Earth System
Science
Data
Open Access
Discussions

DeepOWT: A global offshore wind turbine data set derived with deep learning from Sentinel-1 data

Thorsten Hoeser[1], Stefanie Feuerstein[1], and Claudia Kuenzer[1,2]

[1]German Remote Sensing Data Center (DFD), German Aerospace Center (DLR), 82234, Wessling, Germany
[2]Department of Remote Sensing, Institute of Geography and Geology, University of Wuerzburg, 97074, Wuerzburg, Germany

Correspondence: Thorsten Hoeser (thorsten.hoeser@dlr.de)

Abstract. Offshore wind energy is at the advent of a massive global expansion. Driven by carbon neutral alternatives for energy generation, offshore wind energy receives growing attention as a renewable energy source. Despite the large amount of unused wind energy capacities worldwide, offshore wind farms have to be integrated into already intensively used maritime economic areas. The optimal choice of offshore wind farm locations is as crucial as compatibility with other stakeholders while minimising ecological impacts. Thus, a spatiotemporal data set for offshore wind turbine deployment is necessary to involve all stakeholders and exchange knowledge during the upcoming massive expansion of offshore wind farms. To that end, we introduce the DeepOWT data set (global offshore wind turbines derived with deep learning; available at: https://doi.org/10.5281/zenodo.5933967 (Hoeser and Kuenzer, 2022)), which provides 9,941 locations of offshore wind energy infrastructure along with their deployment stages on a global scale. DeepOWT is based on freely accessible Earth observation data from 2016 until 2021. The locations were derived from radar imagery of the Sentinel-1 mission by applying deep learning based object detection, trained on synthetic training examples. The entire deployment process is reported in a quarterly frequency and spatially contextualised for each single wind turbine location in a ready to use GIS format. Therewith, the DeepOWT data set can directly be used to enable spatial planning, environmental investigations and to optimise location decisions and the deployment process.

1 Introduction

Lately, the expansion of carbon neutral energy is being strongly promoted (COP26, 2021). Offshore wind energy is an efficient and reliable energy source and appears to be an important cornerstone for a renewable energy mix (Esteban et al., 2011). For example, the EU plans to increase the installed offshore wind energy capacity from 12 GW in 2020 to 300 GW in 2050 towards a carbon neutral energy production. Most of this expansion is planned to be realised in the North Sea Basin, an already established hot spot for offshore wind energy production. Nevertheless, new sites in the Mediterranean Sea will be developed to achieve the stated goals (EC, 2020). The plans of the European Union are exemplary for a global trend of expanding offshore wind energy projects. Offshore wind farms (OWF) will be deployed to already established offshore wind energy production sites like the East China Sea. At the same time, the development of new sites for large OWFs, for example, in the Atlantic Ocean on the East coast of the United States, is ongoing (Rodrigues et al., 2015). Thus, today the offshore wind energy sector

25   is starting a phase of massive expansion worldwide which also affects stakeholders active in the same marine economic zone (Henderson et al., 2003; Wever et al., 2015; Gusatu et al., 2020; Gușatu et al., 2021) as well as the marine ecosystem itself (Drewitt and Langston, 2006; Wilson and Elliott, 2009; Bailey et al., 2014; Bergström et al., 2014; Slavik et al., 2019). In order to document this starting point of OWF expansion and provide a database accessible for all stakeholders, we introduce the DeepOWT data set, which reports offshore wind turbine (OWT) locations along with their deployment time series for five

30   years from 2016 until 2021 on a global scale.

With the increase in freely available Earth observation data over the last years, today well structured archives of remote sensing data offer the opportunity to investigate time series on a global scale. Due to the recent successes in machine learning, especially deep learning (LeCun et al., 2015; Zhu et al., 2017) with convolutional neural networks (Hoeser and Kuenzer, 2020), it has become possible to apply novel data processing techniques to these data archives to reveal new insights (Hoeser et al.,

35   2020; Zhu et al., 2021).

The proposed DeepOWT data set has been derived from the Sentinel-1 spaceborne radar archive by applying deep learning based object detection, which was completely trained on synthetic training examples generated by the novel SyntEO approach (Hoeser and Kuenzer, 2021). Therewith, DeepOWT is based on a methodological workflow that allows to extract highly detailed information from large Earth observation archives. The most important novelty of the resulting data set is that, along

40   with global OWT locations, it provides deployment stages for each detected object over the entire time series from 2016 until 2021. This is of major interest for investigations of the impact and optimisation measures during the deployment processes, which is an eventful and critical period in an OWT lifecycle. Besides OWT locations, DeepOWT also provides locations of OWF substations like transformer stations and meteorological masts to increase further the information depth and precision of a global OWT data set. Besides this novel information, DeepOWT is openly accessible, comes along with valuable ground

45   truth data sets for spatial and temporal validation and can easily be used in GIS software due to its lightweight size of 4.1 MB and established .geojson format.

2   Related Research

The detection of persistent offshore infrastructure has been investigated by applying different approaches. The constant false alarm rate (CFAR) (Zhang et al., 2019) and an adapted version of it in combination with a preprocessing by the difference of

50   Gaussians (DoG) approach (Wong et al., 2019) were used to detect marine infrastructure in radar imagery. Xu et al. (2020) applied order statistic filtering in combination with a derived set of thresholds to extract marine infrastructure from the Landsat and Sentinel-2 missions. Zhang et al. (2021) developed a handcrafted morphologic approach with thresholds to identify single OWTs in Sentinel-1 data. Thereby, all studies strongly depend on the high contrast between offshore infrastructure and the surrounding open sea. In a preceding study to this data set, Hoeser and Kuenzer (2021) proposed an adaptive deep learning

55   based object detection approach that takes multiscale spatial patterns of the target objects into account to further distinguish different types of offshore infrastructure in a single model.

In 2022, three data sets exist that describe OWT locations on a global scale and parts of which are accessible to the public. These are the 4C offshorewind data set (4C Offshore, 2021), OWT locations in open street map (OSM) and the global offshore wind turbine (GOWT v1.3) data set (Zhang et al., 2021). The 4C offshorewind data set collects information about offshore wind
60 projects provided by OWF operators, project descriptions and contracts. It gives an overview of OWF boundaries, single OWT locations and specifications, and OWF substations like transformer stations and export cables (4C Offshore, 2021). However, the private company 4C Offshore maintains the 4C offshorewind data set and sells the information. Therewith, the data set is only partly accessible to the public and can not freely be used in planning and research to its full extent.

The first open source variant is the OSM project which provides OWT locations in its spatial database. Since OSM data
65 relies on the activity of its community, the accuracy and completeness for OWT locations vary from region to region. Hence, the data set accuracy is not spatially homogenous. Furthermore, there is only a limited temporal consistency in this data set. An entry made in OSM on a specific date does not necessarily correlate with its first appearance in the real world, especially when the temporal accuracy is narrowed down to weeks or months.

The GOWT v1.3 data set demonstrated the possibility of large scale OWT detection by investigating Earth observation
70 data. In its published version 1.3 it provides OWT locations from 2014 until 2019. It shows OWT locations from the time the construction started. Thereby, it does not distinguish between turbines under construction and those completed (Zhang et al., 2021). This class indifference leads to a temporarily shifted overestimation of the number of power generating OWTs since the construction phase is not provided separately. GOWT v1.3 does not separately classify OWF substations like transformer stations or meteorological masts and has difficulties differentiating them from OWTs, which results in false positive detections
75 of OWTs within OWFs areas. The published data set GOWT v1.3 does not include all recently built large offshore wind farm projects. One example is the Hornsea project of the United Kingdom, the currently largest OWF project in the North Sea Basin. Despite over already 170 available OWT records in the underlying Earth observation data in 2019, they do not appear in GOWT v1.3, which causes a high number of false omissions.

From these existing data sets and their limitations, the following characteristics have been derived for a global OWT data
80 set:

- Access to all OWT information like location and construction stage in a single file

- Global extent with homogenous reliability

- Temporal consistency regarding the date given in the data set for a single data point and its real world appearance

- Differentiation between OWTs which are under construction and OWTs which are completed

85 - Differentiation between OWTs and OWF substations like transformer stations or meteorological stations

- Inclusion of latest OWF projects independent of their size, location and construction type

To that end, we introduce DeepOWT (deep learning derived global offshore wind turbines), an independent and openly accessible data set of offshore wind energy infrastructure locations and their temporal deployment dynamics on a global scale.

It was derived by applying deep learning based object detection on ESA's spaceborne Sentinel-1 synthetic aperture radar (SAR)
90   archive. DeepOWT provides OWT locations along with their quarterly deployment stages from July 2016 until June 2021. It
differentiates between OWTs under construction and OWTs which are readily deployed. Furthermore, it includes locations of
OWF substations as a third class. By using two cascading deep learning convolutional neural network (CNN) object detectors,
first OWFs and upon them, single OWTs and substations are detected. Previously, the CNNs were trained on large synthetic
data sets, which were generated with the lately introduced SyntEO approach (Hoeser and Kuenzer, 2021). That way, new
95   developments of OWFs in size and geometry can be introduced during the training process even when their number in real
world data is small, which successfully decreases false omissions of recently developed OWFs projects. To our knowledge,
DeepOWT is the first data set that provides spatiotemporal information at a single object level and multiple deployment stages
on a global scale derived from spaceborne SAR-data with deep learning.

3   Data

100   ESA's Copernicus program provides open access to continuously acquired Earth observation data (Aschbacher, 2017). As part
of this program, the spaceborne Sentinel-1 SAR mission covers mainland and coastal areas on a global scale with a 10 m pixel
spacing. The active C-band radar system with a wavelength of 5.6 cm is independent of cloud coverage and can acquire images
day and night (Torres et al., 2012), which together makes it an excellent tool to monitor coastal environments and provide a
data source for a global OWF investigation. All Sentinel-1 acquisitions with the specification IW (interferometric wide swath),
105   GRD (ground range detected) and VH (vertical horizontal) polarised were chosen as underlying data for this data set. Figure
1a) shows how often a location on Earth is sensed by the two Sentinel-1 satellites A and B for this specification in the second
quarter of the year 2021. The European focus of the Copernicus program becomes visible with a higher number of acquisitions.
Here the satellites acquire data at both ascending and descending orbits, with an inclination of 98.18°, resulting in the X like
pattern and a higher revisit rate than other parts of the Earth. In order to harmonise all quarterly acquisitions of the entire Earth
110   to a single global mosaic with a pixel spacing of 10 m, all acquisitions were stacked and reduced to a single band median
image, see figure 1b).

The extent of the data set was defined by a buffer of 200 km towards the sea from the global coastline, which originates
from OSM. In order to systematically manage and process the data, a 1.8° regular grid was generated, and all grid cells which
intersect the 200 km buffer were selected. The final grid, pictured in figure 1b, defines the area where the DeepOWT data set
115   was detected.

The Google Earth Engine (Gorelick et al., 2017) python API was used with the prepared grid cells to query the Sentinel-1
data as specified above. The Sentinel-1 image collection in Google Earth Engine provides the IW GRD VH acquisitions with
additional preprocessing for accurate orbit information, border and thermal noise removal, radiometric calibration and terrain
correction. The global median mosaic with a pixel size of 10 m $\times$ 10 m was generated and downloaded from this image
120   collection. To reduce the amount of data before downloading, the 16 bit floating-point number with a range of 0 to -40, which
describes the Sentinel-1 backscatter signal in decibel, were rescaled to 0 to 255 and downloaded as 8 bit integers. The global
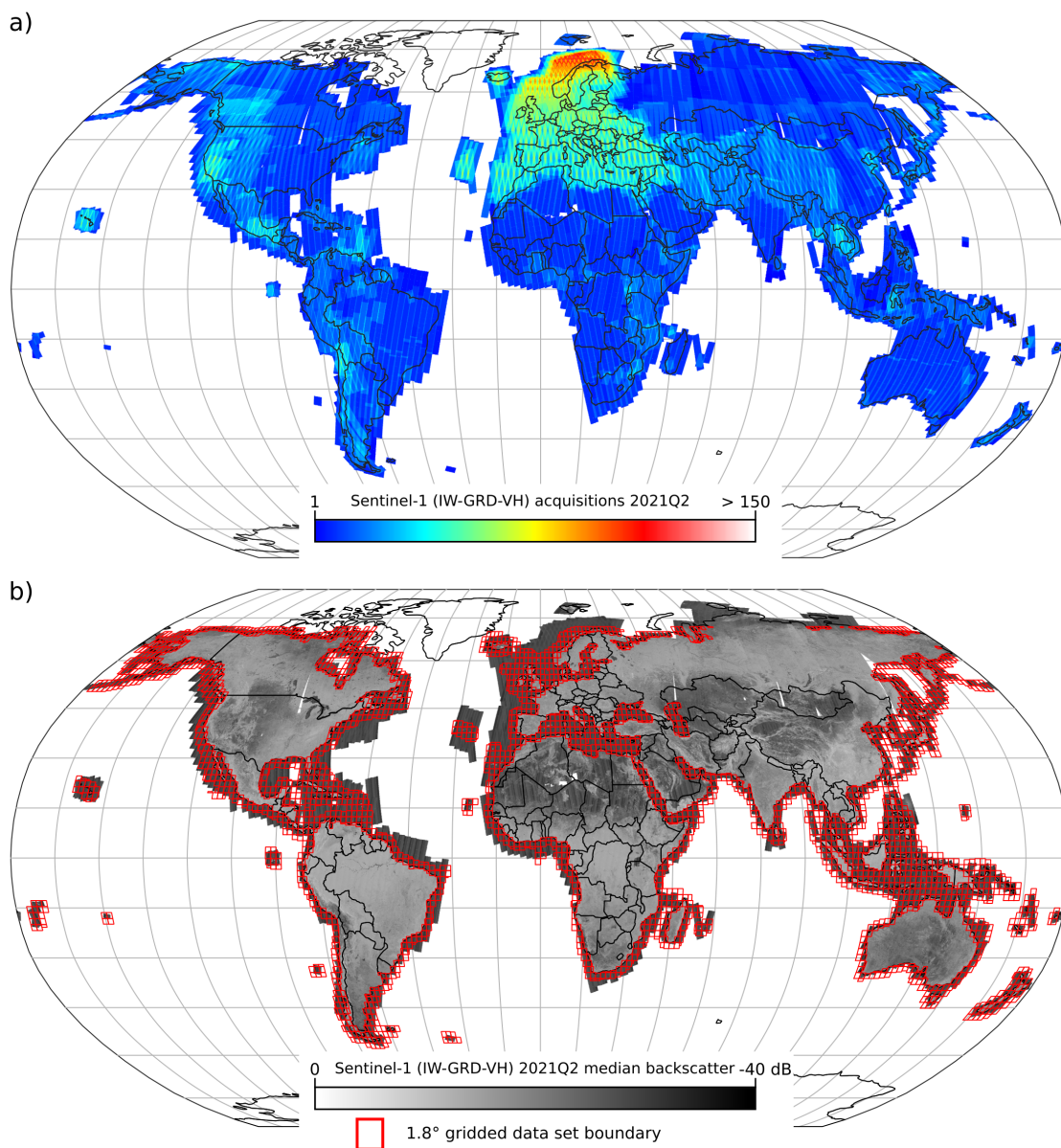
Earth System
Science
Data
Open Access
Discussions

a)

b)

Figure 1. a) Global distribution of the number of all available Sentinel-1 (IW-GRD-VH) acqusitions for the second quater in 2021. b) The corresponding median backscatter amplitude and the data set boundary as 1.8° grid within a buffer of 200 km of the global coastline.

mosaic was queried for the second quarter of 2021 (2021-Q2). Upon this mosaic, the global OWF and OWT locations were derived.

Quarterly subsets from 2016-Q3 to 2021-Q1 were generated the same way as the 2021-Q2 global mosaic to investigate
125 temporal deployment dynamics. However, for the 19 quarterly subsets, only the grid cells were queried, which include OWF areas detected in 2021-Q2. Thus, the underlying data for the DeepOGWT data set are 20 quarterly sets from 2016-Q3 to 2021-Q2. The latest period in 2021-Q2 holds a Sentinel-1 median mosaic of the entire global coastline. All other 19 sets only store Sentinel-1 median images for OWF areas present in 2021-Q2.

4   Methodology

130 Figure 2 shows an overview of the methodological workflow to derive the OWT locations and their temporal deployment dynamics of the proposed data set. The core of the workflow is a cascade of two CNN object detectors, where the first CNN detects the OWF boundaries and the second CNN detects OWT locations on the 2021-Q2 global input data. In order to train both object detector CNNs, two synthetic training data sets were generated in a preceding step. This is the first application of the recently introduced SyntEO approach (Hoeser and Kuenzer, 2021) embedded in a complete workflow to generate a
135 global data set. After the OWT locations have been detected in the 2021-Q2 data, the second input data set, a time series of 19 quarterly time steps, is investigated to derive deployment dynamics for each OWT location from 2016-Q3 until 2021-Q1. In the next step, OWT locations are refined from a bounding box to accurate point locations. Finally, these point locations and the time series information corresponding to each OWT location are combined in the final DeepOWT data set in a single file.

4.1   Synthetic training data set generation with SyntEO

140 In deep learning, layers of untrained weights are stacked to build deep neural networks. In order to adjust the weights in a neural network to succeed in a given task, a large data set with annotated examples is necessary for supervised training (LeCun et al., 2015). The annotation of such large training data sets by hand is a resource expensive task and, in the case of only a few real world examples, which applies for OWFs, even impossible. In order to solve these problems, the SyntEO approach was introduced to automatically synthesise large annotated training data sets with a special focus on the needs of
145 Earth observation data. In SyntEO, a domain expert formulates an ontology that describes entities in a remote sensing scene and their spatiotemporal interrelations. The SyntEO ontology is a complex description of nested entities that are additionally related by using topological rules to describe their spatial dependencies. A synthetic image is generated upon the formulated ontology by composing spatially meaningful geometries of all entities in an abstract scene composition. Hereinafter, sensor specific texture is added to the geometries of the abstract scene composition to generate the final remote sensing image.
150 Furthermore, annotations are derived from the discrete geometries of the scene composition simultaneously. That way, large deep learning ready data sets can be created fast and automatically. Figure 3 shows a visual summary of the SyntEO process for a better intuition. For an in-depth explanation of the SyntEO framework and the underlying ontology concept for automatic data generation in Earth observation, we refer to (Hoeser and Kuenzer, 2021).
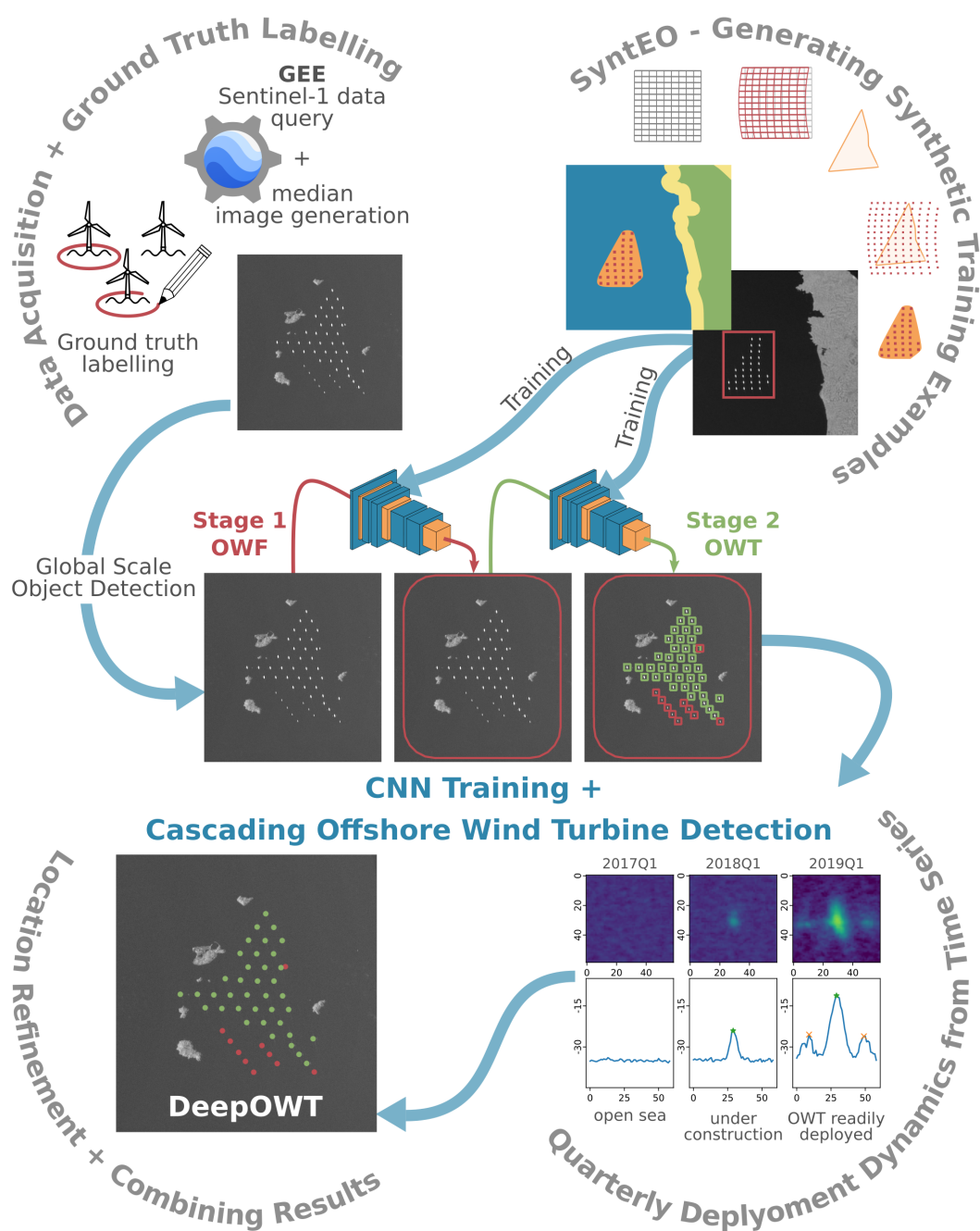
Figure 2. Methodological workflow to generate the DeepOWT data set from the Sentinel-1 archive. Two CNNs detect OWF and OWT locations, trained on synthetic examples. These spatial detections are used to define their temporal dynamics and accurate location. Finally, DeepOWT combines these spatiotemporal results.
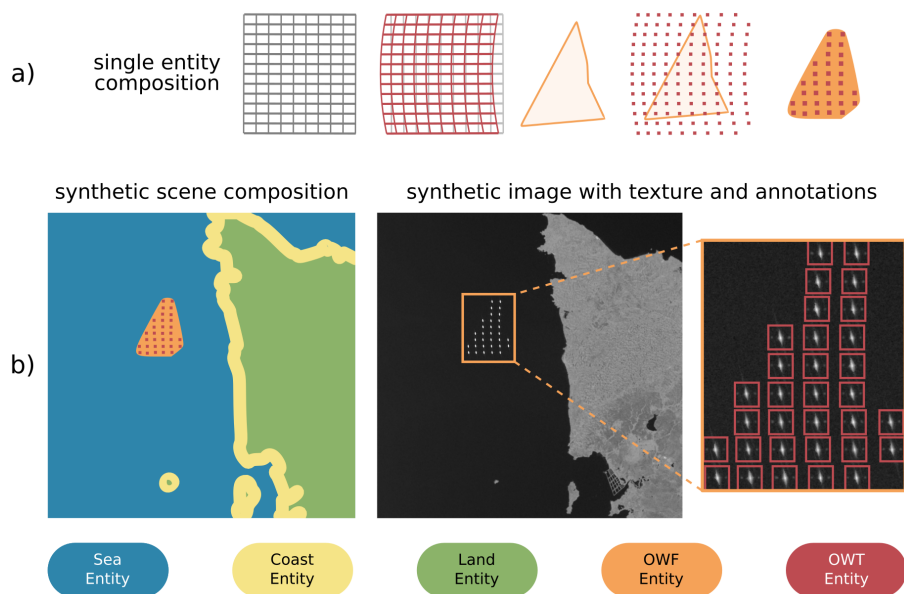
Figure 3. Simplyfied overview of a SyntEO workflow for automatic training example generation (Hoeser and Kuenzer, 2021). a) visualises how the structure of a single OWF entity is generated. b) shows how multiple entities are composed to a scene composition. The final image is generated by adding texture to the composed scene. The bounding box annotations are derived from the scene composition.

For this study, the ontology drafted by Hoeser and Kuenzer (2021) was extended to generate training examples for OWF and OWT detection at the same time. In addition to the origin ontology, which only describes polygon shaped OWFs, linear shaped OWFs were introduced to increase the target object variance of the training examples. To enrich the training data set, non-target classes like synthetically generated oil rigs and mainland were added with explicitly no annotations to provide negative training examples.

Furthermore, the ontology which has been formulated to generate training examples for OWF detection was reused for OWT detection. Most importantly, the size of the generated images and the annotation were changed. Instead of large scale bounding box annotations for OWFs in synthetic images with a dimension of 2048×2048 pixel, small scale bounding box annotations for OWTs were derived from synthetic images with a dimension of 512×512 pixel, see figure 3b).

In order to include other targets besides readily deployed OWTs, the non-target class oil rigs in OWF detection was reused to provide annotated examples for OWF transformer stations and a third class showing OWFs under construction was introduced. The mainland examples without any annotation were kept to provide false positive training examples for OWT detection.

Two pools of training-annotation pairs with additional metadata were generated to create the synthetic training data sets. Two balanced training data sets were compiled from these pools by taking the synthetic image type described in the metadata into account. To enable the TensorFlow deep learning framework (Abadi et al., 2016), the selected image annotation pairs were parsed to the TFrecord binary format. That way, a first training data set with 90,000 examples for OWF detection and a second training data set with 275,000 examples for OWT detection were created.

Earth System
Science
Data

Open Access

Discussions

4.2 Global OWT Object detection with CNNs

Deep learning became an important driver for new insights and methodological developments in Earth system science (Zhu et al., 2017; Reichstein et al., 2019). In Earth observation, the convolutional neural network (CNN) is the most widely used deep learning model type, and recent developments of CNNs allow to detect objects in large images by taking spatial context into account (Ma et al., 2019; Hoeser and Kuenzer, 2020). For object detection, the two stage R-CNN models are the most commonly used architecture in Earth observation applications (Hoeser et al., 2020). To derive the DeepOWT data set, we used a cascade of two ResNet-50 (He et al., 2016) Faster R-CNN (Ren et al., 2015) models, where the first stage detects OWF areas on a global scale and the second stage detects single OWT and non-OWT objects within the OWF areas of the first stage.

The first CNN for OWF detection ingests images with a 1024×1024 pixel dimension. That requires an on the fly downscaling of the training examples for OWF detection with a dimension of 2048×2048. For OWT detection, the training examples already have a dimension of 512×512 that matches the CNN architecture. Another difference in the model architectures for OWF and OWT detection are the adapted scale factors for the region proposal network (RPN), a submodule of the Faster R-CNN object detector (Ren et al., 2015). In order to adjust the sensitivity for specific sizes of the target objects, the scale ratios were set to [0.25, 0.5, 1, 2, 3.5] for OWF and [0.25, 0.5] for OWT detection. The scale factors were calculated by applying the approach introduced by Hoeser and Kuenzer (2021).

The entire training process of the two CNNs was done on four parallel NVIDIA RTX 2080Ti GPUs. The training schemes are the same for both architectures. A 0.95-0.05 training-validation split was used for both synthetic data sets. A single epoch without any data augmentation was used, which was possible due to the large size and variability of the synthetic data set. The learning rate was scheduled to smoothly decrease by implementing the cosine decay method (Loshchilov and Hutter, 2017). After a warm up phase, the base learning rate of 0.01 is reached and then decreased for all remaining training steps.

The trained models were used in a cascading manner. The first stage detects OWF by applying a threshold of 0.5 on the prediction score of the CNN. Thus the first stage tends to allow a higher false positive rate in order to include more areas as necessary but therewith avoids false omissions of OWF areas. The second stage detects OWTs and non-OWTs within the potential OWF areas by applying a threshold of 0.75 to consider a prediction valid. With the second stage's results, which focus on a smaller spatial scale, the predictions of the first stage are refined. Potential OWF areas with a share of 90% or more non-OWTs are belatedly dropped as false detection of stage one. This self-checking property of the cascade leads to a high detection rate with a low number of false omissions by simultaneously decreasing false detections. This property was highly suitable for precisely scanning Earth observation archives to find sparsely scattered target objects in a large amount of image data.

$$
\mathcal{B}_1 := \sum_{b=1}^{B-1} \begin{cases} \mathcal{B}_1 \cup \mathcal{B}_{b+1}, & \text{if} \quad \mathcal{B}_{b+1} \subseteq \mathcal{B}_1 \\ \mathcal{B}_1 \cup \mathcal{B}_{b+1}, & \text{if} \quad \frac{\mathcal{B}_1 \cap \mathcal{B}_{b+1}}{\mathcal{B}_1 \cup \mathcal{B}_{b+1}} > \tau_{\text{IoU}} \\ \mathcal{B}_1, & \text{otherwise} \end{cases}
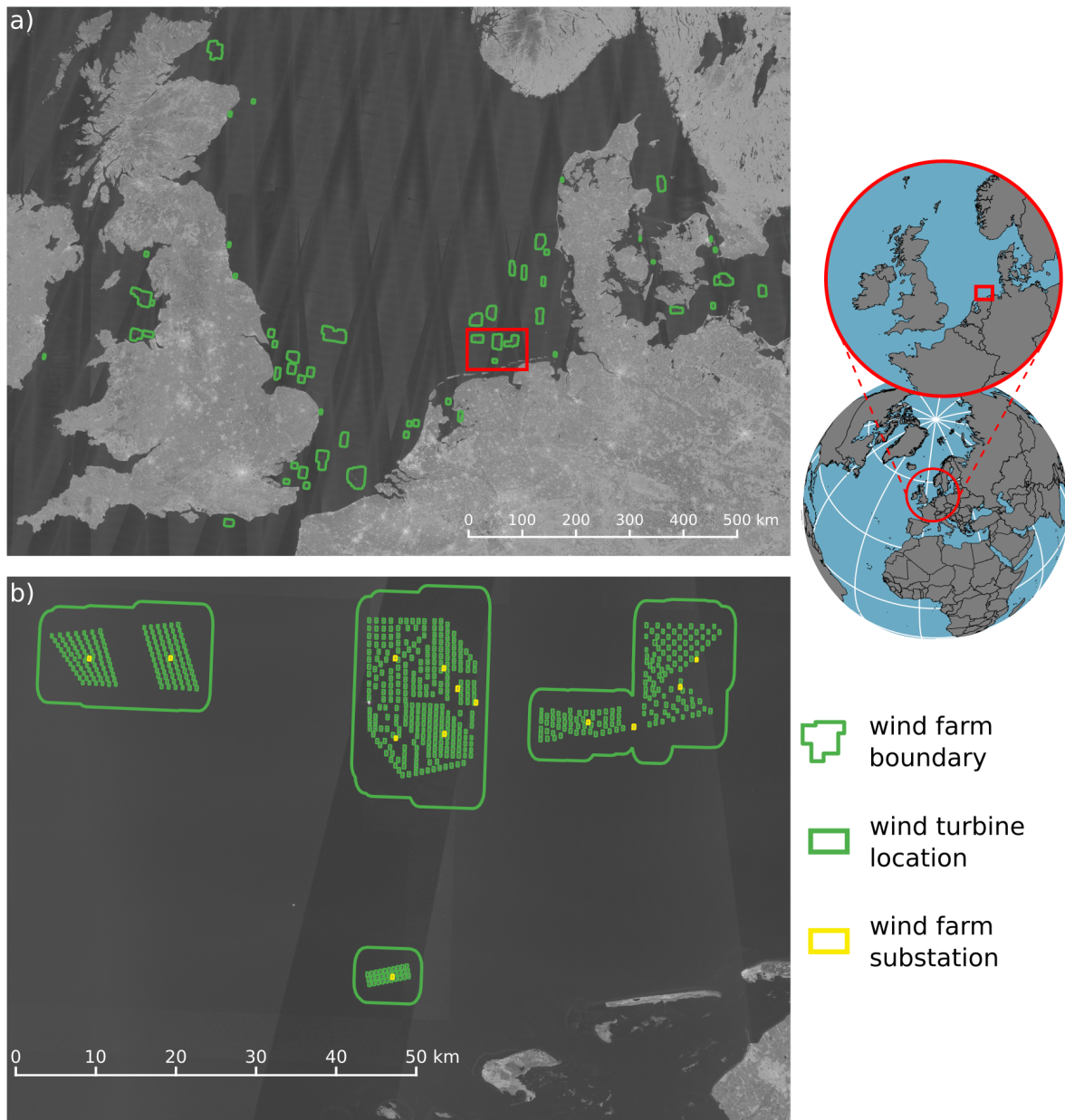\tag{1}
$$

Figure 4. Detection results of the CNN cascade for the North Sea Basin. a) shows the OWFs detected by stage 1, b) the OWT and OWF substations detected by stage 2 within the OWF boundaries of stage 1 in the German Bight.
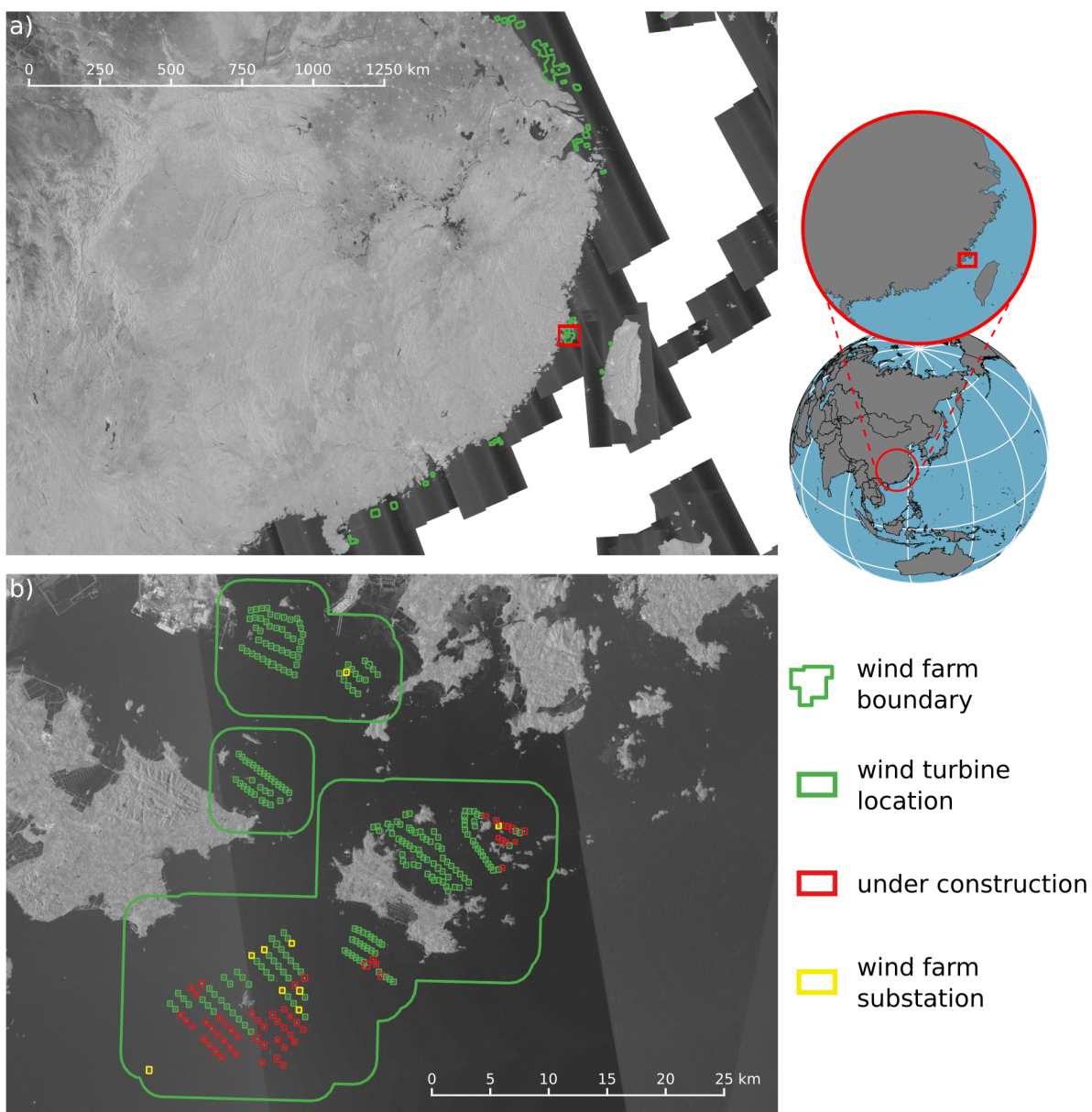
Figure 5. Detection results of the CNN cascade for the South and East China Sea. a) shows the OWFs detected by stage 1, b) the OWT, OWF substations and such under construction detected by stage 2 within the OWF boundaries of stage 1 in the Taiwan Strait.

Since all predictions are made on overlapping input tiles, the same object can be detected multiple times. To summarise all bounding boxes in one file, the pixel coordinates of the bounding boxes are translated to the geographic coordinate system WGS84. Furthermore, to consolidate the bounding boxes $\mathcal{B}$ which belong to the same object, they were sorted in descending order by their prediction score and indexed with $b = 1, 2, 3 ... B$. This sorted list of bounding boxes $\mathcal{B}_b$ was unified in a cascading

205   manner if the next box of the list is completely within the unified box or together they have an intersection over union larger
than $\tau_{\mathrm{IoU}} = 0.333$, see equation (1).

In order to calculate the exact area of each $\mathcal{B}_b$, the bounding box polygons were temporarily reprojected to their correspond-
ing UTM coordinate system depending on their UTM zone. Finally, the bounding box from the unified polygon describes
each object location. Figure 4 and 5 show the detection results of the CNN cascade for North Europe and South-East China

210   respectively.

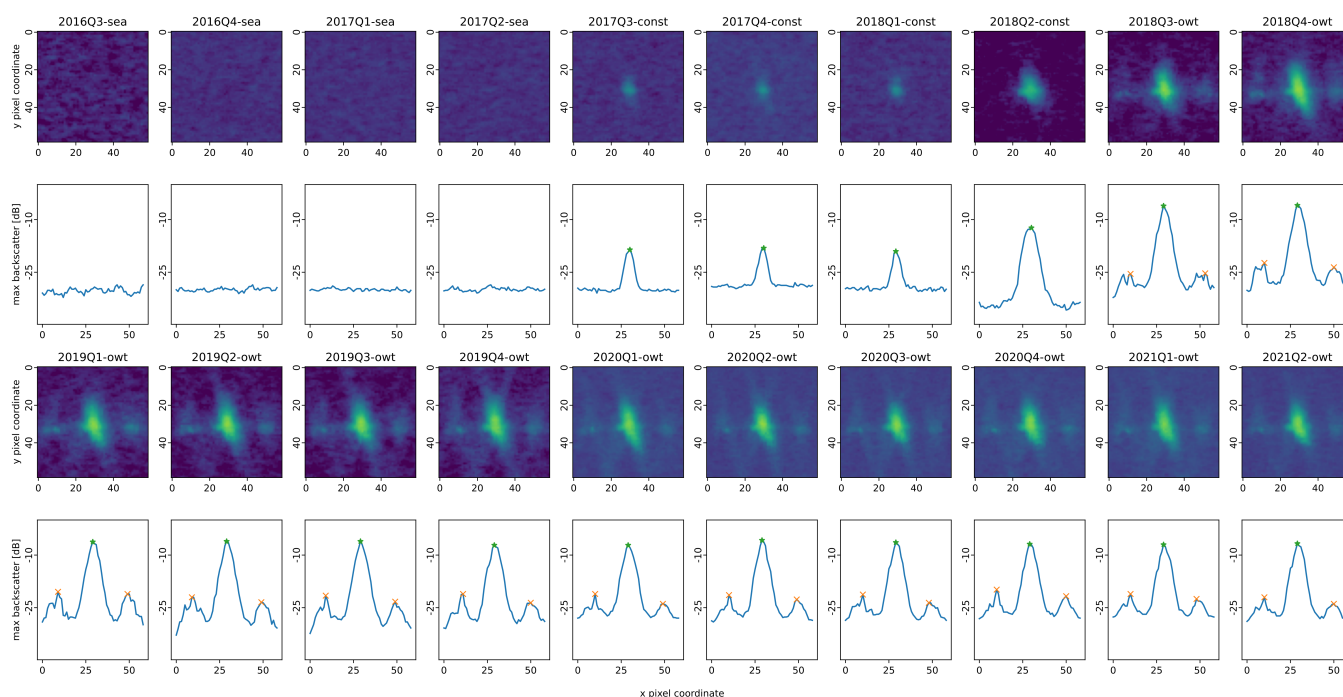## 4.3   Time series analysis for deployment dynamics and location refinement

Figure 6. Deployment time series from 2016Q3 until 20201Q2 of an OWT. First and third row show quarterly Sentinel-1 median images of
the same detected bounding box over time. Second and fourth row show the corresponding maximal swath profiles with detected peaks. The
automatically derived construction stages are given for each image: sea = open sea/no turbine, const = under construction, owt = offshore
wind turbine.

In order to describe the temporal dynamics of OWT deployment, a backwards looking time series analysis of the detected
locations in 2021-Q2 was performed. For each bounding box of an object location, a multi temporal stack of 19 quarterly
Sentinel-1 median images was analysed to determine if an image shows open sea, an OWT or OWF substation under con-

215   struction, or an OWT or OWF substation readily deployed. Swath profiles that describe each column's maximum value along
the horizontal axes from each image in the multitemporal stack were generated. Figure 6 shows quarterly images and corre-

sponding swath profiles for an OWT deployment time series. By applying two consecutive peak finder algorithms with a high and low prominence (Virtanen et al., 2020) to the swath profile, the centre peak and adjacent peaks to the left and right of the centre peak were detected. With the analysis of changes in the number of peaks, centre peak amplitude and width by taking the class predicted by the CNN for the 2021Q2 period into account, the 19 preliminary image patches of the time series were automatically classified in the classes mentioned above.

To finally refine the bounding boxes to point locations, all bounding box images of the 2021-Q2 period were also analysed to find the centre peak of the CNN detected object. With the x coordinate and amplitude of the centre peak coming from the swath profile analysis, the column at the x location of the image patch was queried to find the corresponding y coordinate. By taking the geographic origin of the image patch into account, the WGS84 coordinate for this pixel was derived. Thus, the object is no longer described by a bounding box but by an accurate point location that is precisely at the amplitude maximum of the detected object. By merging spatial and temporal processing results, point locations with the classified quarterly time series from 2016-Q3 until 2021-Q2, the final content for the derived data set is completed. Figure 7 provides an overview of all detected objects and their corresponding class of the entire time series in the DeepOWT data set.
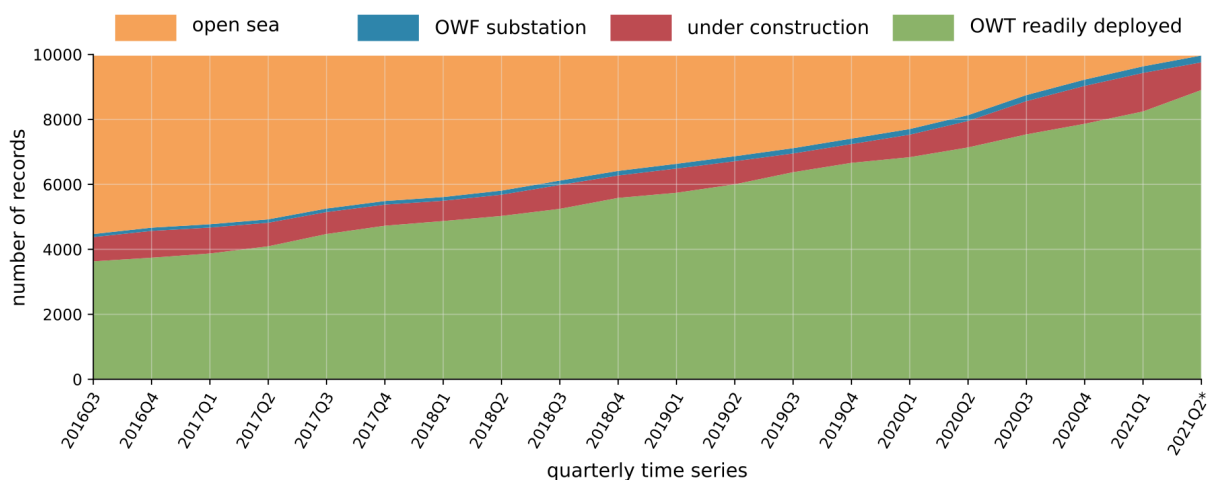

Figure 7. The temporal development of all 9941 objects and their corresponding class in the DeepOWT data set. Class and location in a period with an * were derived by a CNN. In all other periods, the class was derived by the swath profile analysis upon the location detected by the CNN, see figure 6.

5 Ground truth data sets

In order to evaluate the DeepOWT data set, ground truth data sets for two areas were generated. The areas are the North Sea Basin (NSB) and the East China Sea (ECS), pictured in figure 8, which boundaries were aligned with the processing grid in this study. These two areas were chosen due to their importance for offshore wind generation and their differences concerning the underlying Sentinel-1 data and the interaction of coasts, coastal infrastructure, and islands. Where the majority of OWFs in the
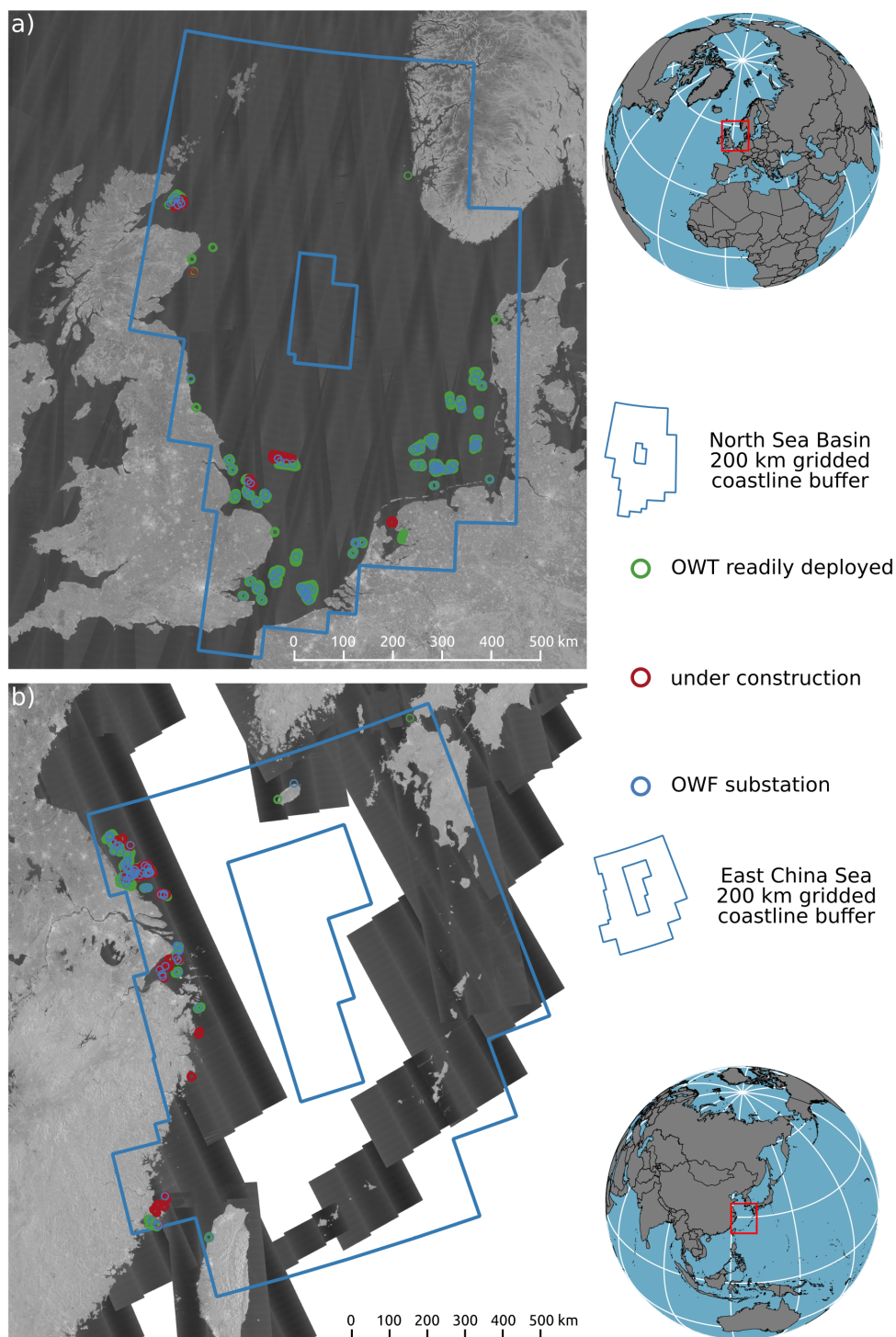
Figure 8. Overview of the ground truth sites boundaries and the hand labelled object locations and deployment stages for the second quarter in 2021. a) shows the North Sea Basin (NSB), b) the East China Sea (ECS).

235 NSB are located within a considerable distance to the coast, small islands and other infrastructure like bridges and harbours, OWFs in the ECS are built close to those features, see figure 5. Furthermore, the NSB has a higher number of quarterly Sentinel-1 acquisitions at both orbits, whereas the ECS is mainly detected at a single orbit and thus has lesser acquisitions at the same time interval, see figure 1. That results in different characteristics of the global Sentinel-1 median image in both areas. Together, the NSB and ECS represent various OWF types and how they appear in their natural environment and are

240 representational ground truth sites.

Two types of ground truth data sets were generated and included in the proposed DeepOWT data set. The first type describes the locations of the target objects at a single point in time. The second type describes quarterly time series of deployment dynamics.

Two data sets of the first type were generated, one for 2021-Q2 and one for 2019-Q4. The 2019-Q4 data set will later be
245 used to compare DeepOWT to OSM and GOWT v1.3 records since the latter ends in 2019. For both ground truth data sets, all locations of OWTs and OWF substations, both readily deployed or under construction, were annotated by hand for the entire NSB and ECS. Afterwards, these point locations were buffered with a radius of 100 m, which is the final area that defines a true positive point location prediction. About 15% of the 2021-Q2 ground truth locations were selected for both the _nsb and the _ecs variant. For them, the entire quarterly time series from 2016-Q3 until 2021-Q2 was annotated by hand. An overview
250 of the different ground truth data sets and their number of annotated target objects is provided in table 1.

Table 1. Overview of all ground truth data sets, their corresponding time stamp and number of objects in each class. * are the numbers for the start in 2016Q3 and end in 20201Q1 of the ground truth time series. † is the number of all hand labelled classes of the entire ground truth time sereis with 19 periods.

Site	Time stamp	Data set name	OWT	Construction	Substation	Open sea	$\sum$ label
NSB	2021Q2	2021Q2_nsb	4,016	253	85	-	4,354
ECS	2021Q2	2021Q2_ecs	2,208	574	62	-	2,844
NSB	2019Q4	2019Q4_nsb	3,571	172	78	-	3,821
ECS	2019Q4	2019Q4_ecs	1,208	214	47	-	1,469
NSB	2016Q3-2021Q1	16Q3-21Q1_nsb	352-583*	59-47*	12-19*	227-1*	12,350†
ECS	2016Q3-2021Q1	16Q3-21Q1_ecs	40-311*	12-87*	3-11*	375-21*	8,170†

During manual ground truth data labelling, all locations and time steps were visually examined and cross checked with different data sources. Therefore, Sentinel-1 images were investigated in combination with additional RGB images from Sentinel-2 and Google Earth and public information concerning the deployment dynamics provided by official planning documents, OWF operators and news portals.

Earth System

Open Access

Science

Data

Discussions

255 ## 6 Technical description

DeepOWT (Global Offshore Wind Turbines derived with Deep Learning) is introduced with version number 1.21.2, where the first number is increased when significant changes are made to the methodology used to detect the OWF and OWT objects automatically or additional content is appended to the data set. The second and third numbers describe the year and quarter of the latest entry included in the data set. Hereby, methodological and temporal information is incorporated into the data set

260 name and versioning.

Table 2. Overview of the metadata of all files included in the DeepOWT data set. The data set is available as repository on Zenodo, see Hoeser and Kuenzer (2022).

File	Extent	Time stamp	Periods	Geometry	OWT	const	sub	sea	Entries
DeepOWT.geojson	global	2016Q3-2021Q2	20	points	✓	✓	✓	✓	9,941
gt_2021Q2_nsb.geojson	NSB	2021Q2	1	polygons	✓	✓	✓	✗	4,354
gt_2021Q2_ecs.geojson	ECS	2021Q2	1	polygons	✓	✓	✓	✗	2,844
gt_2019Q4_nsb.geojson	NSB	2019Q4	1	polygons	✓	✓	✓	✗	3,821
gt_2019Q4_ecs.geojson	ECS	2019Q4	1	polygons	✓	✓	✓	✗	1,469
gt_2016Q3-2021Q1_nsb.geojson	NSB	2016Q3-2021Q1	19	polygons	✓	✓	✓	✓	650
gt_2016Q3-2021Q1_ecs.geojson	ECS	2016Q3-2021Q1	19	polygons	✓	✓	✓	✓	430
gt_nsb_gridded.geojson	NSB	-	-	polygons	✗	✗	✗	✗	1
gt_ecs_gridded.geojson	ECS	-	-	polygons	✗	✗	✗	✗	1

The data set contains the automatically derived target location and the earlier described hand labelled ground truth data sets. Table 2 provides an overview of which file of the data set contains which information. DeepOWT 1.21.2 describes the deployment stages of OWT and OWF substations on a global scale from 2016-Q3 until 2021-Q2. Each entry holds the information of the deployment state within three months of the respective quarter in a year. 9,737 OWTs were detected for

265 the second quarter of 2021. Of these, 8,885 were readily deployed, and 852 were under construction. Additionally, 204 OWF substations were detected for the same period. The file size of DeepOWT is 4.1 MB.

All automatically detected and hand labelled objects are described as points or polygons using the geographic coordinate system WGS84. The spatial geometries were checked topologically to identify duplicate entries, even if no topological errors were found during this inspection. The checked geometries are stored in .geojson files along with the temporal deployment

270 information in the corresponding attribute table. The quarterly periods of the time series in the attribute table of a .geojson file are named in the format $Y_{yyyy}Q_q$, where $yyyy$ is the year and q the quarter of a year. For each location and time series entry, the object class or deployment stage is described as an integer between 0 and 3. Table 3 provides the mapping of the corresponding semantic labels.

Table 3. Mapping of the key in integer format, used in the data set files, to semantic labels and their abbreviations.

Data set key	Semantic label	Abbreviation
0	open sea	sea
1	under construction	const
2	offshore wind turbine	owt
3	offshore wind farm substation	sub

7 Data set evaluation and comparison

275   The evaluation was performed on two ground truth sites, the North Sea Basin (NSB) and the East China Sea (ECS), see section 5. In order to provide consistent and comparable metrics which take the different numbers of objects for each site into account, first, the metrics were calculated separately. Thereby, a predicted object is considered a true positive (TP) when it is within the polygon, a 100 m radius around the object centre, of a ground label and both have the same class, otherwise, it is a false positive (FP). A wrongly omitted ground truth area is considered a false omission (FO). Upon that, precision Pr and recall Rc

280   were calculated:

$$Pr = \frac{TP}{TP + FP} \tag{2}$$

$$Rc = \frac{TP}{TP + FO} \tag{3}$$

The harmonic mean of Pr and Rc, summarises both metrics as the F1 score:

$$F1 = 2 \times \frac{Pr \times Rc}{Pr + Rc} \tag{4}$$

285   Furthermore, all detections from a CNN with a prediction score were sorted in descending order and indexed with $c = 1, 2, ..., C$. From this ordered list, an all point interpolated precision-recall curve Pr_{interp} is generated.

$$Pr_{interp} = \max_{\tilde{Rc}\,:\,\tilde{Rc} \geq Rc} Pr_{interp}(\tilde{Rc}) \tag{5}$$

Finally, the area under the all-point interpolated precision-recall curve describes the average precision AP (Padilla et al., 2021),

290    $$AP = \sum_{c=1}^{C} (Rc(c) - Rc(c-1)) \times Pr_{interp}(Rc(c)), \tag{6}$$

with $\mathrm{Rc}(0) = 0$.

In order to report overall metrics for all n sites or periods of a time series, the separately calculated metrics were combined by macro averaging:

$$\mathrm{Pr}_{\mathrm{avg}} = \frac{1}{n} \sum_{i=1}^{n} \mathrm{Pr}_i \tag{7}$$

$$\mathrm{Rc}_{\mathrm{avg}} = \frac{1}{n} \sum_{i=1}^{n} \mathrm{Rc}_i \tag{8}$$

$$\mathrm{F1}_{\mathrm{avg}} = \frac{1}{n} \sum_{i=1}^{n} \mathrm{F1}_i \tag{9}$$

Thus the macro $\mathrm{F1}_{\mathrm{avg}}$ score is defined as the arithmetic mean over harmonic means following (Opitz and Burst, 2021).

7.1 Data set evaluation

Table 4. Overview of all calculated metrics for the CNN cascade detections on the 2021Q2 global Sentinel-1 median image for each class separately. The detections were evaluated with the 2021Q2_nsb (North Sea Basin) and 2021Q2_ecs (East China Sea) ground truth data sets.

Site	Class	GT	TP	FP	FO	Pr	Rc	F1	AP
North Sea Basin	OWT	4,016	3,996	1	20	1.0	0.995	0.997	0.995
North Sea Basin	under construction	253	186	7	67	0.964	0.735	0.834	0.72
North Sea Basin	OWF substation	85	74	2	11	0.974	0.871	0.919	0.859
East China Sea	OWT	2,208	2,168	16	40	0.993	0.982	0.987	0.981
East China Sea	under construction	574	393	7	181	0.982	0.685	0.807	0.678
East China Sea	OWF substation	62	56	6	6	0.903	0.903	0.903	0.853
						$\mathrm{Pr}_{\mathrm{avg}}$	$\mathrm{Rc}_{\mathrm{avg}}$	$\mathrm{F1}_{\mathrm{avg}}$	
Combined	OWT	6,224	6,164	17	60	0.996	0.988	0.992	
Combined	under construction	827	579	14	248	0.973	0.71	0.821	
Combined	OWF substation	147	130	8	17	0.938	0.887	0.911	

Table 4 and figure 9 summarise the evaluation of the latest period (2021-Q2) of the data set, which was derived by CNNs. The results show that the performance is stable across both study sites. Thus, the CNN model trained on synthetic data can handle both test site characteristics equally well despite the more challenging conditions in the ECS. In general, OWT detection is of

the highest quality with a $F1_{avg}$ score of 99.2%. Furthermore, the $F1_{avg}$ score for OWF substations is at 91.1%. OWT under construction appear to be the most challenging class and can be explained as follows. An OWT under construction's first real world appearance can be restrained in median images when their onset is at the end of a quarter, and thus, just a small amount

305 of acquisitions are part of the quarterly stack. Also, due to the unspecific spatial pattern of single OWT under construction, they are falsely rejected since they are similar to small islands and other persistent marine infrastructure. This interpretation is supported by the PR-curve in figure 9, which clearly shows that when an OWT under construction is recognised, it always is a true positive resulting in high precision. However, the PR-curve drops sharply at a very high precision level around a recall level of 0.7. That indicates false omissions, which can also be seen in table 4. However, the class OWT under construction still
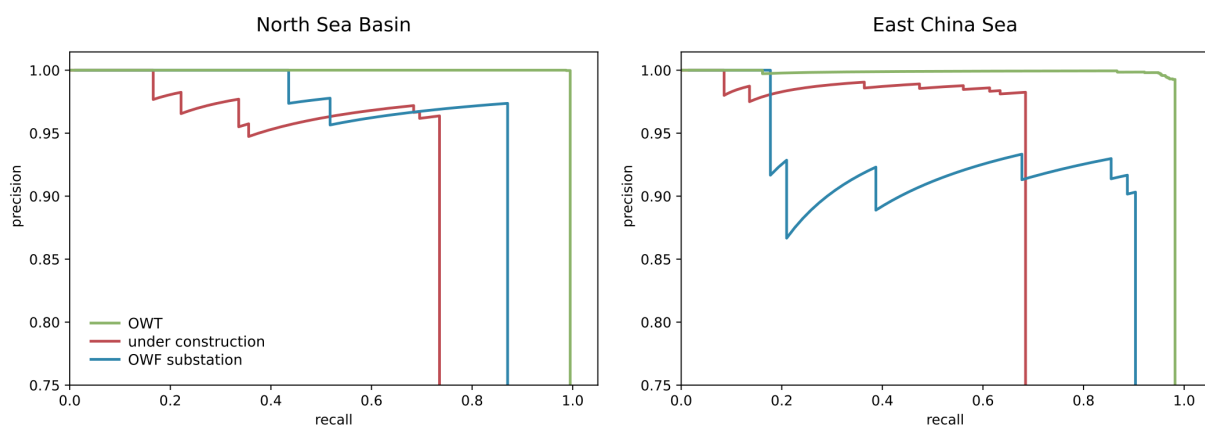
310 has a $F1_{avg}$ score of 82%.

Figure 9. Precision recall curves for the CNN cascade detections on the 2021Q2 global Sentinel-1 median image for each class separately. The detections were evaluated with the 2021Q2_nsb (North Sea Basin) and 2021Q2_ecs (East China Sea) ground truth data sets. The AP values from table 4 are the corresponding areas under the interpolated precision recall curves.

Figure 10 shows the results of the time series evaluation. For each class and period, the F1 scores were calculated. The boxplots on the right side show their distribution and the $F1_{avg}$ over the entire time series. The combined assessment reports the $F1_{avg}$ in each period by averaging the corresponding results from the two ground truth sites. The results show that the time series analysis performs equally well on both sites like the CNN. OWT have a $F1_{avg}$ of 98.1%, OWF substations of 97.6% and

315 OWT under construction of 81%.

7.2 Data set comparison

As reported in section 2, two openly accessible data sets exist, which describe OWT locations, these are OWT records in the OSM database and the GOWT v1.3 data set (Zhang et al., 2021). In order to compare DeepOWT to both data sets, all available entries until 2019Q4 were chosen. Also, for OSM, only such entries were queried, which have an entry date until 31.12.2019.

320 For all 2019 subsets, the evaluation metrics were calculated on the 2019Q4_nsb and 2019Q4_ecs ground truth data sets.
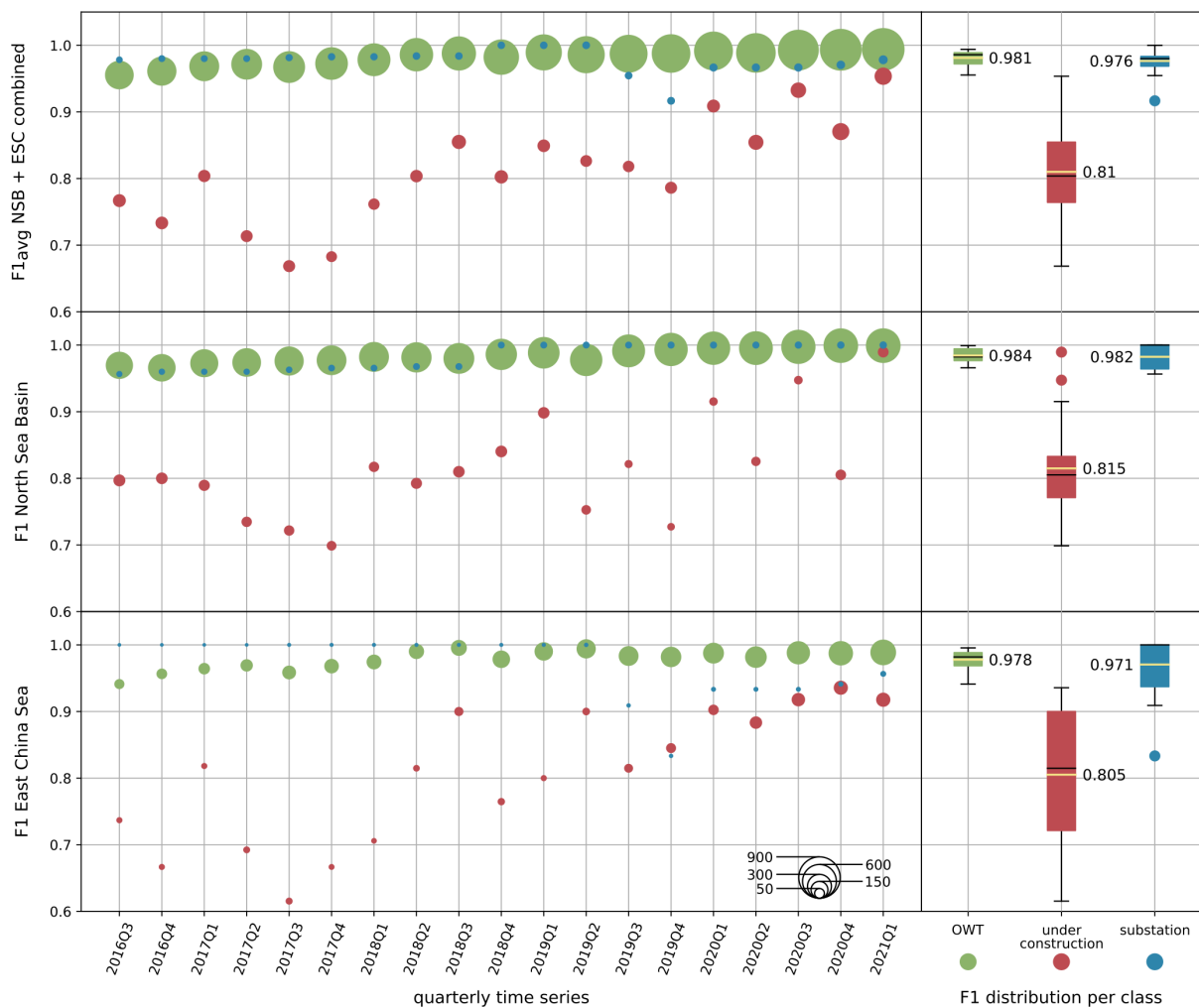
Figure 10. F1 scores for each period pictured as points, where the point size describes the number of ground truth labels. On the right side, boxplots describe the F1 time sereis for each class and provide the temporal $F1_{avg}$ over the entire time sereis. The data was evaluated on the 16Q3-21Q1_nsb (North Sea Basin) and 16Q3-21Q1_ecs (East China Sea) ground truth data sets.

Since OSM only describes locations of readily deployed OWT, only the class 'owt' from DeepOWT was assessed. The comparison in figure 11 shows a consistently better performance of DeepOWT compared to entries of the OSM database. It also becomes clear that the availability of OWT entries in the OSM database differs significantly between the two ground truth sites. In comparison, the remote sensing data derived OWT locations in DeepOWT are independent of their spatial location, and the data set performs equally well on both sites. This clearly shows the advantages of remote sensing data based OWT detection on a global scale.
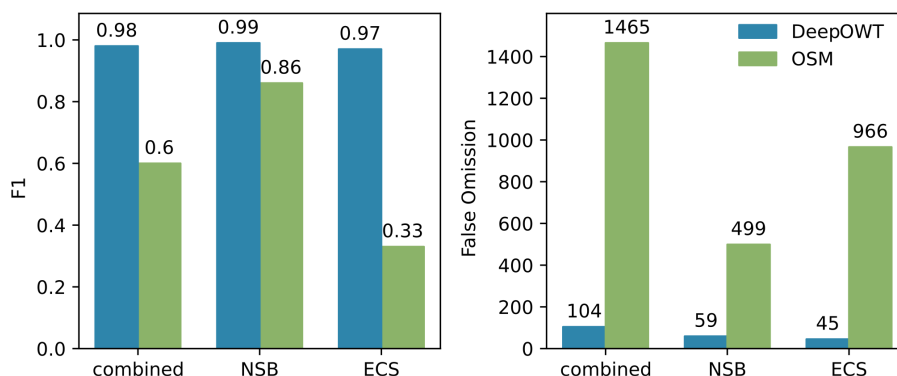
Figure 11. Comparision of the F1 score and false omissions from the OSM and DeepOWT data sets for readily deployed OWT locations in 2019Q4. The data sets were evaluated with the 2019Q4_nsb (North Sea Basin) and 2019Q4_ecs (East China Sea) ground truth data sets.

GOWT v1.3 contains OWT locations beginning at the time when the construction of a turbine foundation was started. However, it does not distinguish between OWTs under construction and those completed. Therefore, the 'under construction' and 'owt' classes of DeepOWT were combined to compare with GOWT v1.3 records. The result in figure 12 shows that both

330   remote sensing based data sets perform consistently on both sites. Nevertheless, this study's deep learning-based DeepOWT data set performs better than GOWT v1.3, which has been derived by applying a handcrafted morphologic approach for OWT detection.
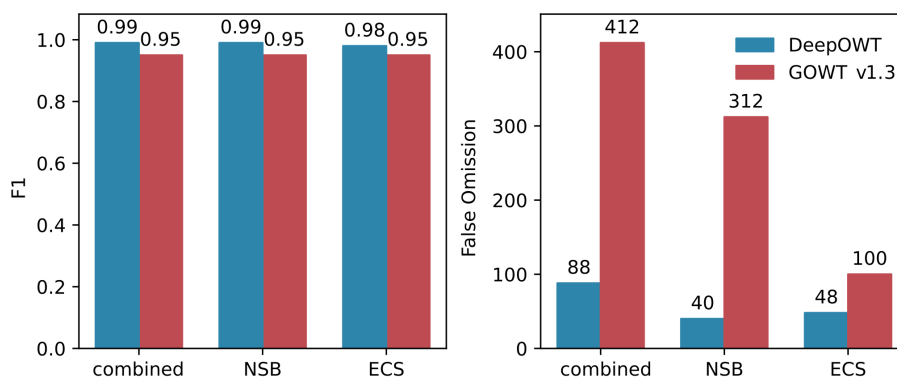

Figure 12. Comparision of the F1 score and false omissions from the GOWT v1.3 and DeepOWT data sets for readily deployed OWT and OWT under construction in 2019Q4. The data sets were evaluated with the 2019Q4_nsb (North Sea Basin) and 2019Q4_ecs (East China Sea) ground truth data sets.

21

8   Potential data set applications

Due to the increasing expansion of OWFs at existing and recently newly developed sites for wind energy production, a holistic
335   understanding, as well as detailed insights of the expansion process, are gaining importance (Fox et al., 2006; Gușatu et al., 2021). The proposed DeepOWT data set enables all stakeholders to access OWT deployment time series globally. Thereby, the division in the pre-, meanwhile, and post-construction phases of the deployment process is particularly important. OWT operators can use this information to develop optimisation measures for the necessary construction process with a specific focus on environmental conditions. Furthermore, single turbine locations on a global scale and under different conditions
340   enable operators to investigate OWFs beyond their own facilities to optimise efficiency during the energy production phase and make better location decisions.

Since OWF often expand in areas that are already used as fishing grounds, shipping routes or are, to some extent, restricted areas like nature reserves or military exclusion zones, potential conflicts have to be recognised early and solved by integrated spatial planning towards multi-use concepts of marine space (Wever et al., 2015; Gusatu et al., 2020). DeepOWT supports the
345   investigation and documentation of past and present OWF projects and potential conflicts to apply the insights to upcoming projects at an early planning phase.

The ecological impacts of OWFs are manifold and have to be differentiated on a spatiotemporal scale (Drewitt and Langston, 2006; Wilson and Elliott, 2009; Bailey et al., 2014; Bergström et al., 2014; Slavik et al., 2019). Thus the spatially contextualised deployment time series in DeepOWT are an important data source for investigations of habitats and migration routes of marine
350   wildlife. For policymakers, DeepOWT offers the opportunity to quickly overview OWF expansion and compare trends on multiple scales in space and time (Rodrigues et al., 2015). Finally, DeepOWT can be used as a database to foster the exchange and transfer of knowledge between these different stakeholders, which was found to be of high importance (Henderson et al., 2003; Fox et al., 2006; Wever et al., 2015; Gusatu et al., 2020).

From a technical perspective, DeepOWT offers a direct integration in analysis made with GIS software and spatial databases
355   (Cavazzi and Dutton, 2016; Gusatu et al., 2020). The lightweight file size and structure of OWT locations, which summarises petabytes of underlying remote sensing images, enables fast processing even on mobile devices. Thus the DeepOWT data set can be used in computational heavy GIS analysis as well as in field campaigns to enrich on-site mapping information.

9   Conclusions

This study introduced DeepOWT, the first openly accessible data set that provides offshore wind turbine (OWT) locations
360   along with their quarterly deployment stages on a global scale. The underlying data originates from the spaceborne C-band radar mission Sentinel-1, from which all acquisitions between July 2016 and June 2021 of the global coastline were used to build quarterly median images. The latest median images from 2021 were investigated using deep learning-based object detection. A cascade of two convolutional neural networks subsequently detects potential offshore wind farm (OWF) locations and single OWT and OWF substations (e.g. transformer stations). The two CNNs were trained entirely on synthetic training
365   data generated by using the novel SyntEO approach for synthetic data generation in Earth observation. Based on the detections

of the CNN cascade, a quarterly time series was derived, which describes the entire deployment dynamics for every single OWT between 2016 and 2021.

The data set covers 8,885 OWTs, 852 platforms under construction and 204 OWF substations for the latest period, the second quarter of 2021. The majority of OWTs are located in the North Sea Basin and the East and South China Sea. With equally
370 good performance on the two ground truth data sets in the North Sea Basin and East China Sea, the quality of the data set is consistent over time and securely describes large OWFs far off the coast as well as small OWFs in a complex near coast environment. The DeepOWT data set contains nine .geojson files, one with the predicted offshore wind energy infrastructure locations and deployment time series, and eight additional files which describe the ground truth data. With a file size of 4.1 MB, DeepOWT is easily portable and ready to use in GIS software.

375 DeepOWT contributes to a holistic understanding as well as detailed insights into the ongoing development of offshore wind energy production, which is at the beginning of a massive increase in OWTs and the development of new OWF sites on a global scale. Furthermore, DeepOWT proves the possibility of automatically detecting small scale objects within large Earth observation archives of radar acquisitions without using auxiliary data by applying state of the art deep learning methods. With the continuation of the Sentinel-1 mission secured, a future detection of OWT deployment time series on a global scale is
380 possible.

10 Data availability

The DeepOWT data set is freely available at https://doi.org/10.5281/zenodo.5933967 (Hoeser and Kuenzer, 2022). To get an overview of the data set, the derived boundaries of the offshore wind energy infrastructure objects along with their temporal deployment dynamics for every single object are presented in the Coastal Explorer (https://coastalx.eoc.dlr.de/) made with the
385 UKIS frontend library (Boeck et al., 2022).

Author contributions. TH designed the study, labelled the ground truth data, developed and implemented the code for data processing, visualisation and evaluation, and prepared the original manuscript, including figures. SF technically implemented the data visualisation in the web mapping service. CK supervised the study, gave suggestions for figures and repeatedly commented and discussed the manuscript.

References

4C Offshore: 4C Offshorewind, https://map.4coffshore.com/offshorewind/, 2021.

395   Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D. G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., Wicke, M., Yu, Y., and Zheng, X.: TensorFlow: A System for Large-Scale Machine Learning, in: 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16), pp. 265–283, USENIX Association, Savannah, GA, https://www.usenix.org/conference/osdi16/technical-sessions/presentation/abadi, 2016.

400   Aschbacher, J.: ESA's earth observation strategy and Copernicus, in: Satellite earth observations and their impact on society and policy, pp. 81–86, Springer, Singapore, https://doi.org/10.1007/978-981-10-3713-9_5, 2017.

Bailey, H., Brookes, K. L., and Thompson, P. M.: Assessing environmental impacts of offshore wind farms: lessons learned and recommendations for the future, Aquatic Biosystems, 10, https://doi.org/10.1186/2046-9063-10-8, 2014.

Bergström, L., Kautsky, L., Malm, T., Rosenberg, R., Wahlberg, M., Capetillo, N. Å., and Wilhelmsson, D.: Effects of offshore wind
405   farms on marine wildlife—a generalized impact assessment, Environmental Research Letters, 9, 034 012, https://doi.org/10.1088/1748-9326/9/3/034012, 2014.

Boeck, M., Voinov, S., Keim, S., Volkmann, R., Langbein, M., and Mühlbauer, M.: Frontend Libraries for DLR UKIS (Map) Applications. Version v8.0.1, https://doi.org/10.5281/zenodo.5835895, 2022.

Cavazzi, S. and Dutton, A.: An Offshore Wind Energy Geographic Information System (OWE-GIS) for assessment of the UK's offshore
410   wind energy potential, Renewable Energy, 87, 212–228, https://doi.org/10.1016/j.renene.2015.09.021, 2016.

COP26: Global coal to clean power transition statement, https://ukcop26.org/global-coal-to-clean-power-transition-statement/, 2021.

Drewitt, A. L. and Langston, R. H. W.: Assessing the impacts of wind farms on birds, Ibis, 148, 29–42, https://doi.org/10.1111/j.1474-919X.2006.00516.x, 2006.

EC, E. C.: An EU Strategy to harness the potential of offshore renewable energy for a climate neutral future, https://ec.europa.eu/energy/
415   sites/ener/files/offshore_renewable_energy_strategy.pdf, 2020.

Esteban, M. D., Diez, J. J., López, J. S., and Negro, V.: Why offshore wind energy?, Renewable Energy, 36, 444–450, https://doi.org/10.1016/j.renene.2010.07.009, 2011.

Fox, A., Desholm, M., Kahlert, J., Christensen, T. K., and Krag Petersen, I.: Information needs to support environmental impact assessment of the effects of European marine offshore wind farms on birds, Ibis, 148, 129–144, https://doi.org/10.1111/j.1474-919X.2006.00510.x,
420   2006.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., and Moore, R.: Google Earth Engine: Planetary-scale geospatial analysis for everyone, Remote Sensing of Environment, 202, 18–27, https://doi.org/10.1016/j.rse.2017.06.031, big Remotely Sensed Data: tools, applications and experiences, 2017.

Gusatu, L. F., Yamu, C., Zuidema, C., and Faaij, A.: A spatial analysis of the potentials for offshore wind farm locations in the North Sea
425   region: Challenges and opportunities, ISPRS International Journal of Geo-Information, 9, 96, https://doi.org/10.3390/ijgi9020096, 2020.

Gușatu, L., Menegon, S., Depellegrin, D., Zuidema, C., Faaij, A., and Yamu, C.: Spatial and temporal analysis of cumulative environmental effects of offshore wind farms in the North Sea basin, Scientific Reports, 11, https://doi.org/10.1038/s41598-021-89537-1, 2021.

He, K., Zhang, X., Ren, S., and Sun, J.: Deep Residual Learning for Image Recognition, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778, https://doi.org/10.1109/CVPR.2016.90, 2016.

430 Henderson, A. R., Morgan, C., Smith, B., Sørensen, H. C., Barthelmie, R. J., and Boesmans, B.: Offshore Wind Energy in Europe— A Review of the State-of-the-Art, Wind Energy, 6, 35–52, https://doi.org/10.1002/we.82, 2003.

Hoeser, T. and Kuenzer, C.: Object Detection and Image Segmentation with Deep Learning on Earth Observation Data: A Review-Part I: Evolution and Recent Trends, Remote Sensing, 12, https://doi.org/10.3390/rs12101667, 2020.

Hoeser, T. and Kuenzer, C.: SyntEO: Synthetic Dataset Generation for Earth Observation with Deep Learning - Demonstrated for Offshore

435 Wind Farm Detection, CoRR, abs/2112.02829, https://arxiv.org/abs/2112.02829, 2021.

Hoeser, T. and Kuenzer, C.: DeepOWT: A global offshore wind turbine data set, https://doi.org/10.5281/zenodo.5933967, 2022.

Hoeser, T., Bachofer, F., and Kuenzer, C.: Object Detection and Image Segmentation with Deep Learning on Earth Observation Data: A Review—Part II: Applications, Remote Sensing, 12, https://doi.org/10.3390/rs12183053, 2020.

LeCun, Y., Bengio, Y., and Hinton, G.: Deep Learning, Nature, 521, 436–444, https://doi.org/10.1038/nature14539, 2015.

440 Loshchilov, I. and Hutter, F.: SGDR: Stochastic Gradient Descent with Warm Restarts, 2017.

Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., and Johnson, B. A.: Deep learning in remote sensing applications: A meta-analysis and review, ISPRS Journal of Photogrammetry and Remote Sensing, 152, 166 – 177, https://doi.org/10.1016/j.isprsjprs.2019.04.015, 2019.

Opitz, J. and Burst, S.: Macro F1 and Macro F1, 2021.

Padilla, R., Passos, W. L., Dias, T. L. B., Netto, S. L., and da Silva, E. A. B.: A Comparative Analysis of Object Detection Metrics with a

445 Companion Open-Source Toolkit, Electronics, 10, https://doi.org/10.3390/electronics10030279, 2021.

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., and Prabhat: Deep learning and process understanding for data-driven Earth system science, Nature, 566, 195–204, 2019.

Ren, S., He, K., Girshick, R. B., and Sun, J.: Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks, IEEE Transactions on Pattern Analysis and Machine Intelligence, 39, 1137–1149, 2015.

450 Rodrigues, S., Restrepo, C., Kontos, E., Teixeira Pinto, R., and Bauer, P.: Trends of offshore wind projects, Renewable and Sustainable Energy Reviews, 49, 1114–1135, https://doi.org/10.1016/j.rser.2015.04.092, 2015.

Slavik, K., Lemmen, C., Zhang, W., Kerimoglu, O., Klingbeil, K., and Wirtz, K. W.: The large-scale impact of offshore wind farm structures on pelagic primary productivity in the southern North Sea, Hydrobiologia, 845, 35–53, https://doi.org/10.1007/s10750-018-3653-5, 2019.

Torres, R., Snoeij, P., Geudtner, D., Bibby, D., Davidson, M., Attema, E., Potin, P., Rommen, B., Floury, N., Brown, M., Traver, I. N.,

455 Deghaye, P., Duesmann, B., Rosich, B., Miranda, N., Bruno, C., L'Abbate, M., Croci, R., Pietropaolo, A., Huchler, M., and Rostan, F.: GMES Sentinel-1 mission, Remote Sensing of Environment, 120, 9–24, https://doi.org/10.1016/j.rse.2011.05.028, the Sentinel Missions - New Opportunities for Science, 2012.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat,

460 İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and SciPy 1.0 Contributors: SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python, Nature Methods, 17, 261–272, https://doi.org/10.1038/s41592-019-0686-2, 2020.

Wever, L., Krause, G., and Buck, B. H.: Lessons from stakeholder dialogues on marine aquaculture in offshore wind farms: Perceived potentials, constraints and research gaps, Marine Policy, 51, 251–259, https://doi.org/10.1016/j.marpol.2014.08.015, 2015.

465 Wilson, J. C. and Elliott, M.: The habitat-creation potential of offshore wind farms, Wind Energy, 12, 203–212, https://doi.org/10.1002/we.324, 2009.

Wong, B. A., Thomas, C., and Halpin, P.: Automating offshore infrastructure extractions using synthetic aperture radar and Google Earth Engine, Remote Sensing of Environment, 233, 111 412, https://doi.org/10.1016/j.rse.2019.111412, 2019.

Xu, W., Liu, Y., Wu, W., Dong, Y., Lu, W., Liu, Y., Zhao, B., Li, H., and Yang, R.: Proliferation of offshore wind farms in the
470 North Sea and surrounding waters revealed by satellite image time series, Renewable and Sustainable Energy Reviews, 133, 110 167, https://doi.org/10.1016/j.rser.2020.110167, 2020.

Zhang, J., Wang, Q., and Su, F.: Automatic extraction of offshore platforms in single SAR images based on a dual-step-modified model, Sensors, 19, 231, 2019.

Zhang, T., Tian, B., Sengupta, D., Zhang, L., and Si, Y.: Global offshore wind turbine dataset, Scientific Data, 8,
475 https://doi.org/10.1038/s41597-021-00982-z, 2021.

Zhu, X., Montazeri, S., Ali, M., Hua, Y., Wang, Y., Mou, L., Shi, Y., Xu, F., and Bamler, R.: Deep Learning Meets SAR: Concepts, Models, Pitfalls, and Perspectives, IEEE Geoscience and Remote Sensing Magazine, Early Access, 0–0, https://doi.org/10.1109/MGRS.2020.3046356, 2021.

Zhu, X. X., Tuia, D., Mou, L., Xia, G., Zhang, L., Xu, F., and Fraundorfer, F.: Deep Learning in Remote Sensing: A Comprehensive Review
480 and List of Resources, IEEE Geoscience and Remote Sensing Magazine, 5, 8–36, 2017.