**Measurement**	
		Positive	***Negative***
Prediction without filtering	***Positive***	TP count: 42	FP count: 7
	Negative	FN count: 44	TN count: 14
Prediction static filtering	***Positive***	TP count: 38	FP count: 5
	Negative	FN count: 48	TN count: 16
Prediction dynamic filtering	***Positive***	TP count: 42	FP count: 6
	Negative	FN count: 44	TN count: 15

Table V: Statistical quality criteria for verification of wake detection without filtering

	Precision	**Recall**	**True negative rate**	**Accuracy**	**F1-Score**
without filtering	0.86	0.49	0.67	0.52	0.62
Static filtering	0.88	0.44	0.76	0.5	0.59
dynamic filtering	0.88	0.49	0.71	0.53	0.63

Static and dynamic filtering both increase the precision. When applying static filtering, the increase of precision comes with a decrease of recall, where the decrease is higher than the increase so that overall performance is decreased. The level of recall is maintained, when dynamic filtering is applied, while the same level of precision is reached as compared to the static filtering. As a result, the overall performance is increased by application of dynamic filtering.

Discussion and Conclusion

This study introduces a dynamic filtering approach, which can be applied during post-processing to reduce false alarms of a wake detection systems. Instead of comparing the probability of detection provided by the wake detection system to a static threshold, the proposed method utilizes detectability models to determine a dynamic threshold and by doing so taking the detection conditions into account.

Different detector variations have been combined with the two filtering methods and varying settings of filtering hyperparameters. The detector variations imply different preprocessing of training data (e.g. masking of ship signatures by minimum or maximum instead of mean backscatter), changes to YOLO-hyperparameters (e.g. different dimensions of width and height of image patches) and activated or deactivated training dataset augmentation as well as training dataset extension by non-wake samples using confirmed locations eddies. While the accuracy performance of all variations differs, the advantage of dynamic filtering over static filtering was constantly observed. As the dynamic filtering is applied during a post-processing step, it is concluded that the proposed method is capable of reducing the amount of false alarms for wake detections systems, which do not already consider influencing parameters or wake component detectability and are also capable of assigning probabilities of detection to their detections.

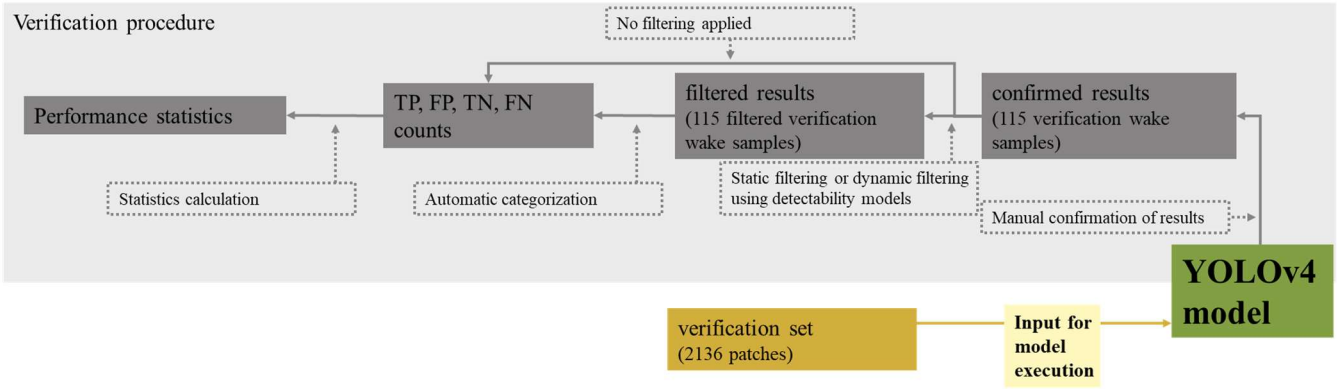

FIGURE 4: OVERVIEW OF VERIFICATION PROCEDURE FROM MODEL INGESTION TO PERFORMANCE STATISTICS

Providing the influencing parameters as an additional input to the training procedure of the wake detector would be an alternative to the dynamical filtering during post-processing. However, resources of deriving influencing parameters might not be available for wake detector training conducted by other researchers. From current literature no wake detector is known, which takes the influencing parameters or detectability into account. Therefore, many wake detectors will benefit from the two-step approach suggested in this study.

Further, it is assumed that the proposed method for dynamic filtering of false alarms is also applicable to other detection tasks. Requirement for the application would be that the detectability of the objects, which are to be detected, is influenced by physical conditions and that those physical conditions can be quantified and used for detectability modelling.

REFERENCES

[1] B. J. Tetreault, "Use of the Automatic Identification System (AIS) for maritime domain awareness (MDA)," *Proceedings of OCEANS 2005 MTS/IEEE,* vol. 2, pp. 1590-1594, 2005.

[2] European Maritime Safety Agency, "Cleanseanet Service - Detecting Marine Pollution from Space," European Maritime Safety Agency, Lissabon, 2019.

[3] S. Lehner and B. Tings, "Maritime Products using TerraSAR-X and Sentinel-1 Imagery," in *The 36th International Symposium on Remote Sensing of Environment*, Berlin, 2015.

[4] D. Crisp, "The State-of-the-Art in Ship Detection in Synthetic Aperture Radar Imagery," DSTO Information Sciences Laboratory, Edinburgh, 2004.

[5] S. Brusch, S. Lehner, T. Fritz, M. Soccorsi and A. Soloviev, "Ship Surveillance With TerraSAR-X," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 49, no. 3, pp. 1092-1103, 2011.

[6] R. Del Prete, M. D. Graziano and A. Renga, "RetinaNet: A deep learning architecture to achieve a robust wake detector in SAR images," in *2021 IEEE 6th International Forum on Research and Technology for Society and Industry (RTSI)*, Naples, 2021.

[7] F. S. Paolo, D. Kroodsma, J. Raynor, T. Hochberg, P. Davis, J. Cleary, L. Marsaglia, S. Orofino, C. Thomas and P. Halpin, "Satellite mapping reveals extensive industrial activity at sea," *Nature,* vol. 625, pp. 85-91, 2024.

[8] B. Tings, "Non-Linear Modeling of Detectability of Ship Wake Components in Dependency to Influencing Parameters Using Spaceborne X-Band SAR," *Remote Sensing,* vol. 13, no. 2, p. 165, 2021.

[9] B. Tings, A. Pleskachevsky and S. Wiehle, "Comparison of detectability of ship wake components between C-Band and X-Band synthetic aperture radar sensors operating under different slant ranges," *ISPRS Journal of Photogrammetry and Remote Sensing,* vol. 196, pp. 306-324, 2023.

[10] A. Bochkovskiy, C. Yao Wang and H. Yuan Mark Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," *arXiv-CoRR,* vol. abs/2004.10934, 2020.

[11] X.-M. Li and S. Lehner, “Algorithm for Sea Surface Wind Retrieval From TerraSAR-X and TanDEM-X Data,” *IEEE Transactions on Geoscience and Remote Sensing,* vol. 52, no. 5, pp. 2928-2939, 2014.

[12] S. Jacobsen, X. Li, S. Lehner, J. Hieronimus and J. Schneemann, "Joint Offshore Wind Field Monitoring with Spaceborne SAR and Platform-Based Doppler LiDAR Measurements," *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences,* vol. 40, p. 959–966, 2015.

[13] F. Monaldo, C. Jackson, X. Li and W. G. Pichel, "Preliminary Evaluation of Sentinel-1A Wind Speed Retrievals," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* vol. 9, no. 6, pp. 2638-2642, 2016.

[14] A. Pleskachevsky, B. Tings, S. Jacobsen, S. Wiehle, E. Schwarz and D. Krause, "A System for Near Real Time Monitoring of the Sea State using SAR Satellites," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 62, pp. 1-18, 2024.

[15] W. C. Skamarock, J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, M. G. Duda, X.-Y. Huang, W. Wang and J. G. Powers, “A Description of the Advanced Research WRF Version 3,” NCAR Technical Notes, Boulder, 2008.

[16] M. D. Graziano, M. D’Errico and G. Rufino, "Wake Component Detection in X-Band SAR Images for Ship Heading and Velocity Estimation," *Remote Sensing,* vol. 8, no. 6, p. 498, 2016.