

Response to the reviewers (#essd-2025-695)

Dear Editor and Reviewers,

We sincerely appreciate the valuable and constructive comments you have provided. In response, we have conducted a comprehensive and thorough revision of the manuscript to address all the comments and suggestions. We have responded to each point in detail to ensure that all concerns have been fully addressed. Your insightful feedback has significantly improved the overall quality of this manuscript. For ease of review, the original comments are presented in *italics*, our responses are provided in regular font, and the corresponding revisions in the manuscript are highlighted in red:

Reviewer #1 (Remarks to the Author):

Reviewer #1 Overall comments *The manuscript presents an important and timely contribution by producing the first decade-long, high-resolution (10 m) water surface PV (WPV) inventory for the Yangtze River Delta using a SAR–optical fusion approach. The topic is highly relevant, and the dataset could be valuable for future studies. However, several methodological aspects require clarification, and certain results need deeper interpretation before the manuscript can be considered for publication.*

[Response] We sincerely thank the reviewer for recognizing the importance and timeliness of our study, as well as the potential value of the decade-long, 10 m resolution WPV inventory for the Yangtze River Delta. We are encouraged that the reviewer considers the dataset to be a useful resource for future research. At the same time, we appreciate the reviewer's constructive comments regarding the need for clearer methodological descriptions and deeper interpretation of several results. These suggestions have helped us identify aspects of the manuscript that require further clarification and strengthening.

Reviewer #1 Major Comments

Reviewer #1 Specific comment 1 *Use of spectral indices – lack of quantitative thresholds. The methodology states that several Sentinel 2–derived spectral indices (e.g., NDVI, MNDWI, NDBI, NDPI, SAVI) were used in the Random Forest classifier. However, the manuscript does not provide any quantitative values, ranges, or thresholds that explain how these indices contribute*

to distinguishing: water surfaces, non-vegetated land, rocky or bare surfaces. Given the importance of spectral indices in the fusion approach, the authors should provide, at minimum: typical value ranges for water vs. land features, variable importance scores from the Random Forest model, and examples of how specific indices helped resolve misclassification challenges. This transparency is essential for reproducibility.

[Response] Thank you for this important suggestion. We agree that clearer quantitative information on the spectral indices is necessary to improve methodological transparency and reproducibility. We would like to clarify that the final WPV mapping was not based on fixed thresholds of individual indices. Instead, the Sentinel-2-derived indices were jointly used as predictor variables in the Random Forest classifier. In addition, the classification was constrained within a pre-constructed maximum historical water-extent mask, so the practical discrimination task mainly involved separating WPV from open water and other non-WPV features within water bodies.

Table R1. Typical value distributions of spectral indices for WPV and non-WPV samples.

Index	WPV median (IQR)	Non-WPV median (IQR)	<i>p</i> -value	Interpretation
NDVI	0.01 (-0.04–0.07)	0.16 (-0.07–0.41)	< 0.01	WPV concentrated near zero; non-WPV broader due to vegetation
MNDWI	-0.29 (-0.35–0.22)	-0.10 (-0.33–0.26)	< 0.01	WPV lower than open water due to panel coverage
NDBI	0.29 (0.28–0.36)	-0.10 (-0.23–0.03)	< 0.01	WPV higher because of artificial panel surface
SAVI	0.004 (-0.01–0.03)	0.07 (-0.02–0.22)	< 0.01	WPV near zero; non-WPV elevated where vegetation exists
NDPI	-3.15 (-5.22–0.23)	-0.58 (-0.76–0.23)	< 0.01	WPV higher than bare/open water

Note: P-values are from a two-sided Mann–Whitney U test. Smaller values indicate more significant differences between WPV and non-WPV samples.

To address this comment, we added quantitative summaries of the major spectral indices for WPV and non-WPV training samples in the revised manuscript (Table R1). These results show that WPV pixels generally have NDVI and SAVI values concentrated near zero, whereas

non-WPV samples display broader distributions because they include open water, vegetated margins, and other non-WPV surfaces within the water mask. WPV pixels also tend to have lower MNDWI values than open water because panel structures partially obscure the water surface, while NDBI and NDPI show distinct distributions between WPV and non-WPV samples, reflecting the artificial surface characteristics of photovoltaic panels. We emphasize that these values represent typical sample distributions rather than hard classification thresholds.

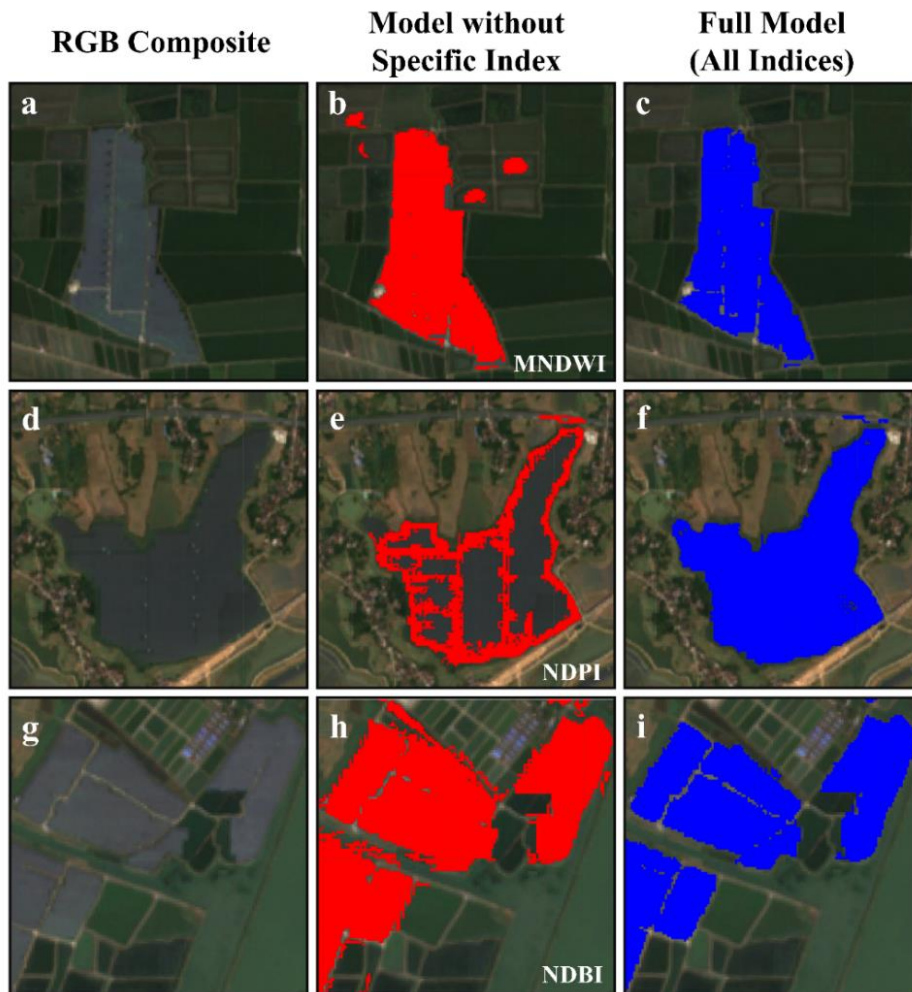


Figure R1. Examples of spectral indices helping resolve local confusion between WPV and non-WPV surfaces. Panels a–c, d–f, and g–i show three representative areas. a, d, g: displays the 2024 Sentinel-2 median RGB composites for each area; b, e, h: the classification results when a specific spectral index is omitted: b: MNDWI, e: NDPI, and h: NDBI; c, f, i: the classification results from the full random forest model using all indices. The three areas are located at a–c: 32°58'37.24"N, 119°37'23.43"E; d–f: 30°52'31.64"N, 117°35'35.81"E; and g–i: 31°29'44.93"N, 119°39'12.34"E (source: Copernicus Sentinel-2 imagery © ESA).

Table R2. Relative importance (%) of input variables in the Random Forest model.

Variable	Importance percent (%)
NDBI	33.04
NDPI	15.31
B11	8.49
MNDWI	6.29
B2	5.99
SAVI	5.91
B12	5.12
B8	5.06

Table R3. Effect of removing spectral indices on overall accuracy (OA, %).

Variable	Overall Accuracy (%)
Full (all features)	97.5
NDBI	97.3
NDPI	97.3
B11	97.3
MNDWI	97.4
B2	96.8
SAVI	97.5
B12	97.4
B8	97.4

We also quantified the contribution of individual predictors using Random Forest feature-importance scores (Table R2) and evaluated the effect of removing individual spectral variables on overall accuracy (Table R3). The results indicate that NDBI and NDPI are among the most important spectral indices, followed by MNDWI and SAVI. Although removing any single index caused only a small reduction in overall accuracy, this does not imply that the indices were uninformative; rather, it reflects the complementary and partly correlated nature of the predictor set in a high-performing ensemble classifier. To further clarify how the indices helped resolve classification challenges, we added explanatory text in Section 2.3.1 and provided representative examples in Fig. R1. The added text states: “**Specifically, NDVI and SAVI helped reduce confusion between WPV and vegetated surfaces such as emergent or floating vegetation near shorelines, MNDWI improved the separation of panel-covered water from open water, and NDBI/NDPI enhanced the identification of artificial panel surfaces that might otherwise resemble dark water in optical imagery.**” (Page 8, Lines 198–201 in the clean version of the manuscript)

Reviewer #1 Specific comment 2 *Interannual variability of water bodies. The surface area of lakes and reservoirs in the YRD can fluctuate significantly due to seasonal or multi year droughts. The manuscript does not explain how these hydrological variations were handled. Please clarify: Were annual water masks independently derived for each year? Did the classifiers incorporate hydrological seasonality? How were changes in water extent prevented from being misinterpreted as WPV presence or absence? This point is critical, especially when estimating decadal trends.*

[Response] Thank you for raising this important point. We agree that interannual and seasonal hydrological variability must be carefully considered when estimating decadal WPV trends. In the revised manuscript, we have clarified both the water-mask strategy and the temporal-consistency control, and we added a workflow figure (Fig. R2) to illustrate the full procedure.

First, annual water masks were not independently derived for each year. Instead, all yearly analyses were constrained using a unified maximum historical water-extent mask, representing the long-term potential water domain. This design reduces the influence of temporary shoreline expansion or contraction caused by hydrological fluctuations and helps avoid interpreting short-term water-extent changes as WPV gain or loss.

Second, the classifier did not explicitly incorporate hydrological seasonality through year-specific hydrological modeling. Rather, we mitigated its influence by generating annual Sentinel-1/2 composites for each year. The use of Sentinel-1 SAR is particularly helpful because it is less sensitive to cloud cover and illumination conditions and provides complementary structural information for WPV detection over water surfaces.

Third, to prevent temporary hydrological variation or image-quality differences from being misinterpreted as WPV disappearance, we did not treat the yearly classification outputs as fully independent final maps. Because WPV installations are generally persistent once deployed, we applied a temporal consistency rule in which previously confirmed WPV areas were retained in subsequent years and only newly detected areas were added. Thus, the final annual WPV series follows a cumulative, non-decreasing pattern. This strategy was intended to reduce spurious year-to-year disappearance caused by classification noise, short-term hydrological variability, or inconsistent image conditions. Finally, installation timing and project boundaries were further checked using Google Earth time-series imagery, which provided an additional

safeguard against false temporal changes introduced by remote-sensing classification alone.

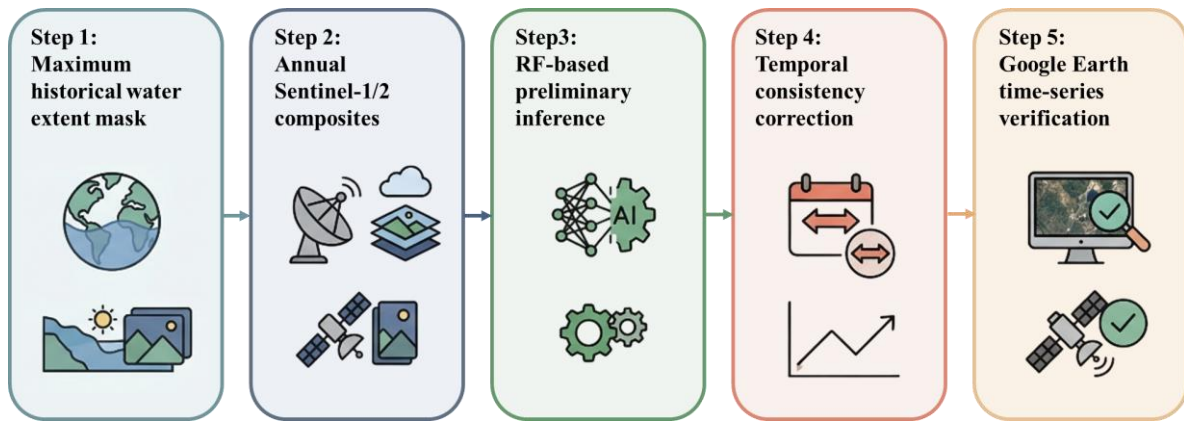


Figure R2. Workflow for annual WPV mapping and temporal consistency control. A maximum historical water extent mask was first used to constrain the analysis to long-term potential water bodies. Annual Sentinel-1/2 composites were then generated for each year and used for Random Forest-based WPV inference. To improve long-term temporal consistency, the preliminary annual maps were further updated using a cumulative temporal-consistency rule, in which previously confirmed WPV areas were retained in subsequent years and only newly detected areas were added. Finally, Google Earth time-series imagery was used to verify installation timing and project boundaries.

In response to reviewer comments, we have added a clarification in Section 2.3.1 (WPV Feature Engineering): “To reduce interference from temporary shoreline fluctuations and to focus the analysis on long-term potential aquatic areas, we filtered Sentinel-2 MSI imagery using a unified maximum historical water-extent mask, rather than deriving an independent water mask for each year.” (Page 8, Lines 185–187 in the clean version of the manuscript)

We have added a detailed clarification in Section 2.4.2 (Manual Refinement and Final Dataset Creation) regarding the temporal consistency of WPV mapping: “Each potential region was then interpreted and corrected using high-resolution satellite imagery from Google Earth (Fig. 3c) to accurately identify and remove misclassified non-WPV areas, thereby substantially improving the reliability of the final dataset. Specifically, each potential WPV region was checked for (1) its location within the water body, (2) the presence of regular and repetitive photovoltaic array patterns, and (3) separation from non-WPV objects such as shoreline buildings, roads, embankments, or floating vegetation. Since WPV installations are typically long-lasting, their installation year was determined by identifying the first year each site visibly appeared in high-resolution Google Earth imagery sequences. To maintain temporal consistency

across the annual series, previously confirmed WPV areas were retained in subsequent years, and only newly detected regions were added. Through this temporal consistency rule, the annual WPV series followed a cumulative, non-decreasing pattern, effectively reducing spurious year-to-year disappearance caused by classification noise, short-term hydrological variations, or image-quality differences, thereby enhancing the reliability of decadal trend estimation.” (Pages 10–11, Lines 252–265 in the clean version of the manuscript)

Reviewer #1 Specific comment 3 *Floating PV movement and texture features. The authors mention the use of texture metrics and SAR backscatter features. However, floating PV (FPV) systems—unlike fixed structures—can move due to wind, currents, or water level fluctuations. Please discuss: whether FPV motion affects texture features, whether SAR temporal variability could introduce classification noise, and whether the method is equally robust for fixed installations and mobile floating platforms. This clarification is important since China hosts many FPV plants.*

[Response] Thank you for this insightful comment. We agree that the potential movement of floating photovoltaic (FPV) systems should be considered when interpreting texture metrics and SAR backscatter features. In practice, however, most utility-scale FPV installations in China are anchored to the shoreline or lakebed. Although anchoring does not completely eliminate short-term movement caused by wind, currents, or water-level fluctuations, it generally constrains displacement to a limited range relative to the 10-m mapping scale used in this study.

To further evaluate the potential influence of FPV motion, we added a time-series analysis for three representative PV sites (Fig. R3), including two FPV plants and one stationary PV (SPV) installation. The results show that the optical texture features remain sufficiently stable for annual-scale classification after installation, although some local interannual variation is present. This suggests that limited FPV motion may affect local texture values, especially near patch boundaries, but does not systematically disrupt the patch-level spatial patterns used in our annual classification.

For SAR features, we acknowledge that Sentinel-1 observations may show temporal variability due to acquisition geometry, speckle, and short-term water-surface dynamics, which

can introduce noise at the single-scene level. However, Fig. R3 shows that the annual median VV and VH backscatter values exhibit a persistent shift relative to the pre-installation period and remain distinguishable thereafter, while the scatter of individual observations mainly reflects short-term variability. This indicates that annual temporal aggregation can effectively suppress SAR-related noise while preserving the installation-related signal.

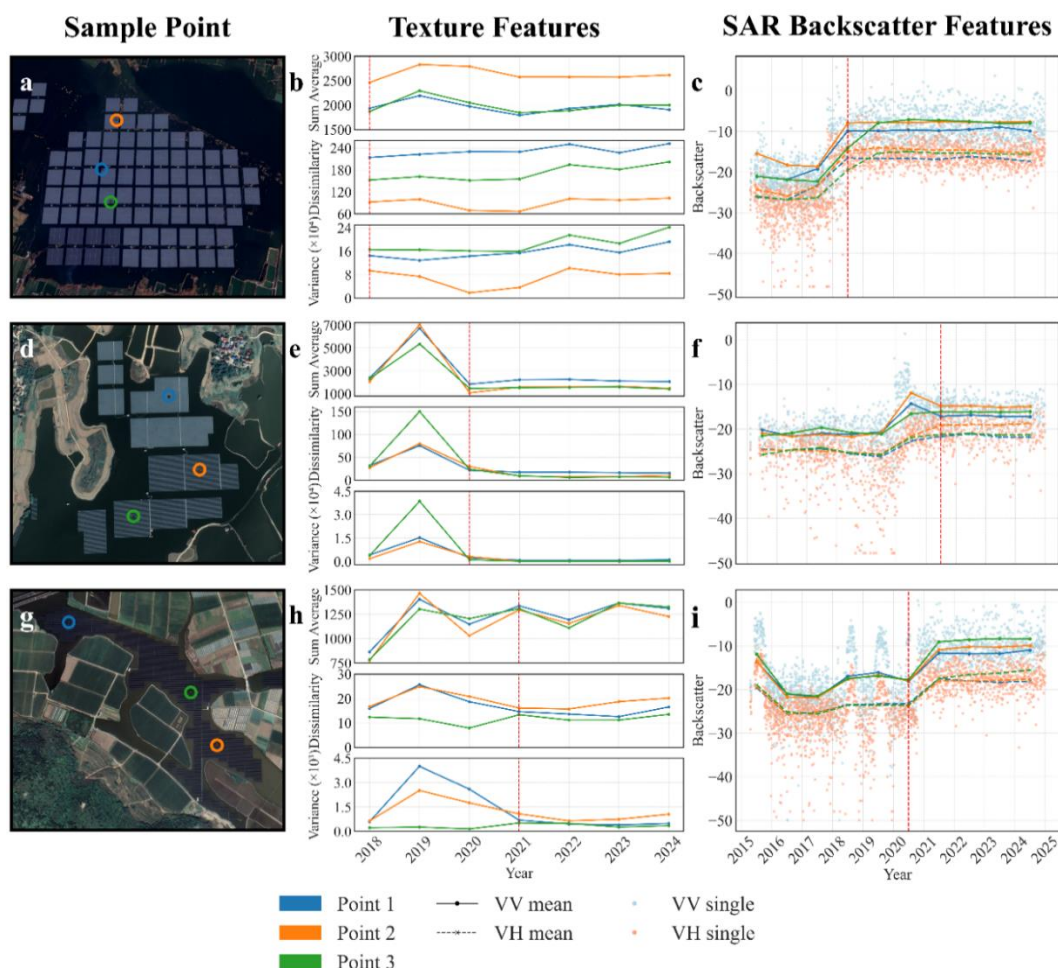


Figure R3. Time-series analysis of optical texture and SAR backscatter characteristics for representative PV sites. a, d, g: High-resolution imagery showing the locations of three representative sites and the selected sample points, where a and d are floating photovoltaic (FPV) plants, and g is a stationary photovoltaic (SPV) installation. b, e, h: Interannual variations in optical texture features derived from the gray-level co-occurrence matrix (GLCM), including B8 variance, B8 sum average, and B8 dissimilarity, for the selected points. c, f, i: Time series of Sentinel-1 SAR backscatter coefficients (VV and VH), where solid lines represent annual median values and dots represent individual observations. The comparison between annual median values and single-scene observations illustrates that, although short-term variability exists, the installation-related SAR signal remains persistent after

deployment. Overall, the figure shows that both texture features and annual SAR backscatter remain sufficiently stable after installation for annual WPV classification, despite limited FPV motion. The three sites are located at a: 32°49'09.88"N, 116°17'42.57"E; d: 30°45'41.47"N, 116°59'35.97"E; and g: 29°15'38.65"N, 121°43'38.65"E (source: Google Earth Pro, © 2025 Airbus, © 2025 Airbus / CNES, © 2025 Maxar Technologies).

We also clarify that the method is not equally robust for all platform types. The current framework is expected to be most reliable for stationary PV installations and for anchored, medium- to large-scale FPV systems, which represent the dominant commercial deployment type in the study region. In contrast, uncertainty may increase for small, loosely arranged, or highly mobile floating platforms, especially where platform displacement changes patch boundaries or weakens texture regularity. We have added this discussion to Section 4.3 (Limitations and Future Research) in the revised manuscript: “**The current framework is expected to be reliable for stationary PV installations and for anchored, medium- to large-scale FPV systems, which represent the dominant commercial deployment type in the study region. However, uncertainty may increase for small, loosely arranged, or highly mobile floating platforms, especially where local displacement alters patch boundaries or weakens texture regularity.**” (Page 18, Lines 467–471 in the clean version of the manuscript).

Reviewer #1 Specific comment 4 *Potential use of the methodology for environmental impact studies. The developed dataset could potentially support research on the environmental effects of FPV installations. Please comment on the feasibility of using this method to investigate: water surface temperature variations due to partial shading; changes in water colour or turbidity, especially related to algae bloom development or suppression; whether SAR–optical fusion offers the sensitivity needed for such environmental applications. These points would strengthen the broader applicability of the work.*

[Response] We sincerely thank the reviewer for this insightful suggestion regarding the broader environmental applicability of the dataset. We agree that the dataset can provide valuable support for future studies on the environmental effects of FPV deployment, but its role is primarily to deliver accurate spatial boundaries and installation timing of WPV systems rather than to directly retrieve environmental variables.

In the revised manuscript, we have clarified in Section 4.2 that the dataset can support environmental impact studies when integrated with other observations: “Its high-resolution delineation of WPV-covered and uncovered areas within the same water body enables spatially explicit comparisons between shaded and unshaded zones, as well as before-and-after analyses based on installation timing. When combined with other observations, these maps may support assessment of possible water-surface temperature differences associated with partial shading using thermal infrared products, as well as potential changes in water optical properties and algal dynamics using water-color or water-quality indicators such as chlorophyll-a, turbidity, or algal bloom proxies (Chen *et al.*, 2025; Chu and He, 2023).” (Page 17, Lines 443–449 in the clean version of the manuscript)

At the same time, we have revised Section 4.2 accordingly to clarify the methodological boundary of the present approach: “It should be noted, however, that the SAR–optical fusion framework is primarily designed to detect WPV extent, boundaries, and deployment timing. Although these attributes are relevant for environmental exposure assessment, the framework does not directly quantify thermal or biogeochemical responses, which require dedicated thermal infrared, water-quality remote sensing, or field observations.” (Page 18, Lines 454–459 in the clean version of the manuscript)

Reviewer #1 Specific comment 5 *Interpretation of high WPV coverage percentages (Fig. 11). Figure 11 shows that several basins have extremely high WPV coverage (85–95%). The manuscript should clarify: Which area was used as the denominator when computing the WPV percentage (e.g., maximum historical water extent, annual water extent, permanent water core). Whether such high coverage is physically accurate, or if classification steps may have overestimated WPV area in small or seasonally shrinking basins. The implications of these very high coverage levels for hydrological, ecological, or energy planning impacts. A deeper interpretation is needed.*

[Response] We appreciate this helpful comment. In the revised manuscript, we have clarified that WPV coverage was calculated using the maximum historical water extent of each water body as the denominator, rather than the annual water extent or a permanent-water core. We selected this reference area to provide a temporally stable basis for interannual comparison and

to avoid artificial inflation of coverage values caused by seasonal or drought-induced contraction of annual water extent.

We also agree that extremely high WPV coverage values (e.g., 85-95%) require careful interpretation. To assess whether such values are physically plausible, we further examined representative high-coverage cases using high-resolution satellite imagery (Fig. R4). These examples show that some small, managed, and highly enclosed water bodies in the study area, particularly aquaculture ponds and other compact artificial basins, can indeed be almost fully occupied by WPV installations, leaving only narrow margins of open water. Therefore, such high values are physically plausible in specific local settings.

At the same time, we acknowledge that coverage estimates for small or seasonally dynamic water bodies are more sensitive to boundary delineation uncertainty than those for large lakes and reservoirs. For this reason, very high coverage values should be interpreted with caution at the individual-basin level. Nevertheless, because the denominator was defined as the maximum historical water extent rather than the annual water extent, our approach reduces the risk of overestimating coverage when water extent temporarily contracts.

Finally, we have expanded Section 4.2, Implications and Potential Applications, to further interpret the significance of these high-coverage cases: “The high-precision, decade-long WPV dataset developed in this study has both practical and scientific relevance. By quantifying the coverage, spatial distribution, and waterbody-specific deployment characteristics of WPV installations, the dataset provides a spatial basis for WPV planning, site selection, and comparative analysis across waterbody types and coverage levels. High-coverage configurations, including those associated with “fishing-solar complementarity” systems, may indicate both high energy-generation intensity and stronger local hydrological or ecological influence (Pringle *et al.*, 2017). Such settings may be associated with greater modification of light availability, air-water exchange, evaporation, and aquatic habitat conditions. From an energy-planning perspective, they also represent an intensive deployment mode, although future expansion should account for environmental carrying capacity and site-specific management constraints (Bai *et al.*, 2024; Château *et al.*, 2019).” (Pages 16–17, Lines 424–434 in the clean version of the manuscript)

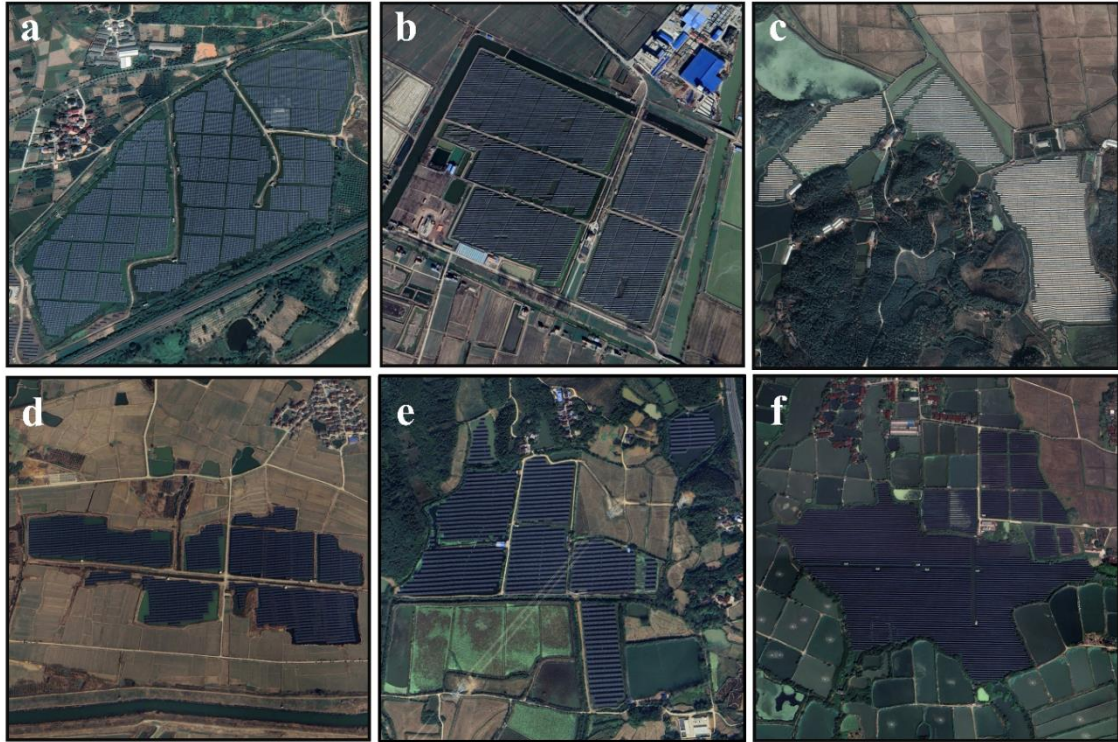


Figure R4. Examples of WPV installations with very high coverage. The six areas are located at a: 28°49'18.32"N, 118°42'34.22"E; b: 33°48'20.33"N, 119°38'24.37"E; c: 30°44'18.35"N, 117°53'33.14"E; d: 31°55'07.12"N, 119°22'59.66"E; e: 30°31'24.49"N, 117°21'25.91"E; f: 30°46'25.86"N, 120°07'53.47"E (source: Google Earth Pro, © 2025 Airbus, © 2025 Airbus / CNES, © 2025 Maxar Technologies).

Reviewer #1 Specific comment 6 *Distinguishing lakes vs. reservoirs. Please provide a clear definition of lake versus reservoir, since the distinction is relevant for WPV siting policies, water level stability, and ownership/management regimes. A short paragraph is needed in the Methods or Study Area section.*

[Response] Thank you for this helpful comment. We agree that a clear distinction between lakes and reservoirs is important for interpreting WPV siting patterns, water-level dynamics, and waterbody management characteristics. In the revised manuscript, we added a short clarification in Section 2.2.2 (Water Body Datasets): “**In this study, reservoirs were defined as water bodies formed by dams or other engineered hydraulic infrastructure and subject to artificial water-level regulation. In contrast, lakes were defined as inland water bodies without reservoir functions, including natural lakes as well as artificial water bodies such as aquaculture ponds and irrigation**”

ponds. This distinction is relevant for WPV analysis because reservoirs generally exhibit more stable water levels and are associated with different ownership and management regimes than lakes.” (Page 7, Lines 157–162 in the clean version of the manuscript)

Reviewer #1 Minor Comments Figure 2 labeling error. There is an inconsistency between the letters shown in the images and those referenced in the caption. Please correct the figure annotations to ensure correspondence

[Response] Thank you for noting this inconsistency. We have corrected the panel annotations in Figure 2 and revised the caption accordingly to ensure consistency between the figure labels and the caption text. The revised caption now reads: “**a–c: Examples of typical sample regions, where red points indicate WPV samples and blue points indicate non-WPV samples; d: Spatial distribution of WPV and non-WPV samples across the study area; e: Proportions of different sample categories.**” (Page 28, Lines 691–694 in the clean version of the manuscript)

Reviewer #2 (Remarks to the Author):

Reviewer #2 Overall comments *This paper is well written and presents solid data and methodological rigor. I have several suggestions that may further strengthen the manuscript:*

[Response] We sincerely thank you for this positive overall assessment of our manuscript. We greatly appreciate your recognition of the manuscript's writing quality, data basis, and methodological rigor. We have carefully considered all your suggestions and revised the manuscript accordingly to further improve its clarity, transparency, and overall contribution. Below, we provide a point-by-point response to the specific comments.

Reviewer #2 Major comments

Reviewer #2 Specific comment 1 *Clarification of Water Mask (Line 176). In Line 176, please clearly specify which waterbody dataset was used as the water mask. If multiple datasets were integrated, describe their respective roles and how they were harmonized.*

[Response] Thank you for this helpful comment. We agree that the construction of the water mask should be described more clearly. In the revised manuscript, we clarified that the water mask was not derived from a single dataset, but was constructed by integrating four waterbody-related datasets: HydroLAKES, GRanD, GOODD, and GeoDar. Specifically, HydroLAKES was used to provide lake boundaries and associated attributes, while GRanD supplied reservoir polygons and reservoir-related information. GOODD and GeoDar were used as complementary dam datasets to support the identification of reservoir systems and to improve the consistency of reservoir records where polygon or attribute information was incomplete or uncertain. These datasets were harmonized into a unified maximum historical water-extent layer for the study area, which served as a stable water mask to constrain the analysis to long-term potential water bodies.

We further clarified in the revised manuscript (Section 2.2.2 “Water Body Datasets”) that this integrated water mask was used as a spatial constraint for WPV candidate areas, rather than as a year-specific annual water mask. “To construct the water mask for WPV mapping, we integrated four waterbody-related datasets: HydroLAKES, GRanD, GOODD, and GeoDar (Lehner *et al.*, 2011; Messenger *et al.*, 2016; Mulligan *et al.*, 2020; Wang *et al.*, 2022). HydroLAKES provided lake boundaries and associated attributes, while GRanD supplied

reservoir polygons and reservoir-related information. GOODD and GeoDar, which contain georeferenced dam records, were used as complementary datasets to support reservoir identification where reservoir boundary or attribute information was incomplete or uncertain. These datasets were harmonized into a unified maximum historical water-extent layer for the study area. This integrated layer was used as a stable spatial mask to constrain the analysis to long-term potential water bodies, rather than as a year-specific annual water mask.” (Pages 6–7, Lines 147–156 in the clean version of the manuscript)

Reviewer #2 Specific comment 2 Multicollinearity Discussion. *Since the model integrates original optical bands together with derived spectral indices, it would be helpful to briefly discuss the potential multicollinearity between these variables. Although Random Forest is generally robust to correlated predictors, explicitly stating that multicollinearity does not adversely affect classification performance under tree-based models would improve methodological transparency.*

[Response] Thank you for this helpful suggestion. We agree that the potential multicollinearity between the original optical bands and the derived spectral indices should be briefly addressed. In the revised manuscript, we clarified that several spectral indices are mathematically derived from the original Sentinel-2 bands and are therefore not fully independent of them.

We also noted that such multicollinearity is unlikely to substantially affect classification performance in this study, because Random Forest is a tree-based ensemble model that is generally robust to correlated predictors. The inclusion of both original bands and derived indices was intended to retain complementary spectral information for WPV discrimination. At the same time, predictor correlation may influence the interpretation of variable importance, even though it does not materially compromise classification accuracy. The relevant text has been added to Section 2.3.1 (WPV Feature Engineering): “**Although some spectral indices are derived from the original optical bands and are therefore correlated with them, Random Forest is generally robust to multicollinearity among predictors in classification tasks. As a result, the inclusion of both original bands and derived indices is unlikely to adversely affect model performance, although it may influence the interpretation of variable importance.**” (Page 8, Lines 202–206 in the clean version of the manuscript)

Reviewer #2 Specific comment 3 *Random Forest Model Specification. Please provide more details about the Random Forest implementation, including key hyperparameters (e.g., number of trees, maximum tree depth, minimum samples per leaf) and whether cross-validation was used.*

[Response] Thanks for this helpful suggestion. We agree that providing more detailed information on the Random Forest implementation is important for methodological transparency and reproducibility. In the revised manuscript, we have expanded the description of the classifier settings and validation strategy.

Specifically, the Random Forest classifier was implemented in Google Earth Engine using the `smileRandomForest` algorithm. We used 80 trees and randomly selected 7 variables at each split. The remaining hyperparameters were left at their default values in the GEE implementation, including minimum leaf population = 1, bag fraction = 0.5, maximum nodes = null, and seed = 0. The detailed Random Forest settings are summarized in Table R4.

Table R4. Random Forest implementation settings used in this study.

Parameter	Setting
Platform	Google Earth Engine
Algorithm	<code>smileRandomForest</code>
Number of trees	80
Variables per split	7
Minimum leaf population	1 (default)
Bag fraction	0.5 (default)
Maximum nodes	Null (default)
Seed	0 (default)
Cross-validation	Used

To further evaluate model robustness, we performed 10-fold stratified cross-validation on the training dataset. The model achieved an average accuracy of 0.9729 ± 0.0030 across the ten folds, indicating stable classification performance. These details have now been added to

Section 2.3.2, Annual WPV Classification: “The Random Forest classifier was implemented in Google Earth Engine using the smileRandomForest algorithm. We used 80 trees and randomly selected 7 variables at each split. The remaining hyperparameters were left at their default values in the GEE implementation, including minimum leaf population = 1, bag fraction = 0.5, maximum nodes = null, and seed = 0.” (Page 9, Lines 223–227 in the clean version of the manuscript)

Reviewer #2 Specific comment 4 Comparison with Existing Global Inventories. *The manuscript would benefit from a clearer introduction and discussion of existing datasets such as A Global Inventory of PV and Global Renewables Watch. Please briefly describe their data sources, spatial resolution, update frequency, and methodological framework. Additionally, discuss how your approach improves upon these traditional inventories (e.g., spatial resolution, temporal consistency, detection accuracy, validation strategy).*

[Response] Thank you for this valuable suggestion. We have revised the Introduction and Discussion to provide a clearer comparison with representative global PV inventories, including *A Global Inventory of Photovoltaic Installations* and *Global Renewables Watch*. In the revised manuscript, we now briefly summarize their data sources, spatial resolution, temporal coverage and update frequency, methodological framework, and validation characteristics, and we added Table R5 to present this comparison more clearly.

We also clarified how the present study differs from these existing products. In particular, our study focuses specifically on WPV rather than general solar infrastructure; provides annual 10-m maps for 2015–2024; uses Sentinel-1/2 SAR–optical fusion to improve detection over water surfaces; and incorporates temporal consistency control together with Google Earth time-series verification. These revisions have been added to Section 1 (Introduction): “Building on these advances, several global photovoltaic inventories have been developed to provide large-scale datasets of PV installations. For example, *A Global Inventory of Photovoltaic Installations* maps commercial-, industrial-, and utility-scale PV facilities worldwide using multi-source satellite imagery, including Sentinel-2 (10 m) and high-resolution SPOT-6/7 imagery, and was produced using deep learning with manual verification for installations identified primarily from 2016 to 2018 (Kruitwagen et al., 2021). Similarly, *Global Renewables Watch* integrates

quarterly PlanetScope imagery at 4.7 m spatial resolution with deep learning–based semantic segmentation to map global solar and wind infrastructure from 2017 Q4 to 2024 Q2, enabling regular updates and temporal tracking of infrastructure development (Robinson et al., 2025). Compared with these products, our study is specifically designed for WPV mapping and combines SAR and optical observations to better distinguish floating PV from surrounding water backgrounds, while also providing annual, temporally consistent WPV inventories for 2015–2024.” (Pages 3–4, Lines 66–79 in the clean version of the manuscript).

Table R5. Comparison of representative global PV inventories and the WPV dataset developed in this study.

Dataset	Source	Resolution	Time	Validation	WPV relevance
<i>Global Inventory of PV</i>	Sentinel-2, SPOT 6/7	10 m to sub-5 m	Up to 2018	Cross-validation	Not WPV-specific
<i>Global Renewables Watch</i>	PlanetScope	4.7 m	2017 Q4–2024 Q2	Comparison with IRENA statistics	Not optimized for WPV
This study	Sentinel-1/2	10 m	2015–2024 annual	Sample-based assessment + time-series Google Earth verification	WPV-focused regional inventory

Reviewer #2 Specific comment 5 *Provincial-Level Proportional Analysis (Section 3.2). In Section 3.2, in addition to reporting cumulative WPV area, I recommend presenting the ratio of cumulative WPV area to total water area for each province. This proportional metric would provide more meaningful insight into deployment intensity and facilitate comparison across provinces.*

[Response] We thank the reviewer for this valuable suggestion. We agree that the ratio of WPV area to total water area is a more informative measure of deployment intensity than cumulative area alone. In the revised manuscript, we added a new inset bar chart to Fig. 9b (upper-left inset) and explicitly reported this proportional metric in Section 3.2. The revised text states: “**The inset**

bar chart further illustrates the proportion of WPV area relative to the total water area in each province in 2024, showing that Anhui had the highest deployment intensity (1.01%), followed by Jiangsu (0.67%) and Zhejiang (0.48%).” (Page 13, Lines 330–332 in the clean version of the manuscript)

Reviewer #2 Specific comment 6 *Overall, this study makes an important contribution. Subject to addressing the points raised above, I support publication of this manuscript.*

[Response] We sincerely thank the reviewer for the positive and encouraging assessment of our manuscript. We greatly appreciate the reviewer’s recognition of the importance and contribution of this study. In response to the comments provided above, we have carefully revised the manuscript and addressed all concerns to the best of our ability.

Reviewer #3 (Remarks to the Author):

Reviewer #3 Overall comments *This manuscript collected RS based reflectance, vegetation, water and built-up indices, texture and SAR data and trained a random forest (RF) classifier to extract floating water photovoltaics (WPV) over the downstream area of Yangtze river. Authors validated the final products and analyzed the recent trend of WPV projects. This manuscript is well written, except for some technical details that are missing. In addition, my major concern on this work is that I'm conservative on the application of this WPV product. The authors mentioned some cases of how others can use this product. But as an earth system science researcher, I strongly recommend the author to provide an application scenario under a broader earth system framework, which is also the main scope of ESSD, e.g., how to use this data product to improve the energy sectors and its interaction with others under an integrated assessment perspective.*

[Response] Thank you for this thoughtful comment. We agree that the broader application value of the WPV dataset should be framed more explicitly within an Earth system science perspective. In the revised manuscript, we expanded Section 4.2 to clarify that the dataset is not only an inventory of WPV distribution, but also a spatially explicit basis for analyzing interactions among renewable-energy development, waterbody use, and environmental change. In this way, the dataset can support integrated assessment of low-carbon energy transitions, particularly in regions where water-surface solar deployment is closely linked to aquaculture, reservoir regulation, and ecosystem management.

In addition, because the dataset distinguishes WPV-covered and uncovered areas within the same water body and provides annual information on deployment timing, it enables spatially explicit comparisons and before-and-after analyses. This creates opportunities to combine the dataset with other Earth observation products to investigate potential hydrological, thermal, and ecological responses associated with WPV expansion. For example, it can be integrated with thermal infrared products to examine possible water-surface temperature differences related to partial shading, and with water-color or water-quality indicators (e.g., chlorophyll-a, turbidity, or algal bloom proxies) to explore possible changes in optical properties and algal dynamics.

These applications are relevant not only for identifying possible environmental risks in high-coverage WPV systems, but also for understanding how energy-sector expansion may

interact with water-resource functions and ecosystem processes under a broader coupled human-environment framework. The relevant text has been revised in Section 4.2 (Implications and Potential Applications). “**The high-precision, decade-long WPV dataset developed in this study has both practical and scientific relevance... which require dedicated thermal infrared, water-quality remote sensing, or field observations.**” (Pages 16–18, Lines 424–459 in the clean version of the manuscript)

Reviewer #3 Major Comments

Reviewer #3 Specific comment 1 *Input data of RF classifier include both direct reflectance of each Sentinel-2 MSI band and the combination of them, i.e., normalized indices, which have strong dependency among these different inputs. Though the classifier is relatively simple (only tell if a grid is WPV or not), which can cause the impact of this issue not reflected in your study, I would still recommend authors to discuss the potential impact from dependency in your inputs, i.e. multi-collinearity, when training your RF classifier.*

[Response] We appreciate this helpful comment. We agree that some spectral indices are mathematically derived from the original optical bands and may therefore be correlated with them, leading to potential multicollinearity among the input features. However, Random Forest (RF) is generally robust to correlated predictors in classification tasks, and such multicollinearity is therefore unlikely to substantially affect model performance, although it may influence the interpretation of variable importance. In our case, the task is a binary classification problem (WPV vs. non-WPV), and we did not observe any evident degradation in classification performance associated with the inclusion of correlated variables. Following the reviewer’s suggestion, we added a brief clarification in Section 2.3.1 (WPV Feature Engineering) of the revised manuscript: “**Although some spectral indices are derived from the original optical bands and are therefore correlated with them, Random Forest is generally robust to multicollinearity among predictors in classification tasks. As a result, the inclusion of both original bands and derived indices is not expected to adversely affect model performance, although it may influence the interpretation of variable importance.**” (Page 8, Lines 202–206 in the clean version of the manuscript)

Reviewer #3 Specific comment 2 *For RF classifier training, how did you gain a robust model? Did you consider 10-fold cross-validation? How are parameters determined when training the model? It will also help if authors can think about using the SHAP value to reflect the importance of each feature on FPV detection.*

[Response] This is an important suggestion. We agree that additional details on model validation, parameter settings, and feature interpretation improve methodological transparency. In the revised manuscript, we clarified that the sample dataset was divided into 80 % training and 20 % testing subsets, and that 10-fold stratified cross-validation was conducted on the training subset to assess model stability. The cross-validation yielded a mean accuracy of 0.9729 \pm 0.0030 across the ten folds, indicating stable classification performance.

We also added the Random Forest parameter settings. The model was implemented in Google Earth Engine using the `smileRandomForest` algorithm, with 80 trees and 7 variables randomly selected at each split. The remaining hyperparameters were kept at the default settings of the GEE implementation, including minimum leaf population = 1, bag fraction = 0.5, maximum nodes = null, and seed = 0. In addition, following the reviewer’s suggestion, we calculated SHAP values to further interpret feature contributions. The SHAP-based ranking (Table R6) is broadly consistent with the RF variable-importance results, with NDBI, B2, NDPI, B11, and MNDWI emerging as the most influential predictors for WPV detection.

Table R6. Global Feature Importance for WPV Detection Based on SHAP Values.

Feature	Mean Absolute SHAP
NDBI	0.1802
B2	0.0630
NDPI	0.0529
B11	0.0523
MNDWI	0.0457
B12	0.0300
B8	0.0235
SAVI	0.0227

Reviewer #3 Technical comments

Reviewer #3 Specific comment 3 *Line 32: What does "eliminated errors" mean?*

[Response] We intended to indicate that misclassified WPV areas were identified and removed through visual interpretation of Google Earth time-series imagery. Accordingly, “eliminated errors” has been revised to “**removed misclassified areas**” in the revised manuscript. (Page 2, Line 32 in the clean version of the manuscript)

Reviewer #3 Specific comment 4 *Line 90: "unprecedented accuracy" can be changed to "high".*

[Response] Agree. We have replaced “unprecedented accuracy” with “**high accuracy**” to avoid overstatement in the revised manuscript (Page 5, Line 103; Page 16, Line 415 in the clean version of the manuscript).

Reviewer #3 Specific comment 5 *Line 131: "To reduce cloud interference, a cloud-masking algorithm was applied, and annual median composites were generated from all available images." Repeated sentence. Please delete.*

[Response] We thank the reviewer for pointing this out. The repeated sentence has been removed from the revised manuscript.

Reviewer #3 Specific comment 6 *Line 134: "These composites ensure radiometric consistency and provide a stable spatial baseline for dynamic WPV detection and temporal analysis." Repeated sentence. Please delete.*

[Response] We thank the reviewer for pointing this out. The repeated sentence has been removed from the revised manuscript.

Reviewer #3 Specific comment 7 *Line 138: It can be helpful to have a flow chart describing how you merged different products into a unified water mask.*

[Response] We thank the reviewer for this valuable suggestion. We agree that a flow chart helps clarify how the unified water mask was constructed, and we have therefore added a new flow chart (Fig. R5) in the revised manuscript. Specifically, HydroLAKES provides lake boundaries and associated attributes, while GRanD supplies reservoir polygons and related information;

these were first integrated to form the preliminary waterbody dataset. GOODD and GeoDar were then used as complementary dam datasets to support reservoir identification and improve the consistency of reservoir-related records where polygon or attribute information was incomplete or uncertain. The resulting waterbodies were subsequently classified into two categories: reservoirs, for those identified as dam-controlled systems, and lakes, for the remaining waterbodies. Finally, all datasets were harmonized into a unified maximum historical water-extent layer for the study area, which served as a stable water mask to constrain the analysis to long-term potential water bodies.

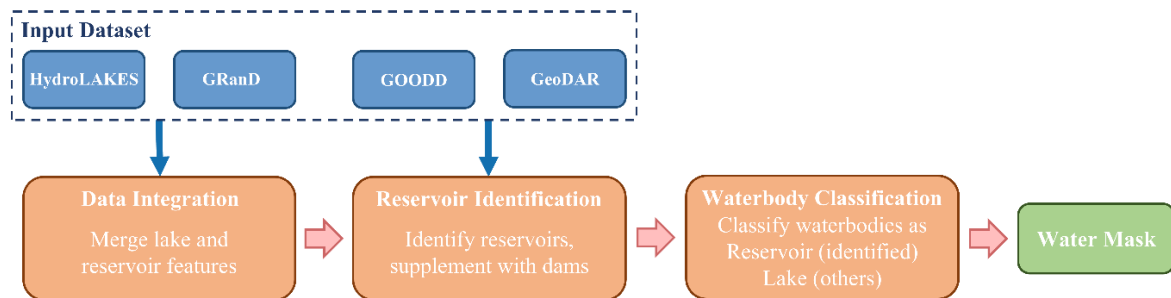


Figure R5. Flow chart of water mask construction. HydroLAKES provides lake boundaries, and GRanD supplies reservoir polygons; these are integrated to form the preliminary waterbody dataset. GOODD and GeoDar are used to improve reservoir identification. Waterbodies are then classified into two categories: identified reservoirs and the remaining lakes. All datasets are finally harmonized into a unified maximum historical water-extent layer, which serves as a stable water mask.

Reviewer #3 Specific comment 8 *Figure 2: "a" shall be "d", "b" shall be "e", "c-e" shall be "a-c".*

[Response] Thank you for pointing out this error. The panel labels and caption of Figure 2 have been revised accordingly in the revised manuscript.

Reviewer #3 Specific comment 9 *Line 193: Did you calculate texture features for all 6 bands?*

[Response] Thank you for this question. Texture features were calculated only for the B8 (NIR) band, as it provides a strong contrast between water surfaces and WPV. We have clarified this in Section 2.3.1 (“Texture features”) of the revised manuscript: “**Calculated from the B8 (NIR) band using the Gray Level Co-occurrence Matrix (Haralick et al., 1973). These features capture**

the characteristic spatial patterns of WPV arrays, which typically exhibit clear, regular boundaries that contrast with natural water bodies.” (Page 9, Lines 212–214 in the clean version of the manuscript)

Reviewer #3 Specific comment 10 Line 233: *"Each potential region was then subjected to rigorous manual interpretation and correction using high-resolution satellite imagery from Google Earth (Fig. 3c)." Please clarify what "manual interpretation" method you used. Based on expert judgement?*

[Response] We agree that the term “manual interpretation” should be defined more clearly. In this study, the interpretation was based on expert visual judgement using high-resolution Google Earth imagery, but it followed explicit visual criteria and temporal consistency checks rather than purely subjective assessment. Specifically, each potential WPV region was evaluated based on (1) its location within the water body, (2) the presence of regular and repetitive photovoltaic array patterns, and (3) clear separation from non-WPV objects such as shoreline buildings, roads, embankments, or floating vegetation. In addition, temporal consistency in Google Earth historical imagery was examined to confirm plausible installation timing and continued presence after deployment. Based on these criteria, false positives were removed and installation timing was further verified.

We have revised the manuscript (section 2.4.2 Manual Refinement and Final Dataset Creation) to clarify this manual interpretation and correction procedure. “Each potential region was then interpreted and corrected using high-resolution satellite imagery from Google Earth (Fig. 3c) to accurately identify and remove misclassified non-WPV areas, thereby substantially improving the reliability of the final dataset. Specifically, each potential WPV region was checked for (1) its location within the water body, (2) the presence of regular and repetitive photovoltaic array patterns, and (3) separation from non-WPV objects such as shoreline buildings, roads, embankments, or floating vegetation. Since WPV installations are typically long-lasting, their installation year was determined by identifying the first year each site visibly appeared in high-resolution Google Earth imagery sequences. To maintain temporal consistency across the annual series, previously confirmed WPV areas were retained in subsequent years, and only newly detected regions were added. Through this temporal consistency rule, the annual

WPV series followed a cumulative, non-decreasing pattern, effectively reducing spurious year-to-year disappearance caused by classification noise, short-term hydrological variations, or image-quality differences, thereby enhancing the reliability of decadal trend estimation.” (Pages 10–11, Lines 252–265 in the clean version of the manuscript)

Reviewer #3 Specific comment 11 *Line 250: It is not clear how you build a model to account for multi-year data. Did you combine multi-year reflectance, indices and texture and SAR data together and use them as input to train your model? Please explain the methodology.*

[Response] We agree that the multi-year classification workflow should be described more explicitly. In this study, the Random Forest classifier was trained using sample points interpreted from the 2024 annual median composite. We chose 2024 as the reference year because all training samples could be matched to confirmed WPV installations, whereas using earlier-year imagery could introduce samples from locations where WPV had not yet been deployed.

Multi-year reflectance, spectral indices, texture, and SAR features were not combined into a single multi-temporal input stack for model training. Instead, for each year from 2015 to 2024, we separately generated annual Sentinel-1/2 composites and derived the corresponding spectral, index, texture, and SAR features using the same feature-construction procedure. The RF model trained on the 2024 samples was then applied to each year’s annual feature set to produce year-specific WPV classification results.

To improve temporal consistency in the annual series, the initial yearly classification outputs were further refined using a cumulative temporal-consistency rule: once a WPV area was confirmed, it was retained in subsequent years, and only newly detected areas were added. Google Earth time-series imagery was additionally used to verify installation timing and to correct potential temporal inconsistencies. We have clarified this methodology in the revised manuscript.

Reviewer #3 Specific comment 12 *Figures 7, 8: Please change FPV to WPV.*

[Response] Thank you. The captions of Figures 7 and 8 have been updated accordingly.

References

- Bai, B., Xiong, S., Ma, X., and Liao, X.: Assessment of floating solar photovoltaic potential in China, *Renew. Energy*, 220, 119572, <https://doi.org/10.1016/j.renene.2023.119572>, 2024.
- Château, P.-A., Wunderlich, R. F., Wang, T.-W., Lai, H.-T., Chen, C.-C., and Chang, F.-J.: Mathematical modeling suggests high potential for the deployment of floating photovoltaic on fish ponds, *Sci. Total Environ.*, 687, 654 – 666, <https://doi.org/10.1016/j.scitotenv.2019.05.420>, 2019.
- Chen, D., Peng, Q., Lu, J., Huang, P., Liu, Y., and Peng, F.: Assessing effect of water photovoltaics on nearby water surface temperature using remote sensing techniques, *Adv. Space Res.*, 75, 138 – 147, <https://doi.org/10.1016/j.asr.2024.08.040>, 2025.
- Chu, H.-J., and He, Y.-C.: Remote sensing water quality inversion using sparse representation: Chlorophyll-a retrieval from Sentinel-2 MSI data, *Remote Sens. Appl. Soc. Environ.*, 31, 101006, <https://doi.org/10.1016/j.rsase.2023.101006>, 2023.
- Haralick, R. M., Shanmugam, K., and Dinstein, I.: Textural features for image classification, *IEEE Trans. Syst. Man Cybern.*, SMC-3, 610 – 621, <https://doi.org/10.1109/TSMC.1973.4309314>, 1973.
- Kruitwagen, L., Story, K. T., Friedrich, J., Byers, L., Skillman, S., and Hepburn, C.: A global inventory of photovoltaic solar energy generating units, *Nature*, 598, 604 – 610, <https://doi.org/10.1038/s41586-021-03957-7>, 2021.
- Lehner, B., Liermann, C. R., Revenga, C., Vörösmarty, C., Fekete, B., Crouzet, P., Döll, P., Endejan, M., Frenken, K., Magome, J., Nilsson, C., Robertson, J. C., Rödel, R., Sindorf, N., and Wissler, D.: High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management, *Front. Ecol. Environ.*, 9, 494 – 502, <https://doi.org/10.1890/100125>, 2011.
- Messenger, M. L., Lehner, B., Grill, G., Nedeva, I., and Schmitt, O.: Estimating the volume and age of water stored in global lakes using a geo-statistical approach, *Nat. Commun.*, 7, 1 – 11, <https://doi.org/10.1038/ncomms13603>, 2016.
- Mulligan, M., van Soesbergen, A., and Sáenz, L.: GOODD, a global dataset of more than 38,000 georeferenced dams, *Sci. Data*, 7, 1 – 8, <https://doi.org/10.1038/s41597-020-0362-5>, 2020.

- Pringle, A. M., Handler, R. M., and Pearce, J. M.: Aquavoltaics: Synergies for dual use of water area for solar photovoltaic electricity generation and aquaculture, *Renew. Sustain. Energy Rev.*, 80, 572 – 584, <https://doi.org/10.1016/j.rser.2017.05.191>, 2017.
- Robinson, C., Ortiz, A., Kim, A., Dодhia, R., Zolli, A., Nagaraju, S. K., Oakleaf, J., Kiesecker, J., and Ferres, J. M. L.: Global Renewables Watch: A temporal dataset of solar and wind energy derived from satellite imagery, *arXiv*, arXiv:2503.14860, <https://doi.org/10.48550/arXiv.2503.14860>, 2025.
- Wang, J., Walter, B. A., Yao, F., Song, C., Ding, M., Maroof, A. S., Zhu, J., Fan, C., McAlister, J. M., Sikder, S., Sheng, Y., Allen, G. H., Crétauх, J.-F., and Wada, Y.: GeoDAR: Georeferenced global dams and reservoirs dataset for bridging attributes and geolocations, *Earth Syst. Sci. Data*, 14, 1869 – 1899, <https://doi.org/10.5194/essd-14-1869-2022>, 2022.