

Response to Reviewers' Comments

Dear Editor and Reviewers,

We are grateful to the editor and both reviewers for their careful reading of the manuscript and their constructive comments, which have substantially improved the quality and clarity of the work. We have revised the manuscript thoroughly in response to all comments and provide a detailed, point-by-point reply below.

The major revisions are summarized as follows:

(1) The detection pipeline has been clarified, with ambiguous methodological descriptions revised to better reflect the workflow implemented in the study.

(2) A comprehensive parameter sensitivity and uncertainty analysis has been added, providing quantitative justification for the threshold choices adopted across all study regions and propagating the associated uncertainties to the temporal trend figures as error bars.

(3) The discussion of the applicability of the dataset has been expanded to clarify the intended use cases, acknowledge detection limitations in challenging environments, and provide guidance to users on the interpretation of temporal attributes.

(4) Several figures have been improved to enhance readability and clarity.

(5) Minor textual revisions, figure updates, and terminology standardization have been made throughout the manuscript.

All changes to the manuscript are indicated in the tracked changes version, where revised or newly added text is highlighted in **blue** and deleted text is shown in **red**. Throughout the point-by-point response below, reviewer comments are presented in **bold**, and excerpts from the revised manuscript are shown in frame.

Author comment to anonymous Referee 1

Thanks for this work, try to develop a robust algorithm to generate global offshore oil and gas datasets from satellite observations. The methodology and results look reliable and reasonable. I have several comments for your reference.

We sincerely thank you for the positive assessment of our work and constructive comments. The reviewer's comments have prompted us to substantially improve the manuscript in several key aspects, including clarifying the detection pipeline and adding a comprehensive parameter sensitivity and uncertainty analysis. We have carefully addressed each comment point-by-point below, and we hope that the revised manuscript meets the expectations of the reviewer and the journal.

1. Fig. 3, why do you only use Sentinel-2 satellite imagers during Jan-Mar for OOGP generation?

Thank you for pointing this out. We apologize for any confusion caused by this. The original Fig. 3 in the manuscript contained a display issue in the x-axis labeling, which incorrectly suggested that only January-March Sentinel-2 images were used. In reality, Sentinel-2 imagery covering the full study period (2017-2023) was used, consistent with the description in Section 2.2. The figure has been corrected in the revised manuscript to accurately reflect the complete temporal coverage of the Sentinel-1 and Sentinel-2 datasets.

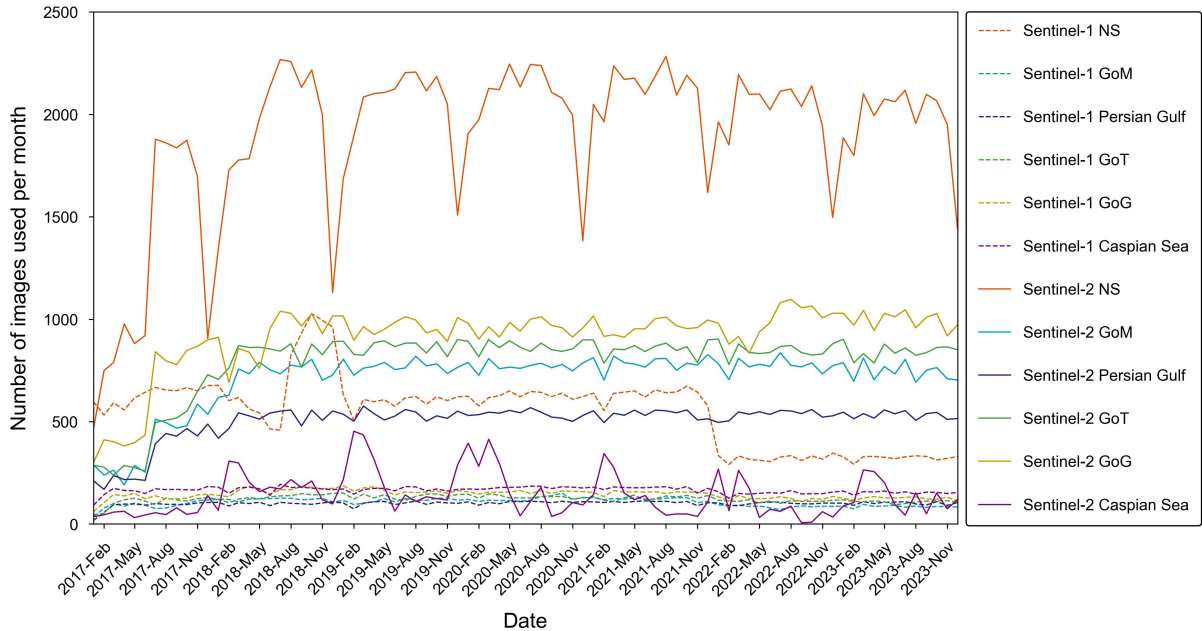


Figure 1: Statistics of the number of Sentinel-1 and Sentinel-2 images used in this study.

2. L155-160, the FIRMS with 1-km spatial resolution, Near-real time frequency was used to identify platform activity status. This is an interesting point and skillful solution. The key challenge is how to address the discrepancy in spatiotemporal resolutions between

this fire data and Sentinel-1/2 data (with 10-60 m resolution). Please add more details on this point. For example, how is proximity defined in the proposed approach?

We clarify that FIRMS data were used exclusively for spatial validation of platform activity status, not for attribute generation. The spatial resolution discrepancy between FIRMS and the primary datasets was addressed through a vector-based spatial overlay. Annual FIRMS hotspot point clouds were aggregated into cluster polygons representing areas of persistent thermal activity, and OOGPs whose centroids fell within these polygons were considered spatially validated. This polygon-based approach does not require pixel-level resolution matching. All detections were further confirmed through multi-source visual interpretation using Sentinel-2 MSI imagery, Google Earth high-resolution images, and authoritative databases (NOAA, BSEE, OGIM, and OESNS), as detailed in Section 4.1.1. To address this point, we have revised the relevant sentences in Section 2.3.2 to explicitly describe the spatial aggregation approach and clarify how the resolution discrepancy was handled. The revised text is as follows:

L167-L173: Annual FIRMS hotspot point clouds were spatially aggregated into cluster polygons representing areas of persistent thermal activity, and OOGPs whose centroids fell within these polygons were considered spatially validated as active platforms. This vector-based spatial overlay approach is inherently robust to the resolution discrepancy between FIRMS and the primary satellite datasets, as it operates on vector geometries rather than requiring pixel-level resolution matching. By overlaying these thermal cluster polygons and dense ship lane polygons onto satellite-detected platform locations, the auxiliary datasets ensure the accuracy of platform identification and distinguish operational platforms from other structures or false positives. All detections were further confirmed through multi-source visual interpretation as described in Section 4.1.1.

3. L160, as mentioned above, FIRMS has a different spatiotemporal resolution from the primary satellite data (Sentinel-1/2). Please clarify this point.

Annual FIRMS hotspot point clouds were aggregated into cluster polygons representing areas of persistent thermal activity, and OOGPs whose centroids fell within these polygons were considered spatially validated. This approach is inherently robust to resolution differences, as it operates on vector geometries rather than raster pixels. A clarifying sentence has been added to Section 2.3.2 of the revised manuscript. The revised text in Section 2.3.2 is provided above.

4. Eq.(1), although two local examples are given to demonstrate the rationale regarding the BC_{max} threshold of -20db, an adaptive threshold like Eq. (2) is more reliable.

We thank the reviewer for this constructive suggestion. We agree that presenting a fixed absolute BC_{max} threshold (original Eq. 1) alongside the adaptive percentile threshold (Eq. 2) created unnecessary ambiguity regarding the actual detection workflow. We wish to clarify that

the fixed threshold was considered during early method development but was ultimately not adopted in the final detection pipeline of this study. To avoid confusion, the corresponding text and Eq. 1 have been removed from Section 3.1 of the revised manuscript. The adaptive percentile threshold, formerly Eq. 2, has been renumbered as Eq. 1 in the revised manuscript.

To provide the quantitative justification requested by the reviewer, we conducted a sensitivity analysis of five absolute $BC_{\max}$ thresholds (-15 , -18 , -20 , -22 , and -25 dB) across 12 representative OOGP locations in the GoM (Fig. 2 of this response document, corresponding to Fig. S2 of the revised manuscript). Two key findings were obtained.

First, platform targets consistently produce $BC_{\max}$ values well above all tested thresholds, confirming robust detectability regardless of the exact cutoff value.

Second, the proportion of sea surface background pixels falling below any given fixed threshold varies substantially across regions and seasons, demonstrating that a fixed absolute cutoff lacks the regional adaptability required for consistent OOGP detection.

These findings collectively justify the exclusive use of the adaptive percentile threshold as the initial detection criterion in the revised manuscript. The revised text in Section 3.1 is as follows:

L185-L195: The backscattering coefficients in the VH band of Sentinel-1 effectively distinguish OOGPs from open water. As demonstrated by a sensitivity analysis of absolute $BC_{\max}$ thresholds across 12 representative OOGP locations in the GoM (Fig. S2), fixed absolute cutoffs lack regional adaptability due to substantial variability in sea-surface background backscatter across regions and seasons. Therefore, the detection framework adopts the percentile-based adaptive threshold described in Section 3.2, which dynamically adjusts to the local backscatter distribution of each region and month. The detection framework is based on the physical differences between fixed OOGPs and mobile objects, such as vessels: OOGPs produce stable and temporally persistent backscatter signals, whereas vessels generate decorrelated and spatially variable signals with low Occurrence Frequency (OF) values. Exploiting these differences, platform detection proceeds in two complementary stages. In the first stage, a percentile-based adaptive threshold was applied to identify candidate targets with high completeness. In the second stage, the OF filtering removed the transient signals. Together, the percentile threshold governs candidate set completeness, while the OF filter governs false positive suppression. The theoretical justification and sensitivity analyses for each stage are detailed in Sections 3.2 and 3.3.1 respectively.

5. Fig. 10, how can we control the effect of the uncertainty in satellite data and algorithm (like threshold selection) on these temporal variations?

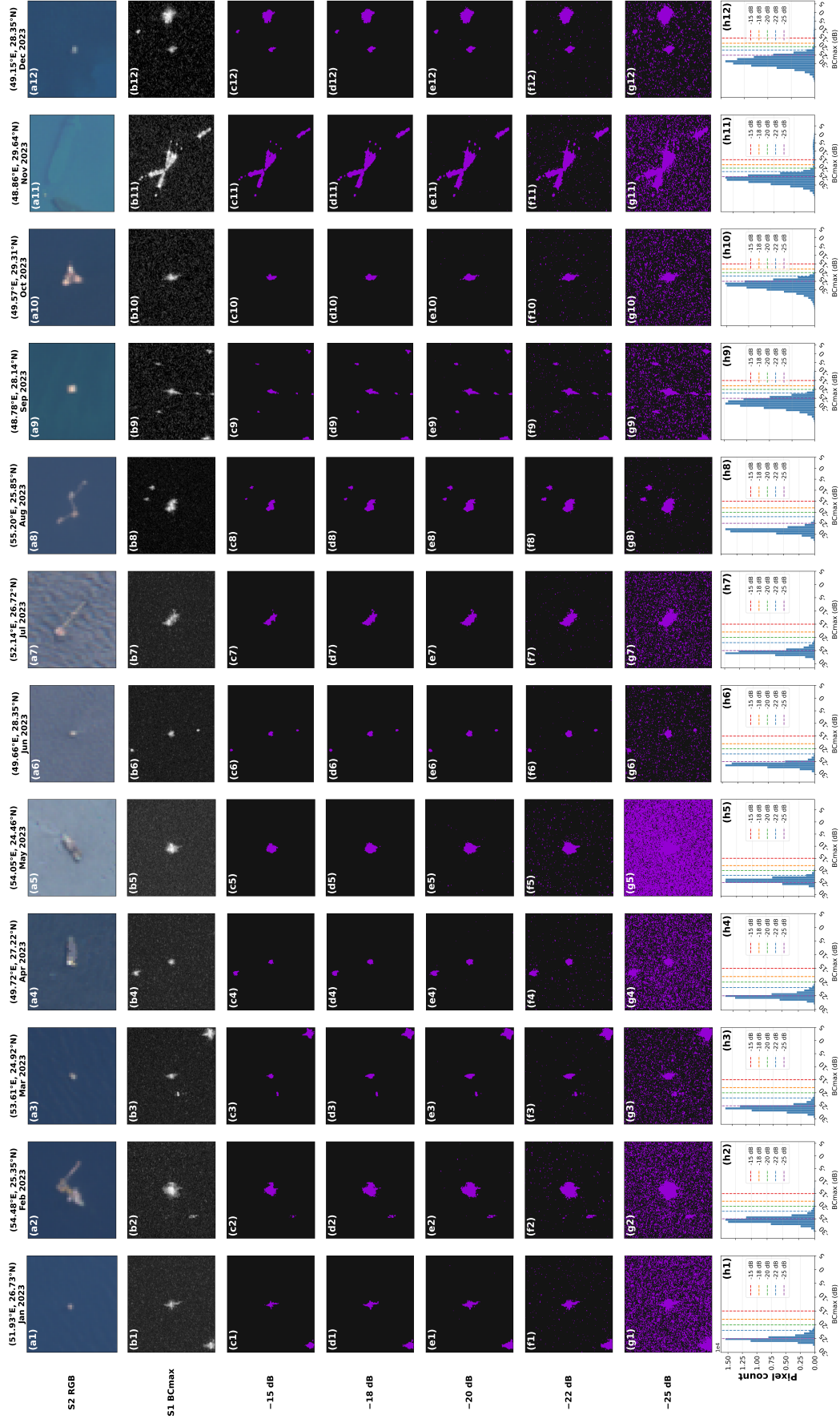


Figure 2: Sensitivity of absolute BC_{max} thresholds at 12 representative OOGP locations in the GoM (2023). Each column corresponds to one platform location, labelled by coordinates and acquisition month. Rows show: (a1-a12) Sentinel-2 true-color imagery using a 200 m buffer; (b1-b12) Sentinel-1 monthly BC_{max} in VH polarization using a 1000 m buffer, shown in grayscale; (c1-g12) binary retention masks under absolute thresholds of -15 , -18 , -20 , -22 , and -25 dB, respectively; and (h1-h12) histograms of BC_{max} pixel values within the 1000 m buffer, with the five candidate thresholds indicated by vertical dashed lines. Purple pixels indicate values exceeding the threshold and are therefore retained, whereas black pixels indicate values below the threshold and are therefore excluded.

As shown in Fig. 10 in the revised manuscript and Fig. 3 in this response letter, we conducted a parameter sensitivity analysis of the region-specific OF thresholds and propagated the results as error bars.

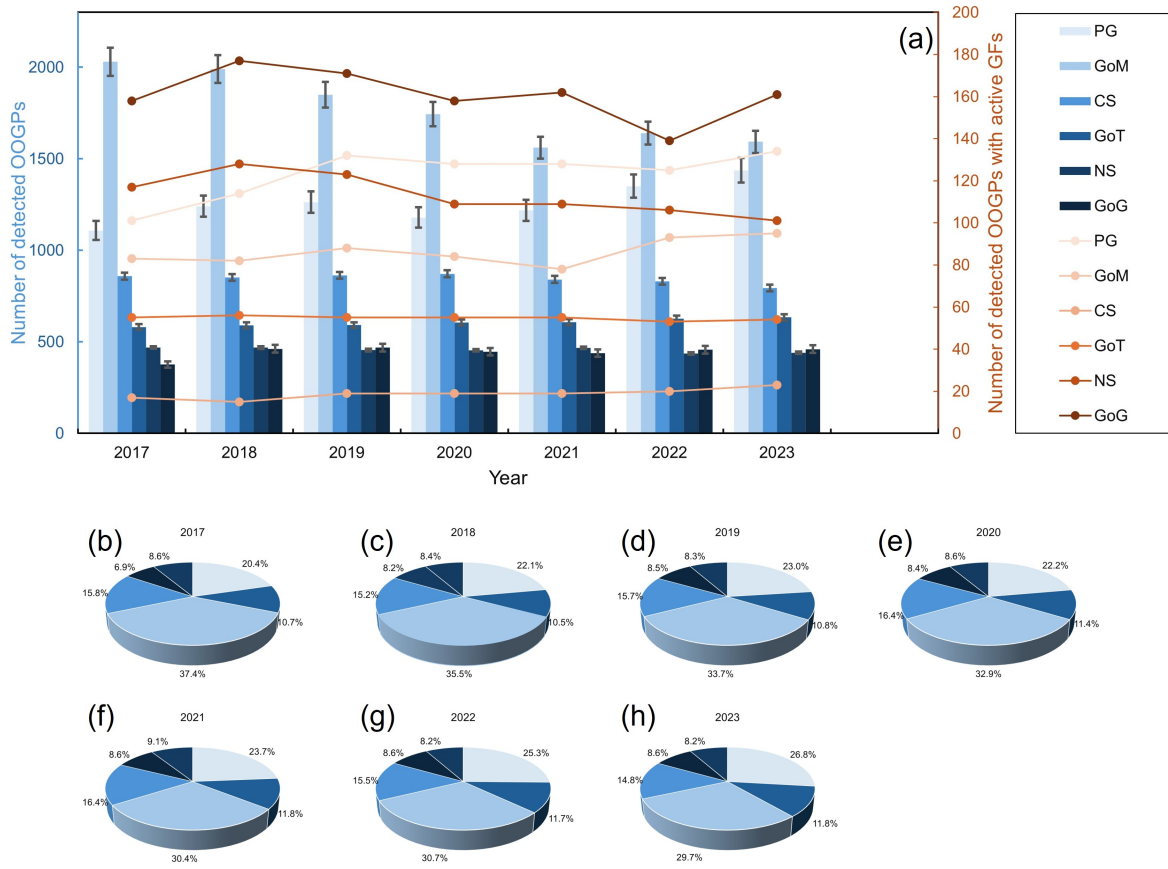


Figure 3: (a) Temporal evolution of detected OOGPs and OOGPs with active gas flarings (GFs) in select regions. (b)–(h) Proportion of each offshore basin based on the available data from 2017 to 2023. Error bars represent algorithmic uncertainty estimated from OF threshold sensitivity analysis ($\sigma_{OF} = 1.6\text{--}4.7\%$; see Supplementary Material, Table S3.)

We varied each basin's OF threshold within ± 2 steps of the adopted value and computed the relative standard deviation of TP counts as the sensitivity index, σ_{OF} . The results are summarized in Table 1 (Table S3 in the Supplementary Material).

Table 1: OF threshold sensitivity analysis results by basin.

Basin	Adopted OF	Test range	σ_{OF}
GoM	3	2 - 5	3.8%
CS	3	2 - 5	2.2%
GoT	2	2 - 4	2.7%
NS	6	4 - 7	1.6%
GoG	2	2 - 4	4.6%
PG	3	2 - 5	4.7%

Error bars are calculated as

$$\delta N_{\text{year}} = N_{\text{year}} \times \sigma_{\text{OF}}$$

The PG increase of 329 platforms far exceeds its ± 68 platform uncertainty and is further corroborated by independent evidence. As shown in Fig. 4 and Fig. 5 (Fig. S10 in the Supplementary Material), new detections are geographically concentrated in the Safaniya, Zuluf, and Marjan fields (Saudi Arabia), the North Dome/South Pars fields (Qatar), and the Upper Zakum fields (UAE), consistent with documented active development (Kaiser, 2022) and reports from the U.S. Energy Information Administration (<https://www.eia.gov/>). In addition, the Middle East accounted for the largest share of global unfinished offshore gas reserves in 2022 (Chen et al., 2024). These results confirm that the temporal trends in Section 4.2 reflect genuine infrastructure changes rather than algorithmic sensitivity.

The sensitivity analysis was based on 2023 validation data applied uniformly across 2017-2023, assuming the temporal stability of σ_{OF} . The results are based on validation samples rather than the full platform population and should therefore be interpreted as approximate estimates. The revised Fig. 10 includes error bars for all basins, and the methodology has been added to the supplementary file.

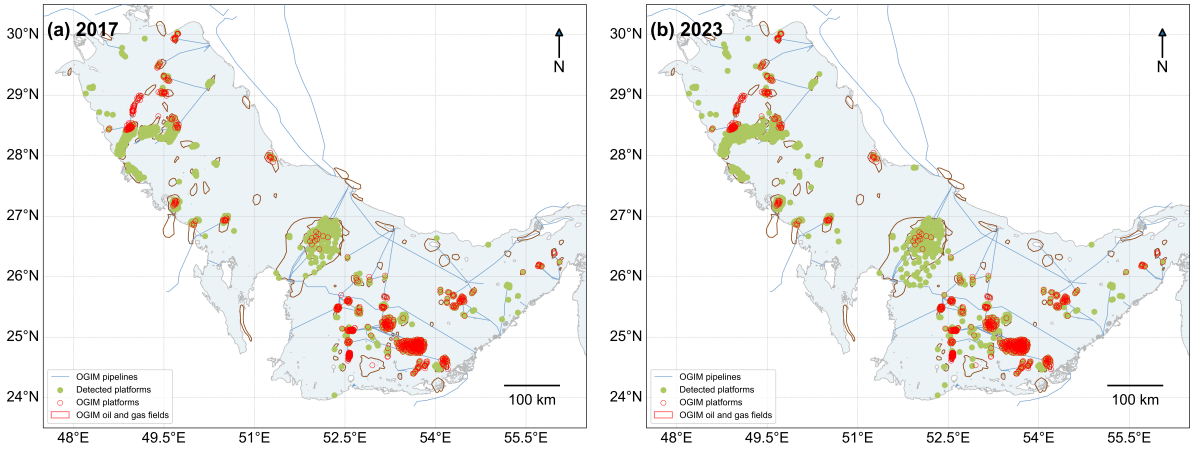


Figure 4: Spatial distribution of detected offshore oil and gas platforms and fields in the Persian Gulf between 2017 and 2023.

L382-L384: The robustness of these trends against algorithmic parameter uncertainty was assessed through an OF threshold sensitivity analysis, the methodology and results of which are provided in the Supplementary Material (Table S3).

Supplementary file: To quantify the sensitivity of the detected platform counts to the adopted OF thresholds, we varied each basin's threshold within 2 steps of the adopted value and computed the relative standard deviation of true positive counts as the sensitivity index σ_{OF} . The results are summarized in Table S3. The error bars in Fig. 10 were calculated as

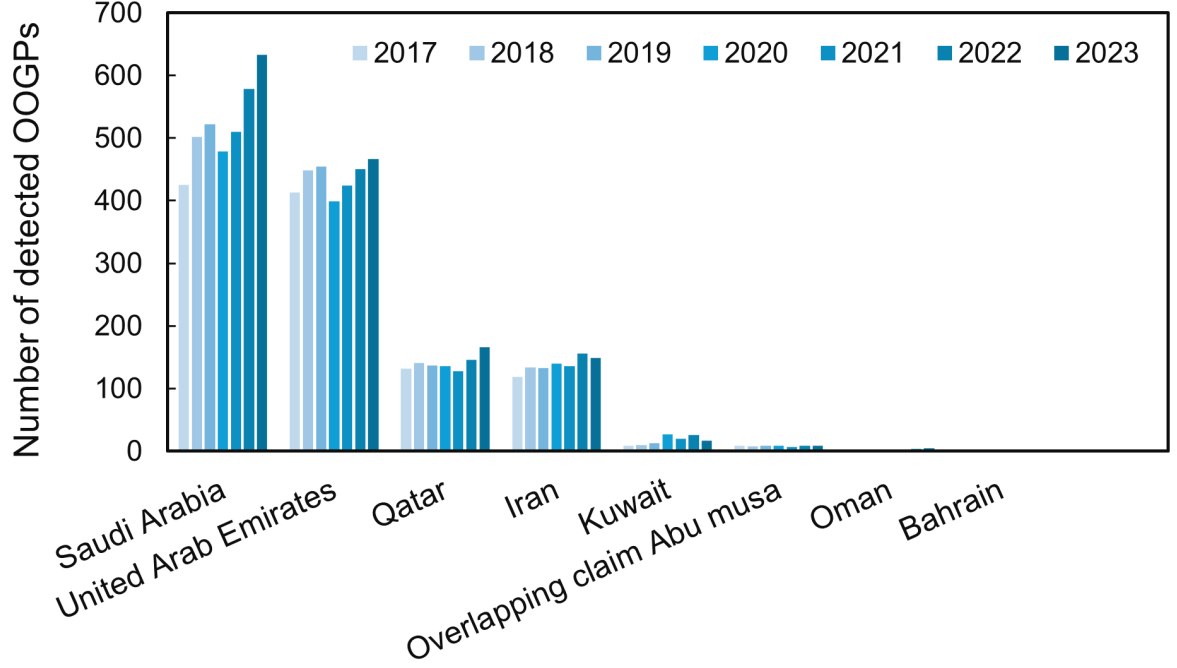


Figure 5: Temporal count changes of detected offshore oil and gas platforms between different countries in the Persian Gulf.

$\delta N_{\text{year}} = N_{\text{year}} \times \sigma_{\text{OF}}$, providing a conservative estimate of the algorithmic uncertainty in the temporal trends.

Author comment to anonymous Referee 2

This manuscript presents a satellite-based framework for detecting offshore oil and gas platforms (OOGPs) using multi-source remote sensing data, primarily Sentinel-1 SAR imagery, and constructs a global dataset covering six major offshore basins from 2017 to 2023. The study addresses an important gap in existing infrastructure inventories by providing a more temporally consistent and spatially comprehensive dataset, with potential applications in environmental monitoring and methane emission attribution. However, in its current form, the manuscript still suffers from methodological ambiguity, insufficient validation rigor, and limited discussion of uncertainty and generalizability.

We sincerely thank Reviewer 2 for the thorough and constructive evaluation of our manuscript. We appreciate the reviewer's recognition of the study's contribution to addressing an important gap in existing offshore infrastructure inventories. The comments prompted substantial improvements to the manuscript, particularly with methodological transparency, validation comprehensiveness, uncertainty quantification, and expanding the discussion on dataset applicability. We hope that the point-by-point responses below demonstrate that all the concerns raised have been carefully considered and fully addressed in the revised manuscript.

Major Revision:

1. The manuscript claims to develop a “robust” and “automated” framework, yet the methodological components (e.g., thresholding, morphological filtering, occurrence frequency filtering) largely rely on established techniques.

We agree that the framework requires region-specific threshold configuration involving manual inspection and expert judgment, and have replaced "automated" with "semi-automated" throughout the revised manuscript.

To substantiate the claim of robustness, we conducted a parameter sensitivity and uncertainty analysis (Section 3.4), varying each region's OF threshold across 2-4 steps from the adopted value. TP counts remained stable across this range in all six basins, with relative standard deviations of 1.6-4.7%, confirming that the primary effect of threshold variation is the suppression or admission of false positives rather than the loss of genuine platform detections. The relevant descriptions in the manuscript have been revised accordingly.

We note that the primary contribution of the framework lies in its operational scalability. By combining adaptive thresholding with temporal persistence filtering and multi-source cross-validation, it enables consistent and annotation-free platform mapping across six basins over seven years without requiring labelled training data or concurrent multi-sensor acquisitions, which are constraints that limit the applicability of supervised and fusion-based approaches at this scale.

2. What is the clear methodological advancement beyond existing SAR-based OOGP detection studies? How does this approach quantitatively outperform prior methods (e.g., deep learning-based or multi-sensor fusion approaches)?

We have revised the Introduction and Discussion sections to articulate the methodological positioning and limitations of the proposed framework relative to existing approaches more clearly.

Deep learning methods have demonstrated considerable progress in SAR target recognition, including convolutional networks for general SAR automatic target recognition (Chen et al.), multiview deep learning frameworks (Pei et al.), and semisupervised generative adversarial approaches (Zhang et al., b). Applied more directly to offshore platform monitoring, VGG16 and VGG19 architectures were evaluated on Sentinel-1 GRD imagery for oil rig recognition in the Campos Basin, reporting mean global accuracies of 86.4% and 84.1%, respectively (Falqueto et al.). However, model performance may degrade when applied to regions not represented in the training data (Spanier et al.), which limits the applicability of deep learning approaches for annotation free, multi-basin, and multiyear mapping without extensive region specific retraining.

Multisensor fusion approaches require the concurrent availability of multiple data streams across all study areas and years. In practice, cloud contamination, sensor availability gaps, and inconsistent acquisition schedules reduce the feasibility of such approaches for temporally consistent basin scale monitoring over long time periods. Similar limitations apply to optical methods (Zhu et al.). The importance of robust all weather observations has also been highlighted by the recently published METER database (Jackson et al.), which identified limited offshore imagery availability as a major constraint for quality control.

We wish to clarify that the primary objective of this study is not to propose a new state-of-the-art detection algorithm, but to construct a comprehensive and openly available spatiotemporal OOGP dataset covering six major offshore basins from 2017 to 2023. The unsupervised SAR framework adopted in this study does not require annotated training data or auxiliary data streams, enabling efficient basin-scale mapping that would be impractical with supervised or multi-sensor approaches requiring extensive data preparation. This makes it practically suited for the specific objective of large-scale and temporally consistent dataset production. Future integration of deep learning and semi supervised approaches remains a promising direction for improving detection completeness, particularly in high density regions such as the GoM and the PG.

L62-L68: While supervised machine learning and deep learning methods have been increasingly applied to offshore platform detection and can achieve high local accuracy when sufficient labelled data are available (Chen et al.; Pei et al.; Zhang et al., b; Ma et al.; Fal-

queto et al.; Zhang et al., a; Spanier et al.), they require extensive annotated training datasets and may experience reduced transferability across regions not represented in the training data (Spanier et al.). Similarly, optical-based methods (Zhu et al.) and approaches using multi-sensor data (Jackson et al.) face operational limitations associated with cloud cover, data availability, and inconsistent acquisition schedules, which reduce their suitability for temporally consistent long-term monitoring across geographically diverse offshore basins.

L440-L445: Although recent supervised machine learning, deep learning, and multi-sensor fusion approaches have demonstrated strong performance for offshore platform detection in specific regions, their applicability to long-term basin-scale monitoring is associated with several practical challenges. Deep learning approaches generally require large annotated training datasets and may experience reduced transferability across regions with different environmental conditions or platform characteristics. Optical and multisensor approaches can be further affected by cloud contamination, inconsistent acquisition schedules, and varying data availability among regions and years.

In comparison, the unsupervised SAR-based framework adopted in this study prioritizes temporal consistency, operational scalability, and reproducibility across geographically diverse offshore basins. Although the method may not always achieve the same level of local optimization as region-specific supervised approaches, its independence from annotated training data and all-weather observational capability make it particularly suitable for constructing long-term, large-scale spatiotemporal OOGP datasets. Future integration of deep learning or multi-sensor information may further improve detection completeness, especially in high-density offshore production regions.

3. The thresholding strategy, including the use of a fixed backscatter cutoff and a percentile-based adaptive threshold, lacks sufficient theoretical justification. It also remains unclear how sensitive the results are to these parameter choices.

We agree that both the backscatter and percentile-based adaptive thresholds require explicit theoretical justification and sensitivity evaluation. We address each of these issues in turn.

This approach is based on the physical differences between fixed offshore platforms and mobile objects such as vessels. Offshore platforms produce stable and temporally persistent backscatter signals, whereas vessels generate decorrelated and spatially variable signals with low OF values. Exploiting these differences, the detection framework adopts a two-stage design. In the first stage, a percentile-based adaptive threshold was applied to monthly Sentinel-1 VH composites to identify candidate bright targets with high completeness, intentionally accepting some false positives from transient objects such as vessels. In the second stage, these transient signals are removed using OF filtering. Together, these two stages function as com-

plementary controls: the percentile threshold governs candidate set completeness, while the OF filter governs false positive suppression.

An earlier version of the manuscript presented a fixed absolute $BC_{\max}$ threshold of -20 dB as a pre-screening criterion. However, this threshold was considered during early development and was not adopted in the final detection pipeline. To eliminate the ambiguity between the fixed and adaptive thresholds, the corresponding text has been removed from Section 3.1 of the revised manuscript. We conducted a sensitivity analysis of five absolute thresholds (-15 , -18 , -20 , -22 , and -25 dB) across 12 representative OOGP locations in the GoM (Fig. 2 of this response document, corresponding to Fig. S2 of the revised manuscript). The results demonstrate that platform targets consistently produce $BC_{\max}$ values well above all tested thresholds. However, the proportion of sea surface background pixels falling below any given fixed threshold varies substantially across regions and seasons, confirming that the percentile-based adaptive threshold (Eq. 1 in the revised manuscript) serves as the initial detection criterion, which dynamically adjusts to the local backscatter distribution for each region and month.

To examine the sensitivity of percentile-based thresholding, we systematically tested three percentile thresholds (85th, 90th, and 95th) on individual monthly Sentinel-1 VH composites. The lower threshold (85th percentile) resulted in increased false positives, particularly in regions with moderate sea surface backscatter, which was largely driven by random backscatter fluctuations and ocean clutter patterns. In contrast, the higher threshold (95th percentile) effectively suppressed noise but also failed to detect smaller or weakly scattering platforms, particularly in turbid waters. We quantitatively evaluated the platform detection accuracy in the NGoM under different threshold conditions using a unified reference dataset constructed from the NOAA, BSEE, and OGIM platform inventories, with spatial deduplication performed using DBSCAN clustering with a 200 m buffer. As summarized in Table S2 in the Supplementary Material, all three thresholds yielded highly comparable performances, with F1 scores ranging from 0.795 to 0.808, demonstrating that the detection algorithm is not overly sensitive to the exact percentile selected. The 90th percentile achieved the optimal balance, providing the highest precision (0.965) within the tested range, and was therefore adopted as the standard threshold throughout the study. Notably, although the 90th percentile threshold retains some transient targets in the initial candidate set, such transient detections are subsequently suppressed by the OF filtering stage, substantially reducing false positives and ensuring that the final dataset predominantly reflects persistent offshore platform locations. The revised text in Section 3.1 is as follows:

L185-L195: The backscattering coefficients in the VH band of Sentinel-1 effectively distinguish offshore platforms from open water. As demonstrated by a sensitivity analysis of absolute $BC_{\max}$ thresholds across 12 representative OOGP locations in the GoM (Fig. S2), fixed absolute cutoffs lack regional adaptability due to substantial variability in sea-surface background backscatter across regions and seasons. Therefore, the detection

framework adopts the percentile-based adaptive threshold described in Section 3.2, which dynamically adjusts to the local backscatter distribution of each region and month. The detection framework is grounded in the physical differences between fixed offshore platforms and mobile objects such as vessels: offshore platforms produce stable and temporally persistent backscatter signals, whereas vessels generate decorrelated and spatially variable signals with low Occurrence Frequency (OF) values. Exploiting these differences, platform detection proceeds in two complementary stages. In the first stage, a percentile-based adaptive threshold is applied to identify candidate targets with high completeness. In the second stage, the OF filtering removes transient signals. Together, the percentile threshold governs candidate set completeness, while the OF filter governs false positive suppression. The theoretical justification and sensitivity analyses for each stage are detailed in Sections 3.2 and 3.3.1 respectively.

4. The regional variability of Occurrence Frequency (OF) thresholds (e.g., OF=12 for the English Channel vs. $OF \geq 2$ for the GoG) is a pragmatic solution to maritime traffic noise. However, the manuscript lacks a quantitative justification for these values. Providing sensitivity analysis results for these thresholds would strengthen the methodological rigor.

We thank the reviewer for this valuable comment. We have conducted a comprehensive OF threshold sensitivity analysis across all study regions, expanded the validation dataset in Table 3 in the manuscript to incorporate a more comprehensive set of reference platforms across all six basins, and standardized and corrected the threshold notation to $OF > n$ throughout the manuscript to ensure consistency with the filter implementation.

The OF filtering step is based on a fundamental physical distinction between offshore platforms and transient maritime objects: platforms are stationary infrastructures that maintain stable radar backscatter over time and thus exhibit persistently high OF values, whereas mobile objects such as vessels appear only intermittently across monthly composites and therefore display low OF values. This temporal persistence property forms the physical basis for all the threshold choices described below.

First, we varied each basin's OF threshold within ± 2 steps of the adopted value and computed the relative standard deviation of TP counts as the sensitivity index, σ_{OF} . The complete sensitivity analysis is shown in Fig. 6 and Table 1 in this response, corresponding to Fig. S5 and Table S3 of the revised manuscript respectively. We also added error bars on Fig. 10 in the revised manuscript.

Second, to provide the quantitative justification requested by the reviewer, we systematically varied the OF threshold of each region across different steps of the adopted value and examined

the resulting precision, recall, and F1 curves (Fig. 6).

The region-specific threshold choices were determined by identifying the OF value at which the F1 score was maximized or reached a stable plateau, balancing the competing effects of false-positive suppression and platform detection completeness. In regions with sparse vessel traffic (GoG and GoT), the adopted threshold corresponded directly to the F1 peak. In regions with higher vessel traffic density (GoM, CS, inner PG, and NS), the adopted threshold was selected within the F1 plateau region, where the marginal gain in precision from further increasing OF no longer justifies the associated reduction in recall. The English Channel represents the most demanding case: false positives are overwhelmingly dominant at low OF values, whereas TP counts remain invariant across the entire tested range, confirming that the elevated threshold is both necessary and sufficient to eliminate vessel contamination without affecting platform detection completeness. The three confirmed platforms were verified using high-resolution Sentinel-2 imagery and vessel traffic management zone data from the UK Hydrographic Office.

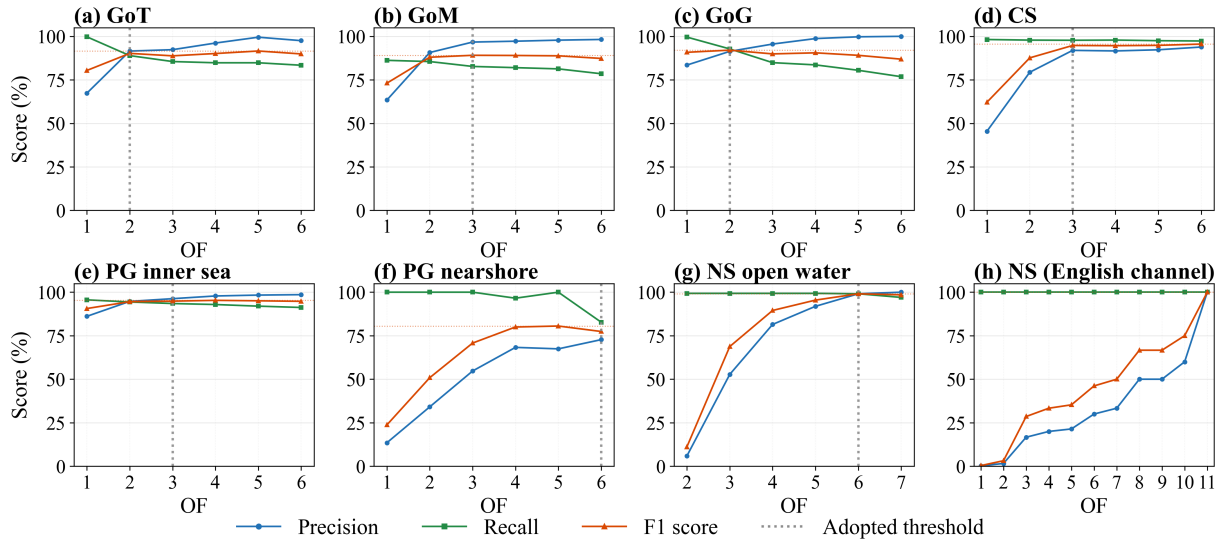


Figure 6: Sensitivity of the occurrence frequency (OF) threshold to platform detection performance across all study regions. Each panel shows the precision (blue), recall (green), and F1 score (orange) as a function of the OF threshold. The vertical dashed line in each panel indicates the adopted OF threshold for the corresponding regions. In all regions, the dominant effect of increasing OF is the suppression of vessel-induced false positives rather than a reduction in true platform detections, supporting the empirical basis of the region-specific threshold selection.

These results collectively demonstrate that the adopted region-specific thresholds are empirically grounded and not arbitrary. An expanded methodological justification has been added to Section 3.3.1 of the revised manuscript.

L234-L240: We applied region-specific OF thresholds determined through sensitivity testing and manual inspection: $OF > 2$ in the GoG and GoT; $OF > 3$ in the GoM, inner PG and CS; $OF > 6$ in the open-water NS and nearshore PG; and $OF = 12$ in the heavily trafficked English Channel (NS). Region-specific OF thresholds were determined through

a systematic sensitivity analysis by varying each threshold within two steps of the adopted value across all study basins. The primary effect of increasing the OF was the progressive suppression of vessel-induced false positives rather than a reduction in true platform detections. The complete sensitivity analysis is presented in Table S3 and Fig. S5.

Supplementary file: To quantify the sensitivity of the detected platform counts to the adopted OF thresholds, we varied each basin’s threshold within 2 steps of the adopted value and computed the relative standard deviation of true positive counts as the sensitivity index σ_{OF} . The results are summarized in Table S3. The error bars in Fig. 10 were calculated as $\delta N_{\text{year}} = N_{\text{year}} \times \sigma_{\text{OF}}$, providing a conservative estimate of the algorithmic uncertainty in the temporal trends.

5. The use of the Mann–Kendall test to infer installation dates from SAR backscatter time series is an interesting approach, but its reliability is uncertain given the limited temporal coverage of Sentinel-1 data. Abrupt changes in backscatter may not uniquely correspond to construction events. The manuscript would benefit from additional validation or uncertainty analysis for these inferred temporal attributes.

We acknowledge that the reliability of Mann-Kendall (MK)-based inference of installation and removal dates from SAR backscatter time series requires careful discussion, and we address the reviewer’s concerns regarding temporal coverage limitations and potential uncertainty below.

The Sequential Mann-Kendall test was applied to the monthly VH backscatter coefficient time series extracted within a candidate buffer zone around each detected platform to identify two types of abrupt change points: a statistically significant upward trend change ($|z| > 1.96$, $p < 0.05$) corresponding to platform installation, and a statistically significant downward trend change corresponding to platform removal or decommissioning. The forward sequential statistic (UFK) and backward sequential statistic (UBK) are computed for each time series, and their intersection within the ± 1.96 significance bounds identifies the timing of the abrupt change. Only platforms whose MK statistics exceeded the significance threshold were assigned an installation or removal date; platforms with no statistically significant trend were recorded with null values in the dataset.

Sentinel-1 data have been available since 2014, with consistent global coverage from 2017 onward. Platforms installed or removed prior to this observation window cannot be reliably captured by the MK method; therefore, their dates are recorded as null values. This is an inherent constraint of satellite-based monitoring rather than a methodological deficiency, and is explicitly noted in Table 2 and the dataset documentation.

A systematic validation of dates inferred using the MK method against reference records is not feasible within the scope of this study. The available reference datasets, including BSEE (NGoM), OGIM, and OESNS (NS), were primarily compiled from industry and regulatory reports rather than continuous satellite monitoring. Consequently, these datasets contain very few records of platforms installed or removed within the 2017-2023 Sentinel-1 observation window. Most platforms in the six study basins were established prior to 2017, meaning that the MK method cannot detect their installation events, while the reference datasets do not provide post-2017 records at sufficient density for a statistically robust comparison. We acknowledge this limitation and propose systematic validation as a priority direction for future work as longer SAR archives and more frequently updated reference datasets become available.

We conducted a qualitative visual assessment of MK-inferred installation and removal dates using a representative sample of platforms randomly selected from different study basins (Fig. 7). For each selected platform, the monthly Sentinel-1 VH backscatter time series was examined together with Sentinel-1 SAR and Sentinel-2 RGB imagery acquired before and after the inferred change point. Visual inspection confirmed distinct SAR backscatter changes associated with platform installation and removal events, while Sentinel-2 imagery further confirmed the appearance or disappearance of the corresponding platform structures. These results provide qualitative support for the reliability of MK-based temporal attribute inference within the constraints of the available data.

The uncertainty in the inferred dates arises from two sources. First, the monthly temporal resolution of the Sentinel-1 composites limited the temporal precision of the inferred dates to approximately one month. Second, for platforms installed or removed near the boundaries of the observation period (2017 or 2023), the available observations before or after the event may be insufficient for robust change-point localization, leading to increased temporal uncertainty. Therefore, we recommend that users treat the installation and removal date attributes as approximate temporal indicators rather than precise engineering records and cross-reference them with reference databases where available. Corresponding clarifications have been added to the revised manuscript and the Supplementary Material.

L275-L282: Representative platforms from different study basins were selected for qualitative visual assessment of MK-inferred installation and removal dates using Sentinel-1 SAR and Sentinel-2 RGB imagery acquired before and after the inferred change points (Fig. S7). The visual inspection confirms distinct SAR backscatter changes associated with platform installation and removal events, while the Sentinel-2 imagery further confirms the appearance or disappearance of the corresponding platform structures. The inferred installation and removal dates should be regarded as approximate temporal indicators rather than precise engineering records. The monthly temporal resolution of the Sentinel-1 composites limits the temporal precision of the inferred dates to approximately one month, while

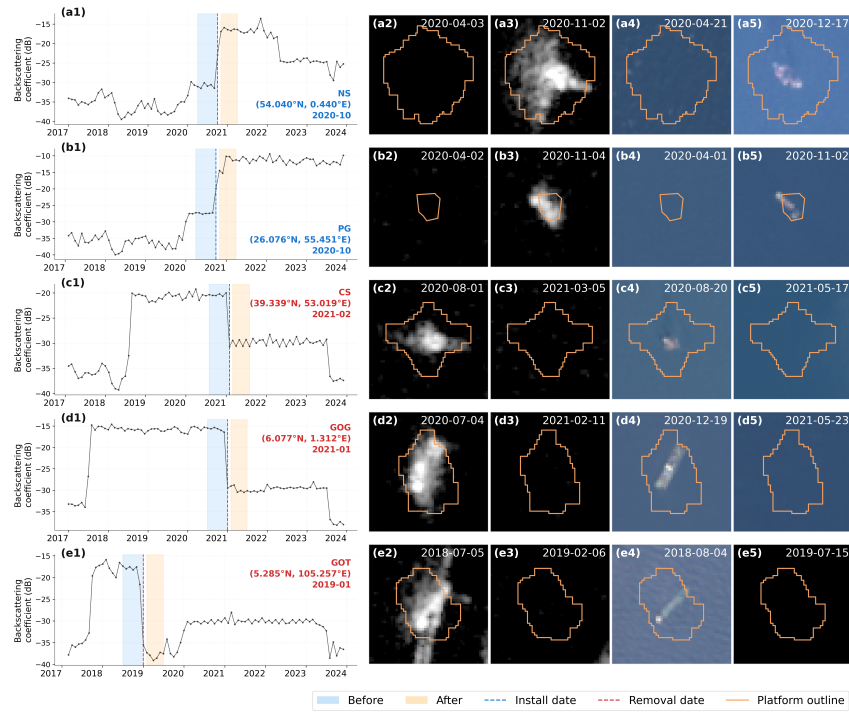


Figure 7: Visual validation of MK-inferred installation and removal dates for representative OOGPs from different study basins. Each row corresponds to one platform: (a1-e1) monthly Sentinel-1 VH backscatter coefficient time series (2017-2023); (a2-e2) Sentinel-1 VH SAR composite before the inferred change point; (a3-e3) Sentinel-1 VH SAR composite after the inferred change point; (a4-e4) Sentinel-2 RGB image before the change point; and (a5-e5) Sentinel-2 RGB image after the change point. Dashed lines indicate the MK-inferred change points. Blue and orange shading indicate the windows before and after imagery acquisition, respectively. Orange outlines delineate the platform polygon boundary.

limited observations near the boundaries of the observation period may further increase temporal uncertainty.

L434-L436: Fourth, the monthly temporal resolution of the Sentinel-1 composites limits the precision of installation and removal date estimates, and platforms installed or removed near the boundaries of the observation period may not be reliably captured by the Mann-Kendall change point detection.

Minor Comments:

1. In the Abstract, "existing OOGP databases are often incomplete or outdated data" should be corrected to "incomplete or contain outdated data".

We have revised this sentence in the Abstract accordingly. The original phrase “existing OOGP databases are often incomplete or outdated data” has been corrected to “existing OOGP databases are often incomplete or contain outdated data.”

2. The abstract and conclusions report slightly different accuracy figures (“98 %” vs. “0.99”). Aligning all statements with the precise metrics in Table 3 (precision = 0.98, recall = 0.92, F1 = 0.95) would eliminate any confusion.

We have revised the performance descriptions in both the Abstract and Conclusion to ensure consistency with the evaluation metrics reported in Table 3. Specifically, the previous expressions “extraction accuracy of 98%” and “overall accuracy of 0.99” have been replaced with the corresponding precision, recall, and F1 score values. We note that the updated metrics (precision = 0.95, recall = 0.89, F1 = 0.92) differ slightly from those cited in the reviewer’s comment, as the validation dataset was expanded and recomputed. The revised sentences are as follows:

L11-L13: An independent validation dataset was used to evaluate the performance of the detection algorithm, which achieved precision of 0.95, recall of 0.89, and F1 score of 0.92.

L470-L471: The OOGPs dataset achieved precision of 0.95, recall of 0.89, F1 score of 0.92, and mean error distance of 0.35 m.

3. P2, lines 25–30: “these aging platforms typically have three potential decommissioning options...” is wordy; tighten.

We thank the reviewer for this helpful suggestion. We have revised the sentence in lines 25–30 to make it more concise and readable. The revised text is as follows.

L28-L30: Aging platforms can be repurposed, recycled, or decommissioned through full removal, partial dismantling, or conversion into artificial reefs or aquaculture facilities. However, both operational and decommissioned OOGPs still pose environmental risks.

4. Terminology is generally consistent, but “OOGP(s)” and “offshore platforms” are occasionally mixed within the same paragraph. Using the acronym uniformly after its first definition would enhance precision.

We have carefully reviewed the manuscript and standardized the terminology throughout. After the first occurrence of “offshore oil and gas platforms (OOGPs),” the acronym “OOGPs” is now consistently used to avoid ambiguity and improve precision.

5. Figure 9’s color legend (red points) could be slightly deepened for better readability in print or PDF format.

We optimized the visualization of the reference platform points in Fig. 9 by applying a deeper red color and enhancing the symbol contrast and visibility. These adjustments improved the readability, particularly in the printed and PDF versions. The revised Fig. 9 has been updated in the manuscript.

6. Table 2 would benefit from a brief note on the exact date format (YYYYMM) and how null values for installation/removal dates are handled.

We have added a note to Table 2 to clarify that the installation and removal dates are reported in the yyyyymm format, representing the year and month. We also clarified that null values indicate unavailable or unreported installation/removal information in the original data. The corresponding note has been added below Table 2 in the revised manuscript to clarify this.

Note: Installation date and Removal date are recorded in yyyyymm format, representing the year and month. Null values indicate unavailable or unreported information.

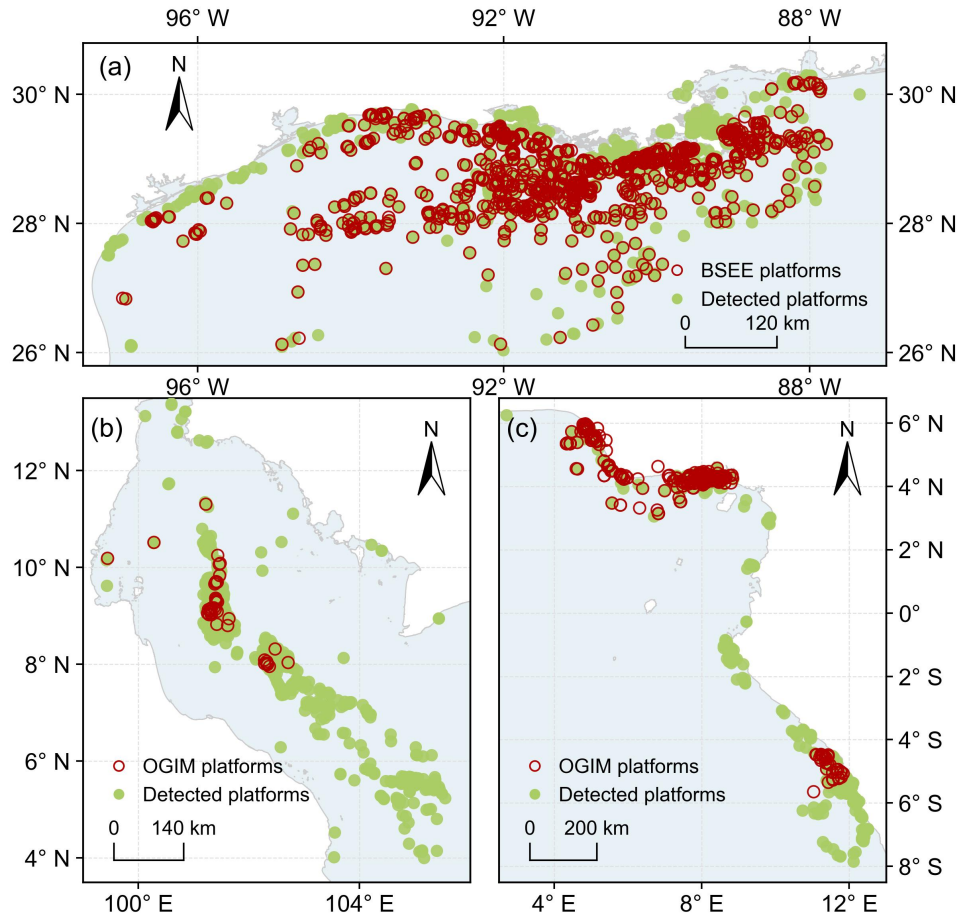


Figure 8: Comparison of the OOGPs map and BSEE or OGIM production in 2023. Green points represent OOGPs detected by Sentinel-1. The red circles indicate the referenced production. Panels (a), (b), and (c) indicate comparisons of the NGoM, GoT, and GoG, respectively.

References

- Chen, S., Wang, H., Xu, F., and Jin, Y.-Q.: Target Classification Using the Deep Convolutional Networks for SAR Images, 54, 4806–4817, <https://doi.org/10.1109/TGRS.2016.2551720>.
- Chen, X., Wang, Z.-q., Li, C.-x., Gao, F., and Wei, Q.: Characteristics of Global Offshore Oil and Gas Development, https://doi.org/10.1007/978-981-97-0475-0_87, 2024.
- Falqueto, L. E., Sá, J. A. S., Paes, R. L., and Passaro, A.: Oil Rig Recognition Using Convolutional Neural Network on Sentinel-1 SAR Images, 16, 1329–1333, <https://doi.org/10.1109/LGRS.2019.2894845>, conference Name: IEEE Geoscience and Remote Sensing Letters.
- Jackson, R. B., Irvin, J. A., Ramachandran, N., Wang, C., Ouyang, Z., Tulloch, P. A., Liu, F. Y., and Ng, A. Y.: MEthane Tracking Emissions Reference (METER): A global database of methane-emitting infrastructure, <https://doi.org/10.5194/essd-2026-124>.
- Kaiser, M. J.: Offshore oil and gas records circa 2020, Taylor & Francis, <https://www.tandfonline.com/doi/abs/10.1080/17445302.2020.1827633>, publisher: Taylor & Francis, 2022.
- Ma, R., Wu, W., Wang, Q., Liu, N., and Chang, Y.: Offshore Hydrocarbon Exploitation Target Extraction Based on Time-Series Night Light Remote Sensing Images and Machine Learning Models: A Comparison of Six Machine Learning Algorithms and Their Multi-Feature Importance, 15, 1843, <https://doi.org/10.3390/rs15071843>.
- Pei, J., Huang, Y., Huo, W., Zhang, Y., Yang, J., and Yeo, T.-S.: SAR Automatic Target Recognition Based on Multiview Deep Learning Framework, 56, 2196–2210, <https://doi.org/10.1109/TGRS.2017.2776357>.
- Spanier, R., Hoese, T., and Kuenzer, C.: Deep learning-based object detection of offshore platforms on Sentinel-1 imagery and the impact of synthetic training data, 47, 2120–2144, <https://doi.org/10.1080/01431161.2026.2612908>, _eprint: <https://doi.org/10.1080/01431161.2026.2612908>.
- Zhang, N., Zhao, H., Jing, P., Gao, Y., Liu, S., Shen, J., Huang, S., Zeng, Q., Liu, Y., and Huang, M.: Deep Learning-Based Intelligent Monitoring of Petroleum Infrastructure Using High-Resolution Remote Sensing Imagery, 14, 28, <https://doi.org/10.3390/pr14010028>, a.
- Zhang, Z., Yang, J., and Du, Y.: Deep Convolutional Generative Adversarial Network With Autoencoder for Semisupervised SAR Image Classification, 19, 1–5, <https://doi.org/10.1109/LGRS.2020.3018186>, b.
- Zhu, H., Jia, G., Zhang, Q., Zhang, S., Lin, X., and Shuai, Y.: Detecting Offshore Drilling Rigs with Multitemporal NDWI: A Case Study in the Caspian Sea, 13, 1576, <https://doi.org/10.3390/rs13081576>, number: 8.