



Supporting offshore wind growth: Automating data analysis in digital aerial surveys to enhance wildlife protection and survey efficiency

Ben Bartlett^{a,b} , Matheus Santos^{b,c}, Petar Trsljic^{b,c}, Gerard Dooly^{b,c}

^a Electronic and Computer Engineering Department, University of Limerick, Castletroy, Limerick, V94 T9PX, Co. Limerick, Ireland

^b CRIS, Centre for Robotics and Intelligent Systems, University of Limerick, Castletroy, Limerick, V94 T9PX, Co. Limerick, Ireland

^c School of Engineering, University of Limerick, Castletroy, Limerick, V94 T9PX, Co. Limerick, Ireland

ARTICLE INFO

Keywords:

Environmental Impact Assessment (EIA)
Automated screening systems
Machine learning (ML) in ecology
Marine mammal and avian surveys
Real-time ecological monitoring
Baseline survey accuracy

ABSTRACT

With Europe projected to install 260 GW of new wind power between 2024 and 2030, much of it offshore, efficient Environmental Impact Assessments (EIAs) are essential. Regulations require 24 monthly aerial digital surveys before development, with continued monitoring during and after construction. This generates massive volumes of ecological data. We present an automated system that drastically reduces the time required for the most labour-intensive task: screening imagery to identify objects or individuals for further species classification. The process is reduced from several months to the 4-hour survey duration. In a 15-month case study (with one month excluded for testing), the system achieved 97.9 % accuracy, outperforming manual screening (68.75 %), and eliminated 99.13 % of frames from requiring manual review. Avian detection matched manual performance but remained limited by current survey conditions and image resolution. Critically, we found that the commonly assumed 2 cm ground sampling distance (GSD) was inconsistent across survey frames, with no part of any image achieving 2 cm/px, due to camera angles and aircraft configuration. This reduces classification confidence and highlights a need for improved data standards and transparency. As the first study to directly examine these assumptions using raw data, our results demonstrate that survey resolution is insufficient for consistent species identification, and that manual screening may miss up to 30 % of individuals. These findings underscore the importance of questioning inherited data assumptions and improving survey methodologies before such outputs are used to inform policy or conservation action.

1. Introduction

1.1. Offshore wind expansion and policy targets

As of 2023, Europe now has a total of 272 GW of wind capacity, with 18.3 GW of new wind power capacity having been installed, with the EU-27 contributing 16.2 GW of this total. This is a record for the region (WindEurope, 2024b). However, this figure represents only half of the capacity needed to meet the EU's 2030 climate and energy targets. Notably, 79% of the new wind capacity was onshore, while offshore installations reached a record 3.8 GW. This creates a cumulative total of 19.38 GW of offshore wind in the EU as of 2023 (European Commission, 2024). Despite this growth, 2/3 of the wind installations expected up to 2030, will remain onshore. Looking forward, Europe is projected to install 260 GW of new wind power capacity between 2024 and 2030, with the EU-27 responsible for 200 GW of this total—an average of 29 GW per year. To meet its 2030 climate and energy goals, the EU needs to ramp up its installation

rate to 33 GW per year. As a result, the EU is likely to fall short by approximately 30 GW of its 425 GW ambition required to achieve a 42.5% renewable energy target (WindEurope, 2024b).

1.2. Survey requirements and ecological constraints

The European goal for offshore wind by 2030 is 111 GW with 317 GW required by 2050 (European Commission, 2024). At the current rates of development this target will never be achieved. While the planning, permission, development and construction stages of offshore wind is a multifaceted orchestra, there are numerous bottlenecks. One such bottleneck is the preconstruction survey requirements during the consenting processes. Balancing the creation of clean, renewable energy and protecting the marine environment is crucial. Environmental Impact Assessments (EIAs) involve the collection of data across multiple domains, including Metocean, Geophysical, Geotechnical, Biological, Socioeconomic, and Cultural Heritage. These are all collected with a goal of creating baseline information about the region of potential

* Corresponding author at: CRIS, Centre for Robotics and Intelligent Systems, University of Limerick, Castletroy, Limerick, V94 T9PX, Co. Limerick, Ireland.
E-mail addresses: ben.bartlett@ul.ie (B. Bartlett), matheus.santos@ul.ie (M. Santos), petar.trsljic@ul.ie (P. Trsljic), gerard.dooly@ul.ie (G. Dooly).

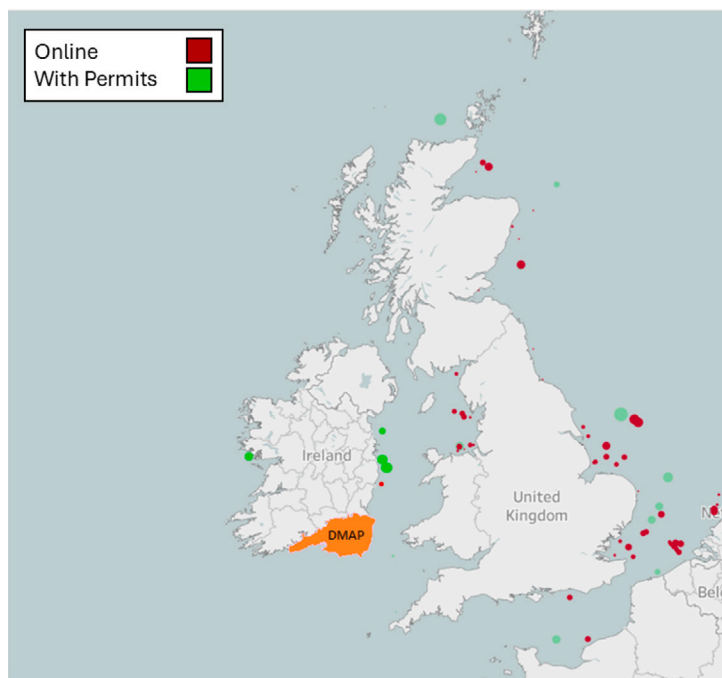


Fig. 1. Planned and operational offshore wind farms in the UK and Ireland, including the DMAP location, with overview statistics for currently connected and permitted offshore wind capacity in Europe (WindEurope, 2024a).

Europe's Offshore Wind

- 34,166 MW connected
- 6,340 turbines online
- 135 wind farms in 13 countries

Country	MW	Turbines
UK	14,742	2,767
Germany	8,464	1,566
Netherlands	4,739	670
Denmark	2,652	672
Belgium	2,261	399
France	1,202	175
Norway	101	13
Sweden	192	80
Finland	71	19
Portugal	25	3
Ireland	25	7
Italy	30	10
Spain	2	1

development. Similar data is collected post construction. While an important part of the preplanning of offshore wind, they do not currently carry a legal requirement to actually avoid or reduce impacts, nor compensate for residual impacts (Institute for European Environmental Policy (IEEP), 2022).

1.3. Role of digital aerial surveys in offshore EIAs

Digital aerial surveys have become the preferred method for ecological monitoring of marine megafauna (e.g., seabirds, marine mammals), due to their speed, consistency, and cost-efficiency compared to traditional manned surveys. This is largely due to the ability to collect and post process the data (Zydalis et al., 2019; NatureScot, 2023). They are also much more repeatable and transparent for as long as the digital images are retained (Weiser et al., 2023). These surveys are primarily used to assess the distribution and abundance of species visible at or above the sea surface, such as seabirds, marine mammals, and some large pelagic species. Since around 2010 (Buckland et al., 2012), digital aerial surveys have been growing in popularity, more recently becoming a prevailing method of survey within the offshore wind industry. Companies such as HiDef, a market leader in this field, have conducted over 800 flights since 2014, generating extensive historical and ongoing datasets (HiDef Surveying Ltd, 2024; Bioconsult SH, 2024). The success of offshore wind energy in regions with the largest cumulative capacity as of 2024, such as the UK, Germany, the Netherlands, Belgium, and Denmark, may be partly attributed to the effective use of these digital aerial surveys enabling this growth (HiDef Surveying, 2023; Hedef Surveying, 2023).

The United Kingdom, with a much longer history in offshore wind than Ireland, is aiming to expand its 22 GW of offshore wind energy to 154 GW by 2030 (Mitchell et al., 2022). Ireland held its first offshore wind auction in 2023 under the Offshore Renewable Electricity Support Scheme (ORESS) 1, beginning its journey to 5 GW of offshore wind by 2030 (Irish Government, 2020). Almost 40 GW of offshore wind projects are in development in Ireland, of which more than 10 GW are in the Atlantic Region, and nearly 4 GW of schemes given “relevant projects” status to allow fast track development (Development, 2022). The winning bids were the North Irish Sea Array (Statkraft, 500 MW),

Dublin Array (RWE and Saorgus Energy, up to 850 MW), Codling Wind Park (EDF and Fred Olsen, up to 1450 MW) and Sceirde Rocks (Corio Generation, 450 MW) as shown in Fig. 1. Three wind farms will be built off the East Coast and one off the West Coast. The first projects could be operational by 2028, depending on the timelines of their planning permission. None of the successful projects have yet been able to apply for planning permission (Development, 2022; WindEurope, 2022).

1.4. Regulatory changes in Ireland and data access opportunities

A recent ORESS 2.1 Industry briefing session, held on 15 January 2024 (Government of Ireland, 2024), highlighted the increasing importance of data in supporting offshore wind development. At the time of writing, a single bid for a 900 MW site is expected to be approved through this process. One of the session's major themes was data acquisition and management. Such data are critical not only for marine spatial planning and reducing development costs, but also for establishing ecological baselines, assessing potential biodiversity impacts, and informing mitigation strategies. The session outlined several key conclusions, including the need for improved data collaboration, standardization, and accessibility. It also emphasized the importance of designing ecological surveys that produce robust, repeatable datasets, while balancing the impact of survey activities on marine life and fisheries. Despite these goals, several challenges persist—most notably, inconsistent standards, limited accessibility, and industry reluctance to share data (Government of Ireland, 2024).

Traditionally, Environmental Impact Assessments (EIAs) were completed by developers applying for development consent. However, Ireland is now shifting this responsibility from private developers to central government authorities. They have agreed that development in ORESS 2.1 will take place within Designated Maritime Area Plans (DMAPs), established according to the Maritime Area Planning Act, 2021. Future developments will be located at specific sites within DMAPs identified for this purpose by the Government (eTenders Government of Ireland, 2024). The first DMAP for ORE will be located off the South Coast, as shown in Fig. 1. As part of this initiative, EIA data is currently being purchased from historically completed surveys or newly commissioned with the partnerships of the Marine Institute (MI)

and Geological Survey Ireland (GSI). Four zones have been proposed within this DMAP for development as of May 2024 (Department of the Environment, Climate and Communications, 2024).

1.5. Cost implications and scalability limits

This purchasing and commission of data acquisition in Ireland presents a new opportunity to begin creating solutions to the challenges identified. With a shift in data analysis away from commercial and industrial partners towards governmental and public bodies, a revisiting of this already collected ecological data is currently taking place to determine the optimal locations to build this infrastructure in Ireland but also validate the methods that have been utilized to date. Based on independently collected figures as a part of this study from a digital aerial survey company in Ireland, a single flight can cost around €20,000, with 24 being completed pre-construction - one a month for two years. The data screening and species identification costs approximately an additional €11,250 per flight with a quoted 600 h allocated to screening and 100 to classification for the 24 flights, cumulatively. This equates to approximately €750,000 for the aerial ecological survey component of an offshore wind EIA, consistent with values reported elsewhere of up to €1,000,000 per gigawatt-scale project (BVG Associates, 2024).

While the figures we received from survey providers were broad and not from the survey provider of the data used in this study, they align closely with those reported in industry publications and informal estimates. However, to the best of our knowledge, no provider has publicly disclosed an exact cumulative breakdown of screening hours or per-project costs. As a result, it remains impossible to precisely quantify the full labour burden or the true economic benefit of automation. In this paper, we make transparent, best-effort estimates based on partial disclosures and publicly advertised screening wages to highlight the scale of the challenge. Although the precise savings achieved by our system cannot be calculated definitively due to this lack of transparency, they are undeniably substantial—especially when considered across dozens or hundreds of surveys needed for large-scale offshore development.

The screening of 150,000–250,000 images per flight is labour-intensive, and cost estimates suggest up to 600 h of manual effort per survey. Past job listings have advertised screening roles at rates at €9/h (HiDef Surveying, 2024) in 2018, with more recent advertisements at annual salaries around £24,800, suggesting a workforce of approximately 30–35 individuals would be utilized within the cost estimates provided. This reinforces the unsustainable nature of current methods and highlights why automation is not just beneficial, but essential.

The lack of standardization and transparency in current survey practices presents additional challenges. Without clearer data structures or oversight, it is difficult to assess the validity of outputs that directly influence ecological risk assessments and conservation decisions.

There is a clear need for automated processing methods (Ditria et al., 2022; Kemper et al., 2016), the creation of common standards (Kraus et al., 2019) and an evaluation of the current state of the art (Diederichs et al., 2008; Bailey et al., 2014). There is significant untapped potential in the current surveying practices, particularly for the analysis of large avian and smaller mammalian species, with considerable room for improvement (Courbis et al., 2023; Largey et al., 2021; Conkling et al., 2021).

In this paper, we have developed a system to screen survey data in near real time, using raw video data of surveys completed off the Irish coast by HiDef, provided by a developer in offshore wind. This system uses commercially human-labelled CSV data from past surveys to create a Convolutional Neural Network (CNN) model that is rapidly expandable through the use of active learning and compatible with the hardware and software platforms currently used by major commercial companies. Sinclair and Booth (2019) and Vilela et al. (2020). This system has demonstrated significant efficiency improvements converting parts of the process from months to hours.

2. Literature review

2.1. Digital aerial survey

To design automation systems that meaningfully improve on current practice, it is essential to understand the operational structure of DAS as currently implemented in the offshore wind industry. Digital aerial surveys have become a crucial method for EIA, particularly in monitoring wildlife, such as avian and non-avian species, in marine environments. Since their introduction, the fundamental methodologies have remained largely unchanged, focusing primarily on two main types: digital video and digital still photography (High Resolution, 2009; Thaxter et al., 2016). Each method has its distinct advantages, determined largely by the camera systems used and the specific requirements of the surveying organizations.

Digital still photography, often employed by companies such as APEM and IfaÖ, typically operates at a frame rate of 0.5 to 1 frame per second (fps). This method is characterized by a significant overlap between images, sometimes up to 50%, which aids in capturing comprehensive coverage of the survey area (Webb and Nehls, 2019). On the other hand, the digital video survey method, branded as the “HiDef method”, exclusively utilized by HiDef Aerial Surveying Ltd, operates at a higher frame rate of 5 to 7 fps (Webb and Nehls, 2019). This increased frame rate allows for the capture of more dynamic scenes and potentially more detailed observations of moving objects, such as birds in flight, but produces more images for analysis.

2.2. The HiDef method

The HiDef method typically employs twin-engine high-wing aircraft, such as the Vulcanair P68 or Diamond DA42 (Webb and Nehls, 2019). These aircraft are chosen for their low fuel consumption and long-range capabilities, making them well-suited for extensive surveys (Wei et al., 2016). The method uses an array of four cameras, where cameras 2 and 3 have a swath width of 129 m, and cameras 1 and 4 have a swath width of 143 m. To prevent double counting of individual species, a 20 m gap is maintained between the ground coverage areas of adjacent cameras, as illustrated in Fig. 2. This gap is crucial due to the higher frame rate of video surveys, which increases the likelihood of a species moving from one camera’s field of view to another while the aircraft is overhead. However, the configuration can be adjusted to eliminate gaps or create some overlap, though this would reduce the overall swath width (Webb and Nehls, 2019).

Additionally, as seen in Fig. 2, the camera array is tilted 30 degrees from Nadir (the point directly beneath the aircraft on the Earth’s surface), either in the direction of flight or against it, to mitigate issues with sun glare (Wei et al., 2016). A nominal GSD of 2 cm is also chosen for these sensors (Webb and Nehls, 2019) utilizing a nominal flight height of 549 m. The GSD, which is the distance between the centre of each pixel on the ground, defines how much of the ground is captured in each image, with a lower GSD value indicating higher resolution.

The aircraft flies parallel transects, with the overall direction of these transects determined by the specific survey type. Often, these transects are oriented perpendicular to the coast, a strategy that enhances the capture of gradual water changes (Wei et al., 2016). For baseline ecological surveys related to offshore wind projects, regulations mandate a minimum coverage of 10% of the study area (Aumüller et al., 2023), which is commonly achieved by spacing the parallel transects 2–5 km apart (Żydulis et al., 2019; Buckland et al., 2012; Wei et al., 2016). This spacing minimizes the likelihood of encountering the same wildlife more than once during the survey, which can occur if the transects are too close together (Buckland et al., 2012). Striking a balance between the accuracy of results, flight time, and area coverage is crucial, as closer transect spacing increases the risk of double-counting species, while wider spacing may reduce overall accuracy and increase costs.

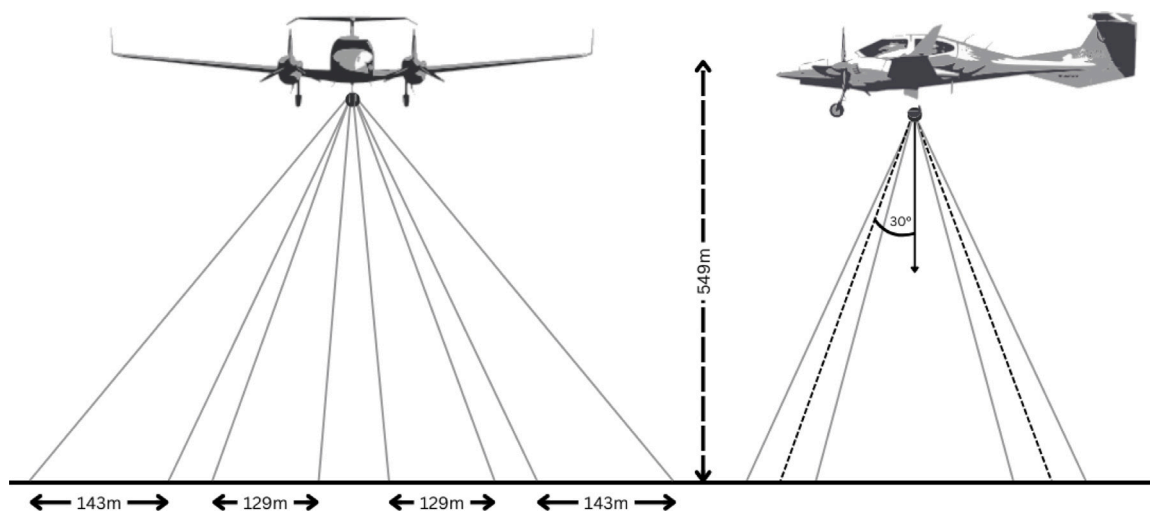


Fig. 2. Survey aircraft camera configuration in the HiDef method.

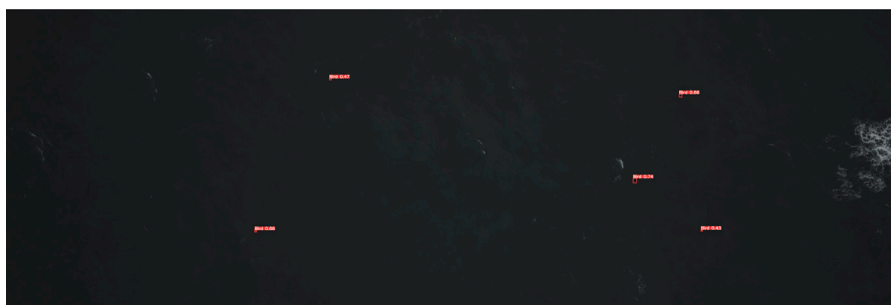


Fig. 3. Sample AI-screened video frame from our developed process.

Aircraft avionics data, including GPS position, compass heading, and barometric altitude, is recorded at a rate of 1 Hz in parallel to the camera frames. This means 6–7 frames receive the same positional data. With interpolation an intermediate position for each frame can be estimated. The lateral offsets of the camera swaths are then used to georeference bird and mammal observations. All observations in a frame receive the same position location corresponding to the centre of the frame. Flights are restricted to certain conditions such as no cloud cover below the flight altitude, no precipitation, wind speeds of less than 30 km/h at sea level, sea state of 6 or less and not less than 1.5 h after sunrise or before sunset (Collier et al., 2022).

2.3. Limitations of manual screening and QA workflows

Understanding how manual screeners process DAS imagery is essential to identifying where automation may yield the greatest efficiency gains and correct for known detection issues. After the survey is complete, the data is screened offline. The video is made available to the reviewers, firstly marking any objects for re-review by specialist ornithologists and marine mammal specialists to classify. A sample frame is shown in Fig. 3. There is no mention in the literature or guidance that zooming is used during this process. Based on our own manual analysis of full transects, substantial undercounting was observed—especially in frames with dense bird presence. Across multiple randomized frames, screeners routinely missed individuals, particularly when more than 10 birds were present. In one example, 77 birds were manually counted by the authors, while only 66 were counted in the official commercial dataset—a shortfall of nearly 15%. In another, only 3 individuals were annotated in a frame where at least 10 birds were visible, a short fall of 70%. These discrepancies were most pronounced for small or tightly clustered species, reinforcing the

reviewer's observation that the lack of zoom or enhanced viewing likely results in missed detections. This is further supported in our results section, where the AI model outperformed manual screening by nearly 30% for all mammals in a survey, highlighting that manual screening is not only time-consuming but also systematically underdetects species. This strongly supports the case for automated screening solutions in large-scale offshore surveys.

The quality assurance (QA) involves a second reviewer, the auditor, to review 20% of the data initially reviewed by the first reviewer. If there is not over 90% agreement between both reviewers, corrective action is taken. The remaining data is reviewed and where appropriate the failed reviewer's data is discarded. The QA process not only ensures accuracy but also serves as a training mechanism, allowing reviewers to continually coordinate and compare their work, thus improving consistency and expertise within the team (Wei et al., 2016; Pavat and Harker, 2022).

If possible, during the initial screening, a broad category is recorded in the form of bird species, mammal species or man-made object, allowing the sorted data to be passed to the ornithologists and mammal specialists. Their efforts are best utilized on a sorted subset of the data since the vast majority of frames will have no objects to identify. The specialists classify the detections to the lowest taxonomic level possible and assess the approximate age and sex, as well as any behavioural traits visible in the image. This can include, sitting vs. flying, what direction the bird is flying, if a mammal is surfacing or submerged. This part of the process also entails an identical QA process where 20% of the data is randomly selected for review with at least 90% agreement required between the specialists. If disputes still exist a third reviewer makes the final decision (Žydelis et al., 2019; Pavat and Harker, 2022).

It should be noted that confidence ratings assigned during species identification are not standardized across taxonomic groups. According

to survey documentation from NECR439 (Pavat and Harker, 2022), the likelihood of achieving a definite or probable identification is not consistent for all component members of a species group. For example, someone undertaking identification of a large auk will find it easier to be confident of guillemot (*Uria aalge*) identification than razorbill (*Alca torda*). Definite was defined as certain as reasonably possible. Probable was defined as very likely to be this species or species group. Possible was defined as more likely to be this species or species group than anything else. Any animals that could not be identified to species level were assigned to a category 'No ID'. If, on occasion, the unidentified bird is suspected of belonging to two possible groups, then a broader group category may be used. Interestingly, this phrasing aligns not only with the published report but also with the internal labelling used in commercial datasets, where "guillemot" is used generically to refer to *Uria aalge*. This reinforces a broader issue—namely, the lack of precision and consistency in classification terminology across both public-facing documentation and operational practices. The interchangeable use of common names without clear disambiguation contributes to confusion in downstream analysis and underscores the need for standardized species identification protocols.

All of this information is stored in a CSV format with examples publicly available (High Resolution, 2009; Marine Data Exchange, 2024). A sample is also shown at the bottom of Fig. 3. The CSV data consists of a number of headings. These are: Survey, Survey date, Camera, Resolution, Reel Ref, Time, Frame Ref, Broad Category, Species Group, Confidence, Species, Confidence, BTO code, behaviour, Submerged, Flight Height, Flight Height Confidence, Age, Sex, Measurements, Additional Comments, Lat and Lon. This has largely been the way the data has been compiled since 2009 (High Resolution, 2009). All classifications are finally outputted as a single output, in ArcGIS shapefile format, using UTM30N projection, WGS84 datum (Pavat and Harker, 2022). After this a series of data analysis techniques are applied for estimating populations, taking bias into consideration and creating density maps.

2.4. Evidence of systemic gaps in survey accuracy and transparency

As highlighted in Section 1, the adequacy of current data collection and screening methods for offshore wind monitoring has been questioned extensively. This concern was notably raised during the ORESS 2.1 industry briefing session in Ireland, where significant doubts were expressed regarding the effectiveness of existing approaches. However, this issue is not unique to Ireland. Similar concerns have been echoed in the broader literature, reflecting widespread apprehension about the current state of data collection and analysis in the field.

While guidance exists for ORE developers in Ireland, particularly for aerial surveys, this guidance is in need of updating to meet current technological and environmental requirements (Department of the Environment, Climate and Communications, 2020). Numerous sources in the literature emphasize these concerns. For instance, recent reviews and reports have scrutinized the sufficiency of data collection and screening methods, calling for improvements (Courbis et al., 2023; Largey et al., 2021). A recent final project report specifically assessing the monitoring of birds and marine mammals in offshore wind environments highlights critical issues, including the need for automation, standardization, and transparency in dataset collection and evaluation (Courbis et al., 2024). The report underscores that high-quality, long-term datasets collected through robust and cohesive methodologies are essential for effectively informing adaptive management of the offshore wind industry. In contrast, disjointed and poorly designed individual efforts fall short in advancing our understanding of OSW impacts (NYETWG, 2024; New York State Energy Research and Development Authority, 2021; Wilding et al., 2017; Bureau of Ocean Energy Management (BOEM), 2017).

One significant observation is that while large datasets have been gathered, they do not necessarily improve the understanding of OSW

effects. This gap highlights the need for a more systematic approach to data collection and analysis (Courbis et al., 2024; NYETWG, 2024). Such a shift would enable the generation of more meaningful insights into the environmental impacts of offshore wind farms, thereby supporting more effective management and conservation strategies.

2.5. Automation in EIA

Despite major advances in AI and computer vision, automation remains largely absent from commercial Environmental Impact Assessment (EIA) workflows. This section reviews the limited efforts to date, critiques their scope, and contextualizes our contributions in response.

While automated wildlife detection systems have been actively developed for videos, satellite imagery, and ecological datasets more broadly (Xu et al., 2024; Hong et al., 2019; Berg et al., 2022; Ke et al., 2024; Borowicz et al., 2019a; Chabot and Francis, 2016), they have yet to be meaningfully adopted in the commercial EIA sector. The reasons for this are structural rather than technical: industry datasets are rarely made available for academic development, and the sector continues to rely heavily on manual labour, in part due to concerns over proprietary methods and competition.

Emerging alternatives to traditional aerial surveys – such as satellite- or UAV-based approaches – offer promise for future automation (Cubaynes and Fretwell, 2022; Borowicz et al., 2019b; Fiori et al., 2017), but require further development and regulatory alignment. Accordingly, our focus is not on introducing new data collection methods, but on critically evaluating and improving the workflows already in place.

Only one study in the literature directly addresses automation in the EIA context. Kuru et al. introduced WILDetect, an intelligent platform developed using data provided by APEM (Kuru et al., 2023b,a). The system employs an ensemble of supervised machine learning and reinforcement learning techniques to classify avian detections. However, the study is best understood as a conceptual demonstration rather than a deployable solution. The model is limited by dataset size – only 1073 instances of gannets were available, later augmented to 4292 – and was developed using outdated GPU infrastructure, with reported training times exceeding 50 days. More critically, WILDetect automates only the final classification step, bypassing the initial data screening stage that accounts for the majority of effort in EIA workflows. This narrow focus prevents it from addressing the true bottleneck of commercial survey pipelines. While the work is a valuable academic exercise, it does not yet offer a viable pathway for operational deployment.

While efforts such as WILDetect represent progress, they also highlight a broader limitation: collaboration between academia and commercial providers remains superficial. In most cases, companies provide only processed outputs or subsets of data, without transparency about how labels were derived, how quality was assured, or what classification confidence means in practice.

This lack of transparency presents a fundamental challenge. Once data and methods are used in a regulatory or research context, they must adhere to scientific standards of reproducibility, openness, and methodological clarity. If commercial confidentiality restricts the disclosure of underlying methods, then such findings – however useful internally – should not be treated as peer-reviewed evidence.

In our study, although the primary goal was to develop an automated screening pipeline, we gained rare access to raw digital aerial survey data from a major provider. This enabled a more thorough audit of current workflows than has previously been possible. Our analysis identified undercounting of detections, inconsistencies in classification confidence, and misalignments between annotations and raw imagery. These issues are not well documented in existing literature, yet they directly impact the reliability of ecological baselines used for offshore wind consenting.

By bridging technical development with methodological critique, this study contributes not just a tool, but a framework for evaluating

the robustness of current EIA practices. Our results support the case for automated methods—not only for their efficiency gains, but also as a route towards greater transparency, consistency, and scientific integrity in ecological monitoring.

3. Methodology

3.1. Data overview

The dataset analysed in this study was generated using the HiDef digital video survey method. The raw video data was commissioned and provided by a developer in offshore wind for the purpose of exploring automated data processing techniques. To the best of our knowledge, this is the first study of its kind as access to this data has proved challenging. Unlike previous studies that often rely only on CSV outputs or selectively published image subsets (Marine Data Exchange, 2024; Vallejo et al., 2017), this research delves into the initial raw data from a detection and screening perspective, taking a close look at the process that produces bio-census information. Rather than serving as a procedural walkthrough, this section outlines the unique technical challenges encountered and our approach to enabling scalable, reproducible analysis. The focus on the provided data is central to this study, as it provides a comprehensive understanding of the potential for automation in processing large volumes of aerial survey data.

This data includes raw, compressed video files in '.seq' format, GNSS data in '.txt' format, and '.csv' files containing the final classified outputs from the aerial survey provider. Given this is an independent study, we developed a system to process the data through detailed analysis and reverse engineering.

It was estimated that the camera data for this project and likely for other surveys of the same type, was captured using NorPix's turnkey airborne video recording station, a system capable of recording multiple camera streams at resolutions up to 50MP with GNSS timestamping and real-time compression (Norpix, 2024). The video files, recorded using StreamPix, can be converted to '.png' or '.jpg' using NorPix's batch processing software. However, this conversion process can significantly increase storage requirements if not completed sparingly. It is also not practical to implement within an automated pipeline due to interface limitations.

3.2. Software overview

To interpret and process the raw survey files, a clear understanding of their underlying structure is required. Each '.seq' file contains a 1024-byte header, followed by JPEG-compressed image data which consists of 3 parts. The first 4 bytes contain the JPEG image data size in bytes +4. This is followed by the actual JPEG image buffer and 8 bytes allocated to the timestamp. Associated with each '.seq' file is an '.idx' file. Given each image is compressed it may occupy a varied amount of data in the '.seq' file. As such the '.seq' file must be read sequentially to determine the length of each image. The associated '.idx' file contains the byte offset for each image which allows indexing to any image in the '.seq' file without sequentially processing. The header information revealed that Huffman lossless compression was used during recording.

While decoding uncompressed '.seq' files is straightforward, extensive effort was required to decode the Huffman-compressed JPEG buffers. This was accomplished using the legacy Intel Integrated Performance Primitives (Intel IPP) libraries, enabling dynamic, targeted image extraction in real-time from large databases of '.seq' video files. This capability is a critical component of our workflow.

In contrast to WILDetect (Kuru et al., 2023b), which relies on APEM's manually curated snag library – a commercial dataset of cropped images with labelled wildlife instances – and is focused on developing new methods for future use, our approach retroactively constructs a comparable dataset directly from raw survey materials. By integrating full-resolution frames with associated detection metadata,

we generate a snag-style dataset in an automated and scalable manner. Our method is designed to incorporate both historical and new data, enabling continuous validation and integration without disrupting existing workflows.

To the best of our knowledge, no automated method currently exists for generating snag datasets from full-resolution survey video. In the absence of access to commercial batch tools or input from providers, we independently developed a reverse-engineered pipeline capable of aligning labelled metadata with raw frames to extract cropped object regions at scale. This system enables the automated generation of snag libraries across multiple terabytes of historical and ongoing survey data—paving the way for scalable, repeatable model training without disrupting existing workflows.

3.3. Historic data translation to ML dataset

3.3.1. Dataset creation through dynamic batch processing

The human-in-the-loop system (Fig. 4) illustrates the core workflow structure underpinning data preparation and model training. The first step to translating the information into an object detection ML model format is to convert the CSV information and Compressed Video files, to image files such as JPEGs. These are then paired with label files (e.g., .txt or .xml) formatted for downstream ML frameworks. The CSV data consists of a number of headings as mentioned in Section 2. During the manual screening process a broad category is recorded in the form of avian, non-avian and manmade object. As such this should be replicated.

The CSV data is filtered by the manual chosen category and a series of sub-directories in the ML database are created, populated with the video file names and frames with an instance of the chosen category. This preparation supported efficient downstream dataset construction. Given 15 surveys associated with one wind farm encompasses three 8 TB hard drives, with varying naming standards creating a robust parsing and organization of species across survey outputs is paramount for scalability. The option exists to choose subcategories with this being utilized to further subdivide the broad categories.

The dataset creation then involves extracting and decoding frames from the '.seq' video files. Using the parsed CSV information, the byte offset for each image is retrieved from the associated '.idx' file. Following this the JPEG buffers are decoded using Huffman decompression for this dataset. Huffman compression involves several steps, including frequency analysis, building a priority queue, Huffman tree generation, and data encoding. To decompress the data, the tree must be known. Substantial development effort focused on reverse-engineering the batch processing software to integrate it with historical datasets. This complex and challenging task involved detailed analysis and customization to ensure compatibility with the existing data format and structures. Compared to manual batch processing tools, this automated, targeted workflow substantially reduces overhead and ensures scalability.

This dynamic batch processing allows for the creation of an image dataset where large volumes of data can be selectively converted, such that each image is known to now at least contain one instance of the broad category class in it.

3.3.2. Labelling process

Because object positions are not encoded in the raw metadata, each image requires manual bounding box annotation. Any labelling software can be used to create label files for a variety of ML models. While some have developed tools specifically for ecological survey data (Kellenberger et al., 2020), many open-source options exist. We have utilized a customized version of Label Studio (Studio, 2024) due to the extensive API, the ability to manage multiple labellers, integration with cloud storage and the ability to connect ML models for real time predictions while labelling.

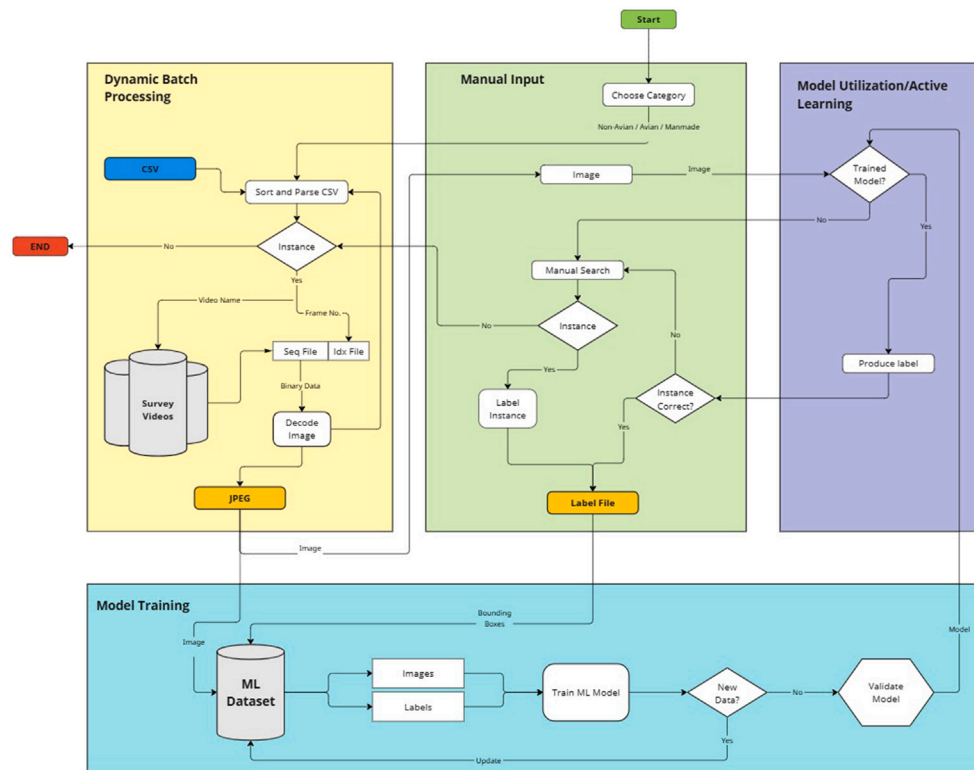


Fig. 4. Automated image batch processing with semi-automated image labelling.

The images are very large, with the dimensions of 6576×2192 pixels. Through experimentation, it was determined that to accurately screen the full image thoroughly, they must be split into 5×5 sub-images. The image is magnified to this resolution, equivalent to zooming in on just 4% of the frame at a time, and the images automatically scrolled through to be viewed by the labeller. On average, our labellers settled on spending 40 s per 6576×2192 image, assuming nothing of interest appeared, as this was the minimum time to comprehensively view the frame. This corresponds to an effective screening rate of just 0.025 Hz—far slower than the 1–3 Hz frame rate used in standard 50% split-screen review workflows. While we acknowledge that screeners in commercial workflows may focus only on objects near a red tracking line, it is not plausible that reliable screening can occur at just 0.33–1 s per frame. As such, it is clear why undercounting is routinely observed in commercial workflows.

3.4. Model creation and use of active learning

Upon the labelling of all of the data by creating bounding boxes for all instances, a model can be trained. The standard outputs for the labels are typically a txt or xml file, including a class number, bounding box coordinates and dimensions. The exact details of this are model dependant.

Once a preliminary model is trained, the model can be used to help label any additional data more efficiently. While the CSV files only list each instance of a detected species once, the collected data is digital video. As such an assumption can be made that the frames before and after the given frame number in the CSV file, will also contain instances relevant to the model for training. As such with the process can be repeated.

The dynamic batch processing method is applied again to the data, with an additional loop added to the frame decompression section. It was determined on average that the third to fifth time an instance was visible in the video was the frame number recorded in the CSV file. As such decoding five images before and after the listed frame means

for every frame listed in the CSV file, 1–5 others exist that should be labelled. While five images before and after the listed frame will lead to frames with no instances, this safety buffer should be utilized to generate the maximum dataset size from the 15 video surveys utilized in this study.

As shown in Fig. 4 the model can operate in parallel to the manual labeller once trained. The model receives the image at the same time as the labeller. The model infers potential bounding box locations on the image and plots them for the manual labeller to adjust, accept or remove. This active learning implementation improved labelling throughput by 95%, allowing screeners to process empty frames in 2 s—compared to 40 s without model assistance.

4. Results

4.1. Final datasets

The decision was made to create datasets for two classes. Class one was the Broad Category of Non-Avian Species which encompassed Species Groups of Dolphin, Cetacean, and Seal species with Species of No_ID, harbour porpoise, common dolphins and grey seals. This labelling structure reflects how the data was categorized in the original commercial outputs; examples are provided in Supplementary Table 2. Class two was chosen to be more specifically Northern Gannets as a subclass of the broad category of Avian Species.

Given the low resolution of the data and inconsistencies in commercial labelling, it was not feasible to create a general “Avian” class. For context, the Borssele report (Collier et al., 2022) documents 56,580 birds and only 1831 marine mammals across 25 surveys, underscoring the imbalance in frequency. Attempting to model all avian detections would have produced an overwhelming and highly imbalanced dataset, making it difficult to validate model performance or maintain annotation consistency for a system designed to screen the data, not classify. Instead, Gannet was selected as the avian target class. This decision was based on three factors: (1) Gannets are among the most

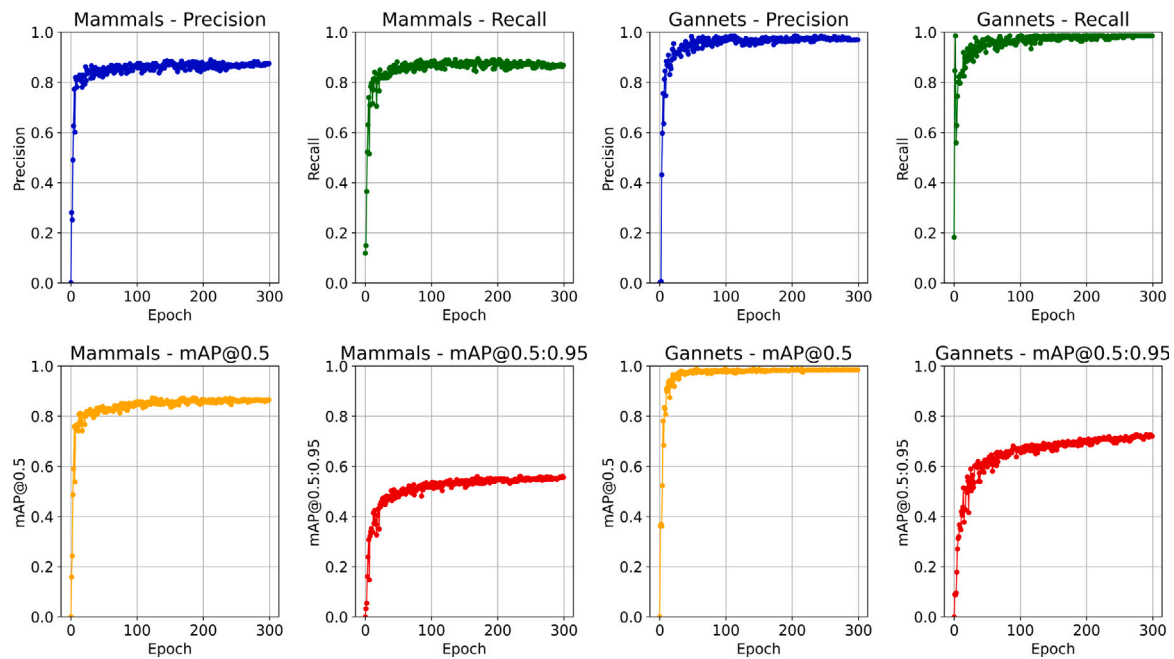


Fig. 5. Training evaluation metrics (Precision, Recall, mAP@0.5, and mAP@0.5:0.95) for the two trained models—Non-avian Species and Northern Gannet. X-axis shows training epochs; Y-axis shows metric score.

consistently labelled avian species in the dataset, (2) they are visually distinct and relatively large, making them more suitable for detection at low resolution, and (3) they offer a direct point of comparison with WILDetect (Kuru et al., 2023b), which also used Gannet as its avian example in snag-based screening. The classification provider used the label “Gannet” at both the group and species ID level (see Supplementary Table 3), further justifying its selection as a clearly defined and reliably annotated target class.

The initial dataset for non-avian species, based on CSV frame numbers, contained 375 instances across 275 frames. Using the active learning approach outlined in Fig. 4, this dataset was expanded to 2382 individually labelled instances. For the Northern Gannets dataset, a total of 1061 instances were achieved.

During training, with hyper-parameters enabled, the datasets were augmented such that each image was used to generate three additional randomly augmented images. Consequently, the training set for non-avian species grew to over 10,000 instances, while the Gannets dataset increased to over 4000 instances.

4.2. Selected model architecture and configuration

The object detection framework selected for this study was YOLOv5 (Jocher, 2020; Jiang et al., 2022), a single-stage real-time model known for its strong balance of detection accuracy, speed, and deployment flexibility. It remains widely used in applied vision tasks due to its ease of training and performance efficiency on high-resolution data. YOLOv5 was chosen because it is fast enough to support near real-time inference, enabling detections to be generated before an aircraft lands—thereby significantly reducing the time needed for downstream review by specialists.

YOLOv5 supports multi-scale detection and uses a CSPDarknet backbone with PANet neck, facilitating robust detection of small objects—a critical feature given that targets in aerial images often occupy less than 2% of the frame. While newer variants such as YOLOv7 and YOLOv11 were released during the development of this study, YOLOv5 was selected early and consistently used throughout all prototyping stages. Its performance was found to be operationally sufficient, with no practical justification to migrate pipelines mid-study.

As mentioned, a total of 15 surveys were used. One survey was entirely excluded from the dataset, with the remaining labelled data split into 70% training, 15% test, and 15% validation sets. An additional 10% of the total image count with no instances of class was added to provide background context to the model.

Given the high resolution of the imagery, it was decided – following Ultralytics guidance – to train at as close to native resolution as hardware allowed. The final training image size used was 2176 px, downsampled from 6576 px due to memory constraints on a Tesla A100 GPU. This corresponds to a compression factor of approximately 3. A batch size of 4 was used at this resolution, which, although small, proved sufficient for robust model convergence in all experiments.

4.3. Traditional training outputs

The results of the training for the mammal and gannet data are shown in Fig. 5. Despite limited optimizations or the application of high-resolution methods such as sliding windows or image segmentation, the test model results are extremely positive.

When assessing the training results of the model there are a number of important parameters. The results are in this case generated on the 15% validation set. Precision (P) of the model is the accuracy of the detected objects, indicating how many detections were correct. Recall (R) is the ability of the model to identify all instances of objects in the images. mAP50 and 95 give the average precision calculated at Intersection over Union (IoU) threshold, the quantification of the overlap between a predicted bounding box and the ground truth bounding box.

4.4. Cross-validation

This section presents the true generalization test of the models: performance on an entirely unseen, real-world dataset. While many studies rely solely on internal validation subsets (e.g., a 15% validation split), such approaches may not reflect practical performance. Here, we evaluate the model against a full-length survey withheld from all training and tuning stages, offering a robust and meaningful cross-validation.

Both models appear to perform very similarly on their individual validation sets. However, the excluded survey set serves as a clear

means to cross-validate the model in a directly comparable way to human screening, on new, entirely unseen data. This validation approach ensures that the model's performance is not only theoretical but also practically applicable for unseen data, enhancing the credibility of the results as often this is an overlooked detail. The results of running the model on this randomly excluded survey set are summarized in the confusion matrices in Fig. 6.

It should be noted that while 4 cameras exist per survey, only 2 cameras had been manually processed in the commercial workflow. The selected survey for validation was randomly chosen and, unfortunately, exhibited considerable overexposure in Camera 1, while Camera 3 showed clearer images. Despite this, the results are compelling and reflect a fair test without attempts to cherry-pick the most favourable conditions. As outlined in Section 2.2, all 15 surveys were conducted under controlled and standardized environmental conditions. As such, survey-to-survey variability is expected to be minimal, supporting the validity of using a randomly selected survey for model evaluation.

The detection parameters of 0.7 for confidence and 0.45 for IoU thresholds were used. The typical default setting is 0.25 for the confidence threshold. A balance between false positives and negatives must be made. These settings resulted in some false positives, including large grey avian species and man-made objects like buoys. Since the model's purpose is screening, it is preferable to have false positives rather than false negatives. As such utilizing the lower confidence levels while detecting minimizes the potential that non-avian species are missed. Further experimental work is needed to fine-tune these parameters for optimal performance.

The developed system parsed the 2 h and 25 mins of total video footage collected in this validation survey, in 3 h and 39 min. The hardware utilized for these tests while not optimal, is accessible when compared to the device used during model training. The CPU was an AMD Ryzen 7, with the GPU being a Nvidia GTX 1080. It is important to highlight that multiple cameras can be processed in parallel, allowing up to 4 cameras to be handled simultaneously within the same time frame as the system does not consume anywhere near the full CPU or GPU capacity when running. It should also be noted all data was decoded from HDD drives post capture which is slower than what the model is capable of achieving directly on the raw camera stream. The total flight duration for this survey from take-off to landing was 3 h and 36 min.

To ensure a fair evaluation, we manually reviewed every frame in the cross-validation dataset to establish an independent ground truth. This process was conducted frame-by-frame by the authors, using zoomed and magnified views in the same fashion as our manual labelling workflow. All detections – whether produced by the model or found in the commercial CSV – were assessed against this independently created benchmark. This is standard practice when evaluating detection systems: the ground truth is created manually, and then both the model and any existing annotations are compared against it. As such, the model and the commercial outputs were not assumed to be correct—they were evaluated against this curated reference.

In this test case, of the 66 726 frames captured by cameras 1 and 3, 581 total detections occurred. It should be noted that this is detections and not individual counts. The system in this case does not account for detecting the same mammal from frame to frame in the video. This is currently a manual step to group detections that occur in sequence into single counts for fair comparison with the commercial CSV data. Commercial screeners labelled 33 non-avian species in this survey, whereas the ML model labelled 47 non-avian species as shown in Fig. 6. The blue indicates 14 instances that were missed by commercial screeners but found by the ML model. These results were independently verified to be of the non-avian class, and an additional 3 detections that could not definitively be classified were excluded. As such the 47 individual mammals were counted in the 581 detections. This equates to approximately 228/581 detections that were correct, creating a final false positive rate of approximately one for every correct detection.

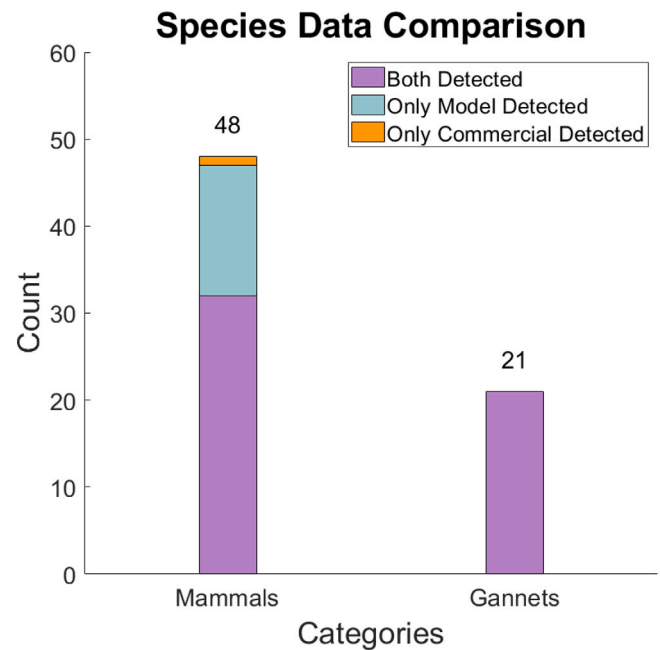


Fig. 6. AI model versus manual commercial comparison and cross validation. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

These false detections included some large grey avian species and man-made objects. However, the ML model screening process, in this case, eliminated 99.13% of the total data, with only 1 false negative almost within the flight duration, shown in Fig. 7. These 581 image snags are classifiable by one marine mammal expert within a day at a conservative figure.

In comparison while the results look very similar for the Gannet model, the reality of the low-resolution data is extremely apparent when trialled on this cross-validation dataset. It should be noted that the Gannet dataset generated from the 15 surveys is over 50% smaller than the mammal dataset in this paper. In this case two separate models were trained. It is preferable to have at least 10,000 instances of class.

In the 66,726 images, 83,207 total detections occurred. While all instances found by commercial screeners were found by the model, the limitations of the image resolution are clear. Sorting the detections by size, all 21 gannets – or approximately 126 detections – were found within the first 1000/83,207 detections. While more instances may exist, screening the 83,207 detections in their entirety was not feasible at this time. While the high-flying gannets are extremely distinct due to the higher GSD and larger size as they are closer to the camera, sitting on the sea they are no more than a group of white pixels. As such, a very large number of false positives exist, particularly for the overexposed Camera 1. This overexposure is unlikely to be caused by glare or sea state, but instead reflects internal camera settings and inconsistent calibration. Indeed, across multiple datasets, the two cameras not selected for commercial screening often showed poorer image quality—suggesting they may be excluded based on exposure or focus issues. These false positives include unknown objects, waves, other bird species, mammals, and man-made objects. An improved result may be possible by combining the non-avian and gannet classes into one model while also creating more classes such as Kittiwakes, Guillemots, and man-made objects to prevent excessive false positives. However, without improved resolution, the feasibility of this is significantly compromised.

5. Discussion

The results clearly demonstrate that the application of automation and machine learning to digital aerial surveys offers a promising and

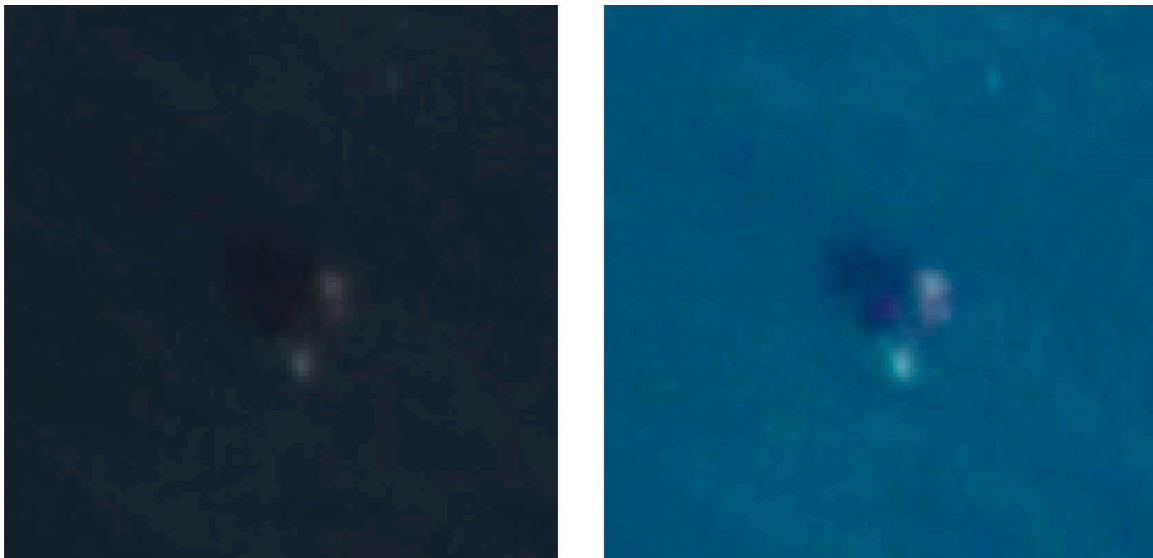


Fig. 7. Only missed instance by model that was labelled by HiDef - Seal Species Probable - NO_ID N/A. Note: The second image is an enhanced version of the original frame, with adjusted saturation, brightness, and contrast for visual clarity only. It is not representative of the raw survey data. Enhancements were applied solely for illustrative purposes.

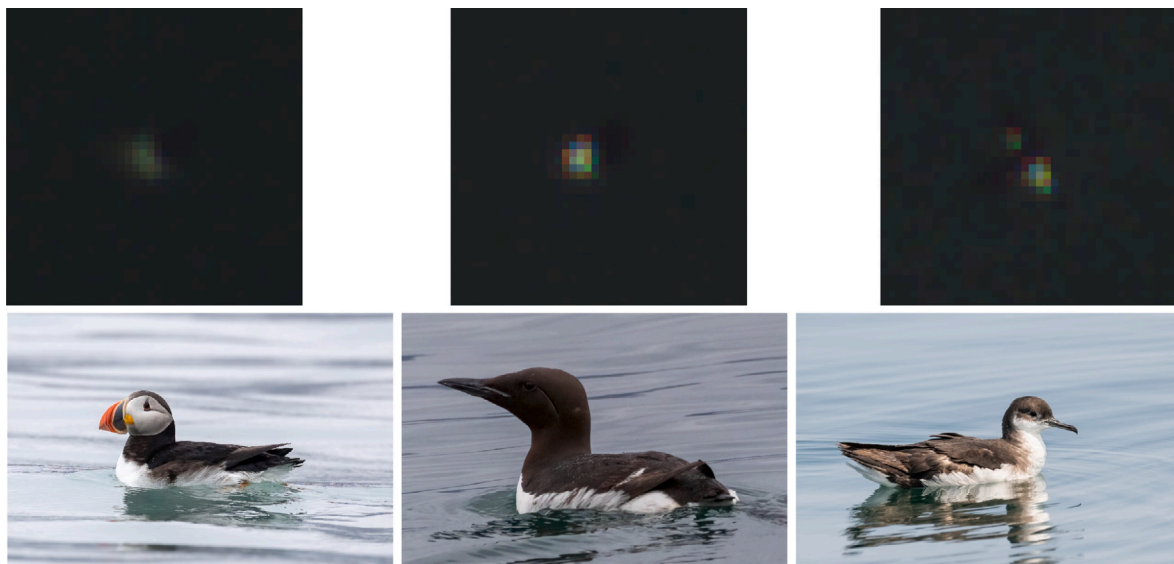


Fig. 8. Sample classified avian species: Left to Right—Definite Auk Species, Definite Puffin (**Fratercula arctica**) – Photo by Daniel Danckwerts (Rockjumper Birding Tours); Definite Large Auk, Definite Guillemot (**Uria aalge**) – Photo by Christine Jacobs; Definite Shearwater Species, Definite Manx Shearwater (**Puffinus puffinus**) – Photo by Linda Cunningham. All images © Cornell Lab of Ornithology—Macaulay Library.

effective solution for wildlife monitoring, particularly for mammals. The performance of the ML model for mammals not only surpassed traditional manual methods but is validated, proving its reliability and improved accuracy in identifying non-avian species in a time frame exponentially faster than human screening.

However, the case of avian species revealed a significant challenge: the existing ground truth data generated by manual methods is not consistent enough to support robust human validation, let alone to develop a robust AI model. The inconsistency in the manual classifications, especially when dealing with small and often indistinct avian species, led to the generation of many false positive predictions by the ML models, as what separates one species from another is not visually discernible in many cases. Representative samples of data are shown in Fig. 8. The top row includes cropped images from HiDef aerial footage, with classification and confidence levels directly sourced from the commercial CSV outputs (see Supplementary Table 4). The bottom row presents reference images from the Cornell Lab of Ornithology's

Macaulay Library, included to show what these species look like under ideal photographic conditions.

Each aerial detection spans approximately 10–15 pixels, representing just 0.001–0.002% of the full frame. These are not isolated examples—such image quality is typical for many labelled species beyond the two classes, non-avian and gannet species, in this study. This widespread limitation precluded our ability to construct additional species-specific classes (e.g., Puffin, Guillemot, Manx Shearwater), despite their inclusion in the commercial dataset, as no verifiable ground truth could be established from such pixel-limited data.

This finding highlights a critical issue: the current image resolution is insufficient for reliable results for avian species, by commercial process or otherwise. The limitations in resolution hinder the development of a comprehensive AI model capable of accurately identifying all species. Additionally, this issue underscores the unreliability of the ground truth data, which is crucial for both human and machine learning-based classifications. This inconsistency compromises species

detection accuracy, especially for avian species, raising questions about the effectiveness of EIAs that rely on this data and similar studies based on historical outputs (Garthe et al., 2023; Fox and Petersen, 2019; Heinänen et al., 2020). These studies should question the initial data used to produce the results they report on.

5.1. Survey parameters and methods

Our analysis affirms several points found in the literature. As discussed in Section 2 it was determined that the surveys were completed with a Vulkanair P68 flying in this case, transects spaced 2 km apart, perpendicularly to the coast. When dealing with sun glare it appears rather than being tilted from +30 degrees in pitch to -30 degrees the rig is rotated 180 degrees in yaw leading to high repeatability in having the camera angles remain at 30 degrees. However, it is never discussed what types of cameras are used, or what angles the 4 cameras are mounted at with regards to their roll angles. The only values discussed are a flight altitude of 549 m leading to a GSD of 2 cm/px and the swath widths of 129 m and 143 m as shown in Fig. 2.

Though not supplied, through thorough evaluation of the survey data, it was estimated that the cameras utilized in the 15 surveys were most likely Allied Vision Prosilica GT 6600C, based on video metadata. While many cameras now use CMOS technology with Sony discontinuing their CCD cameras in 2015 (Allied Vision Technologies, 2024a), this CCD sensor technology (CCD, 2024), still produces high quality images and is often utilized in applications such as fluorescence microscopy, high-resolution microscopy, astronomy and aerial mapping. However, CMOS technology is considered the dominant technology for Machine Vision applications due to their speed (DALSA, 2011) and with the mentioned discontinuation of manufacture, it is expected CCD will be replaced by CMOS technology for this type of application in the future, with the introduction of automated techniques such as those developed in this paper. The nominal GSD of 2 cm, though useful, must be contextualized given the camera's roll and pitch.

Given the known image dimensions of 6576×2192 px one of two options was configured on the cameras: a reduced region of interest (ROI) or vertical binning by a factor of 2. This accounts for the image height dimension of 2192 pixels, compared to the datasheet's stated image height of 4384 pixels. The typical frame rate in the .seq files was 6.51 frames per second. While it is impossible to determine exactly which option was used without confirmation of the survey parameters, but this frame rate matches the theoretical maximum for an ROI halved in image height, according to the user manual (Allied Vision Technologies, 2024b). This would be the preferred option, as enabling vertical binning involves horizontal row summing on the CCD before readout, thereby doubling the nominal ground sample distance horizontally (GSDh) from 2 cm/px to 4 cm/px. The GSDh is unaffected if the ROI is changed. It is typical to centre the ROI on the sensor, but the offset is user adjustable. It is worth noting that if it is not centred, this will also affect the GSDh, leading to further variation given the cameras are pitched to 30 degrees.

There is a question as to the exact benefit of higher frame rates when compared with digital photo stills when only half of the final data is screened. For this site, processing 2 of the cameras, equates to about 12%–14% coverage of the total area surveyed. There is also no clear rationale for camera selection, other than perhaps choosing the ones with the sharpest focus or least over exposed.

5.2. Ground Sampling Distance (GSD)

As mentioned in Section 2 many concerns have been raised in numerous reports about the quality of the data collected and as such what is determined from these reports. The GSD values of 2 cm are utilized for measurements such as species size (Sinclair and Booth, 2019) and flight height (Boersch-Supan et al., 2024; Humphries et al., 2023; Thaxter et al., 2016). This aligns with the information in the CSV

files used for this study also with GSD stated on every line of 2 cm. A category for measurements appears in the CSV, with the non-avian species exclusively having measurements. The average length given is around 1.2–1.4 m with our average non-avian detections by the A.I model at 60–70 pixels \times 2 cm GSD. Flight height while possible to calculate in many instances (Humphries et al., 2023) was not completed for any of the data in the CSV. However, a major oversight is evident within the GSD related calculations.

GSD by definition is the distance between pixel centres measured on the ground. While this is easily calculated for Nadir images (Bartlett et al., 2023), this is not the case for the camera configuration used in these surveys as none of the 4 cameras are pointed directly down. Due to the roll and pitch of the cameras, oblique project must be considered. Utilizing a simple pinhole camera model at the nominal flight height, sensor parameters from the datasheet (Allied Vision Technologies, 2022) and lens specification from the literature (High Resolution, 2009), we observe a theoretical nominal GSD of 2 cm (Borowicz et al., 2019a; Žydelis et al., 2019; Kemper et al., 2016; Humphries et al., 2023; Hatch et al., 2013) for the image height and width as seen in Eqs. (1) and (2) below:

$$GSD_w = \frac{AGL * S_w}{f * I_w} = \frac{549 \text{ m} * 36 \text{ mm}}{150 \text{ mm} * 6576 \text{ px}} = 2.0036 \text{ cm/px} \quad (1)$$

$$GSD_h = \frac{AGL * S_h}{f * I_h} = \frac{549 \text{ m} * 12 \text{ mm}}{150 \text{ mm} * 2192 \text{ px}} = 2.0036 \text{ cm/px} \quad (2)$$

Calibrated GSD is less frequently discussed. For example, Sinclair and Booth (2019) found improved measurements for mammal species with calibrated GSD considering variations in flight height from frame to frame, but the impact of camera angles on GSD variation is often overlooked. The cameras used in this survey are estimated to have roll angles of approximately 7.5 and 22.5 degrees. This results in varying GSD across the image frame. Given the cameras are rolled and pitched, we would expect a variation in GSD from pixel to pixel across the entirety of the image.

5.3. GSD validations

To accurately determine the variation in GSD across each image, in both width and height due to the non-nadir capture angles, a simulation was developed. Through this simulation, the camera angles were adjusted until a 20 m separation between the camera ground footprints was observed. This equated to roll angles of 7.7675 and 23.1748 degrees. While this may not exactly match the physical camera setup it aligns with the stated ground separation and highlights the issue. The separation is 20 m only at the bottom of each footprint, exceeding 21 m of separation at the top of each footprint as shown in Fig. 9. This discrepancy results from the combined roll and pitch of the cameras. If the cameras were only rolled or pitched, the separation would be consistent with trapezoidal footprints.

Utilizing ray tracing and the parameters as defined, all pixels can be projected onto a flat plane which is representative of the survey given the plane like topology of the ocean. As the altitude of the survey is relatively low, the earth's curvature and atmospheric distortions are not considered. The centre-to-centre distance for all pixels can then be efficiently determined using parallel processing techniques. The averages of the side lengths of each pixel, are also used to define an aspect ratio of GSDh to GSDw for each pixel across the entirety of the image frame to create an anisotropic GSD plot for a visual representation (Xie et al., 2016) as seen in Fig. 9.

For the cameras rolled to 7.7675 the GSD variation across the image ranges from 2.53 cm to 2.89 cm for image height, and 2.22 cm to 2.49 cm for image width. For the cameras rolled to 23.1748 degrees the GSD variation across the image ranges from 2.66 cm to 3.36 cm for image height, and 2.4 cm to 3.14 cm for image width. As such no part of the image has a GSD of 2 cm. This means the area of ground that is represented by each pixel can be as much as 5.65 to 7.1745 cm² or 6.37

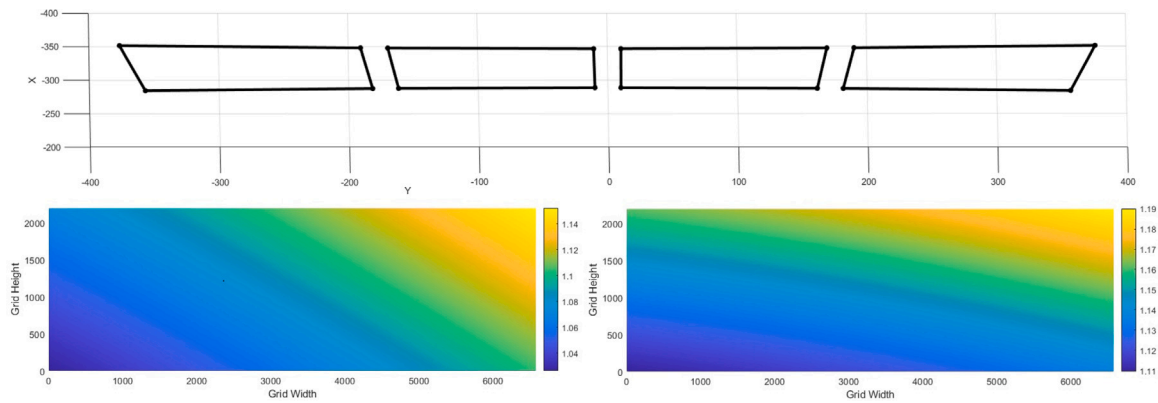


Fig. 9. 4 camera ground footprints with associated anisotropic plots.

to 10.1 cm^2 for the two roll values, instead of the expected 4 cm^2 . This is a 41 to 253% decrease in expected resolution across the images.

These values are also impacted by variations in flight height and aircraft attitude per image, where as much as a 1% difference in altitude or attitude due to the flight dynamics of fixed-wing aircraft will also contribute to further changes in GSD across all cameras. Roll and Pitch of the aircraft is not recorded in the historic datasets; only barometric altitude and aircraft heading are recorded. The largest altitude discrepancy observed in the surveys was 437.5 to 580.5 m. Notably this is not radar altitude creating another source of inaccuracy. As such per image altitude as used for calibrated GSD should be considered on top of the roll and pitch consideration. The effect of this is especially apparent towards the edges of the images as shown in Fig. 9. When looking at video playback this variation in GSD becomes very apparent, affecting detectability of species instances as they move from the top to bottom of the frame as the aircraft flies overhead. This lack of consideration presents a major flaw within Humphries et al.s recent paper (Humphries et al., 2023) on the method used to calculate avian flight heights in HiDef video surveys, which has already received critical review for other aspects of their process (Boersch-Supan et al., 2024).

5.4. Comparisons

Although exact cumulative time figures were not provided by the companies, our estimates – based on the quotes discussed in Section 1 – suggest a potential reduction in total cumulative human screening time from 725 h per survey to less than 4 h using AI-assisted screening. The only direct comparisons available are presented in the four scenarios shown in Fig. 10. Scenario 1 represents the current process using the test case survey, which takes approximately three months from flight to report. The intermediate timeline is estimated since no precise breakdown exists for the manual screening and classification portions. Although WilDetect is designed for the APEM workflow, which differs from HiDef, its role within the process is clearly illustrated. The exact duration of the process remains unknown but can only commence after the manual screening, which consumes the most time and resources. It should be noted that the scale of the X axis for this chart is 50 times larger than all subsequent charts.

Scenario 2 showcases the work completed in this paper, where AI screening occurs after the survey data is collected. Manual classification of the screened data is then performed, which, under generous assumptions, could be completed in a single day by an expert. However, it is important to note that this currently only applies to the non-avian class. If a robust model for all recorded avian species were available, the same timeline could be applied. The total is 72 times faster in this case.

Scenario 3 is the next logical step: employing the system in real-time while the data is being recorded on the aircraft. Given the YoloV5 model's inference time of 39 ms and a video frame rate of 6.51

fps for this data, the model is fully capable of running in real-time with approximately 153 ms between frames at this frame rate. The raw uncompressed images could be passed through the model further enhancing time.

Scenario 4 is our proposed future implementation. In this scenario, data collection must occur at a much higher resolution to enable consistent classification of birds. This implies a narrower swath width (e.g., 63 m vs. 131 m for current systems discussed in Section 5.5), which would require a greater number of transects to maintain coverage—effectively doubling the flight distance. Additionally, UAV platforms are typically slower than manned aircraft (e.g., $\sim 100 \text{ km/h}$ vs. $200\text{--}220 \text{ km/h}$), potentially further increasing flight duration. However, other design options – such as deploying multiple sensors, increasing UAV speed, or using more advanced optics – could offset this. While the estimated timeline assumes a $2\times$ increase in flight time, the actual value will depend on specific mission parameters. Importantly, detection bias due to animal movement remains a known consideration, and future implementations must account for this in transect spacing and timing. Despite these trade-offs, the potential for fully autonomous and high-resolution workflows offers major efficiency gains. Recent developments such as our Modular Detection and Targeting System (MDTS) (Bartlett et al., 2025) demonstrate how adaptive UAV platforms combining wide-area detection with zoom-enabled classification may address many of these challenges in real time—providing a scalable path towards operational deployment of Scenario 4-type systems.

5.5. Recommendations

While this study demonstrates the substantial benefits of integrating automation and artificial intelligence (AI) into the EIA workflow, several fundamental issues must be addressed in future work to ensure the reliability and accuracy of the data collected.

The resolution has been proved to be insufficient both for a ML model but also the traditional manual labelling by mammalian and avian experts. The GSD needs to be decreased to a nominal value of at least 0.5 cm/px . To achieve this there are three controllable parameters. Altitude, sensor dimensions and the focal length of the lens as per Eq. (2). Moving towards cameras with full frame sensors, equipped with 200–300 mm lens or decreasing the altitude of the flight are potential options. Decreasing the altitude is the cheapest measure but increases risk for operators. Increasing the sensor size will allow larger ground coverage areas at the same altitude, minimizing the additional required paths. Increasing the focal length maintains altitude while decreasing GSD. We are currently completing trials utilizing Sony ILX-LR1 cameras equipped with 85 mm lenses on an autonomous fixed wing UAV at flight altitudes of 150 m. This equates to GSD values of 0.6629 cm/px and a nominal swath width of 63 m. This aligns with

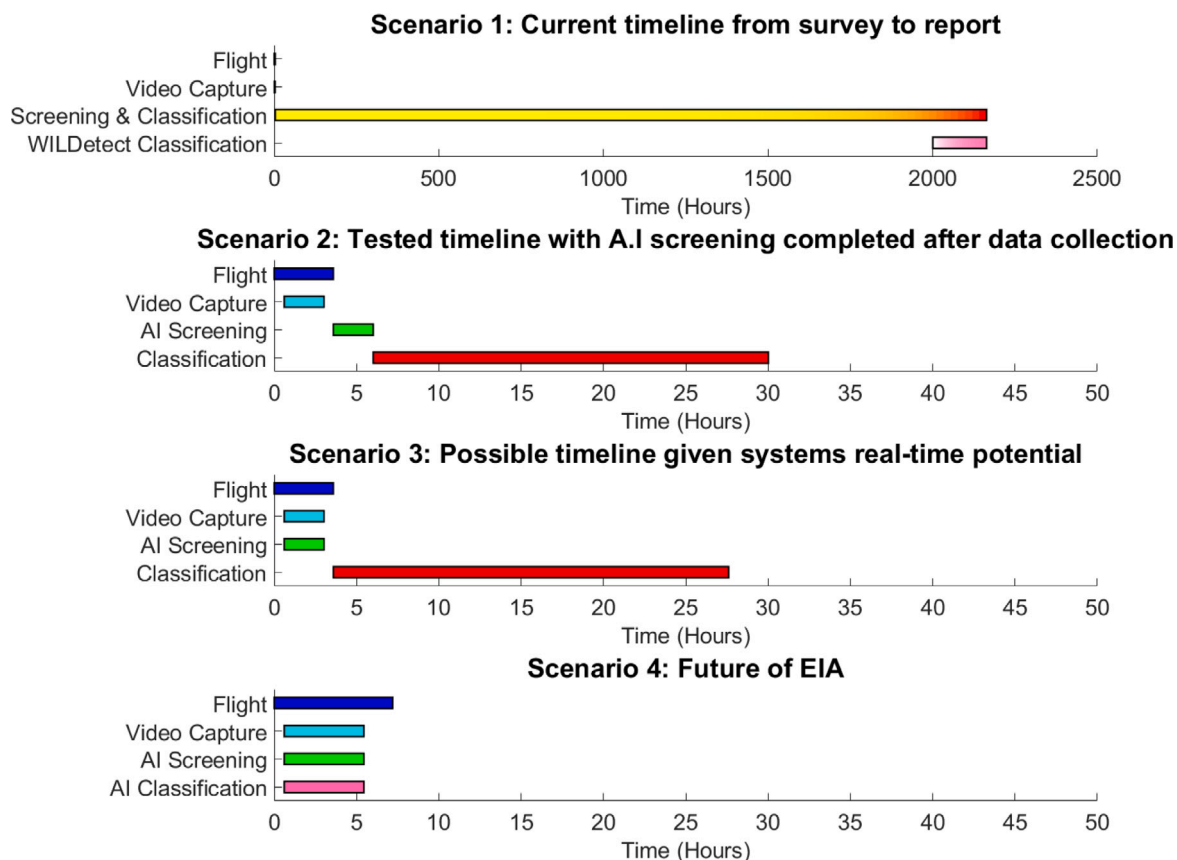


Fig. 10. Timeline comparisons.

scenario 4 in Fig. 10, an approximate doubling of flight time when considering ground coverage, but 9 times increase in ground resolution.

On top of this the assumed GSD needs to be reevaluated under the model discussed here. The recording of aircraft avionic data at 1 Hz is insufficient. As for many traditional high-quality surveys, the utilization of survey grade, high-precision Inertial Measurement Unit (IMU) data should be used when accurate measurements are to be recorded. Along with this, accurate above sea altitude measurements recorded and timestamped to every image frame is required when the data is to be used for measuring aspects such as species sizes or flying heights. Considering these measurements per image as well as considering the installation angles of the camera will lead to much higher accuracy and analysis capability overall.

6. Conclusion

This study demonstrates a clear path towards fully automated screening of marine wildlife in digital aerial surveys. We achieved 97.9% accuracy in mammal classification, eliminated 99.13% of frames from manual review, and recovered 14 mammal detections missed by the original commercial process—exceeding existing screening accuracy by nearly 30%. Our pipeline enabled full dataset parsing and object extraction from raw video, outperforming current practices while operating on low-cost hardware.

The system described here is the first openly documented method for transforming raw HiDef digital video survey data into model-ready SNAG datasets. This reverse-engineered pipeline allows independent users to extract object crops, align detections, and build training sets at scale, without needing access to proprietary vendor tooling. As a result, historical datasets can now be re-analysed and validated using open AI workflows—paving the way for reproducible screening across the industry.

Our findings also reveal critical issues in current commercial workflows, particularly in the classification of avian species. As shown through both visual inspection and model performance, species-level annotations made at the current resolution (2 cm GSD) often exceed what the data supports. This undermines classification reliability, and we urge greater transparency from providers and regulators regarding the confidence and technical validity of species-level detections.

Key contributions of this paper include:

- A scalable and reproducible pipeline for snag dataset generation directly from raw aerial survey footage.
- A validated AI model that outperforms commercial screeners for mammal detection and enables data reduction at scale.
- Clarification of the practical limits of current avian classification methods at standard survey resolution.
- Reassessment of stated Ground Sampling Distances (GSDs) and a call for improved sensor transparency.
- Cost-saving analysis showing potential savings of over €231,000 per pre-construction EIA based on per-flight reductions of €9642.85—without added infrastructure costs.

While this system is currently optimized for detecting and classifying non-avian species, we see no technical barrier to expanding its utility to full multispecies surveys. With improved avian ground truths and access to higher-resolution imaging platforms, we believe future models could support broader classifications—matching or exceeding human consistency across taxa.

Finally, we urge greater scrutiny of the data used to inform environmental policy and offshore planning. Regulatory decisions, licencing, and cumulative impact assessments frequently rely on species abundance and distribution metrics derived from digital aerial surveys.

However, as shown in this study, the underlying classification data – especially for avian species – often lacks the resolution and consistency required for reliable interpretation. Researchers working solely with CSV summaries or aggregated outputs should be aware of these limitations, and whenever possible, revalidate findings using raw imagery. For evidence-based policy to be meaningful, it must be grounded in reproducible, transparent, and technically justifiable data—not just trusted assumptions.

CRedit authorship contribution statement

Ben Bartlett: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation. **Matheus Santos:** Writing – review & editing, Writing – original draft, Validation, Supervision, Methodology, Investigation, Data curation. **Petar Trsljic:** Writing – original draft, Supervision, Project administration, Investigation, Funding acquisition. **Gerard Dooly:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Data curation, Conceptualization.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) used ChatGPT, an AI language model developed by OpenAI, in order to assist with language refinement and editing. After using this tool, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the published article.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was financed and supported by the Marine Institute Ireland under the Postdoctoral Fellowship Programme, Grant Number PDOC/21/03/02. The authors also gratefully acknowledge the Electricity Supply Board Ireland for providing access to data.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.ecoinf.2025.103242>.

Data availability

Data will be made available on request.

References

- Allied Vision Technologies, 2022. Prosilica GT6600 GigE Vision Cameras - Technical Data Sheet. URL https://cdn.alliedvision.com/fileadmin/pdf/en/Prosilica_GT_6600_DataSheet_en.pdf.
- Allied Vision Technologies, 2024a. Discontinuation Prosilica GT with KAI-xxxxx CCD sensors. URL https://cdn.alliedvision.com/fileadmin/content/documents/tqm/PCN/PCN_Discontinuation_Prosilica_GT_with_KAI-xxxxx_CCD_sensors.pdf.
- Allied Vision Technologies, 2024b. GT technical manual. URL https://www.1stvision.com/cameras/AVT/dataman/GT_TechMan_70-0070.pdf. (Accessed 11 June 2024).
- Aumüller, R., Bach, L., Baier, H., Behm, H., Beiersdorf, A., Bellmann, M., Betke, K., et al., 2023. Investigation of the Impacts of Offshore Wind Turbines on the Marine Environment (StUK4).
- Bailey, H., Brookes, K.L., Thompson, P.M., 2014. Assessing environmental impacts of offshore wind farms: Lessons learned and recommendations for the future. *Aquat. Biosyst.* 10, 1–13.
- Bartlett, B., Santos, M., Dorian, T., Moreno, M., Trsljic, P., Dooly, G., 2025. Real-time UAV surveys with the modular detection and targeting system: Balancing wide-area coverage and high-resolution precision in wildlife monitoring. *Remote. Sens.* 17 (5), <http://dx.doi.org/10.3390/rs17050879>, URL <https://www.mdpi.com/2072-4292/17/5/879>.
- Bartlett, B., Trsljic, P., Santos, M., Manduhu, M., Riordan, J., Dooly, G., 2023. Automated close-quarter, high-resolution inspection and 3D reconstruction of unknown infrastructure. In: *OCEANS 2023-Limerick*. IEEE, pp. 1–6.
- Berg, P., Santana Maia, D., Pham, M.-T., Lefèvre, S., 2022. Weakly supervised detection of marine animals in high resolution aerial images. *Remote. Sens.* 14 (2), 339.
- Bioconsult SH, 2024. Digital aerial surveys - hidedf. URL <https://www.bioconsult-sh.de/en/services/digital-aerial-surveys-hidedf>. (Accessed 30 May 2024).
- Boersch-Supan, P.H., Brighton, C.H., Thaxter, C.B., Cook, A.S., 2024. Natural body size variation in seabirds provides a fundamental challenge for flight height determination by single-camera photogrammetry: a comment on Humphries et al. (2023). *Mar. Biol.* 171 (6), 122.
- Borowicz, A., Le, H., Humphries, G., Nehls, G., Höschle, C., Kosarev, V., Lynch, H.J., 2019a. Aerial-trained deep learning networks for surveying cetaceans from satellite imagery. *PLoS One* 14 (10), e0212532.
- Borowicz, A., Le, H., Humphries, G., Nehls, G., Höschle, C., Kosarev, V., Lynch, H.J., 2019b. Aerial-trained deep learning networks for surveying cetaceans from satellite imagery. *PLoS One* 14 (10), e0212532. <http://dx.doi.org/10.1371/journal.pone.0212532>.
- Buckland, S.T., Burt, M.L., Rexstad, E.A., Mellor, M., Williams, A.E., Woodward, R., 2012. Aerial surveys of seabirds: The advent of digital methods. *J. Appl. Ecol.* 49 (4), 960–967. <http://dx.doi.org/10.1111/j.1365-2664.2012.02150.x>.
- Bureau of Ocean Energy Management (BOEM), 2017. Best management practices workshop for atlantic offshore wind facilities and marine protected species. Summary report BOEM 2018-015. Washington, DC: Kearns & West. URL <https://www.boem.gov/Final-Summary-Report-for-BMP-Workshop-BOEM/>.
- BVG Associates, 2024. Guide to an offshore wind farm. URL <https://guidetoanoffshorewindfarm.com/guide>. (Accessed 30 May 2024).
2024. CCD or CMOS - key differences between both sensor technologies. https://www.alliedvision.com/fileadmin/content/documents/landing-pages/Allied_Vision_Whitepaper_CCD_CMOS_Comparison_V1.0_EN.pdf. (Accessed 10 June 2024).
- Chabot, D., Francis, C.M., 2016. Computer-automated bird detection and counts in high-resolution aerial images: A review. *J. Field Ornithol.* 87 (4), 343–359.
- Collier, M.P., Middelveld, R.P., van Bemmelen, R.S.A., Weiß, F., Irwin, C.G., Fijn, R.C., 2022. High-definition bird and marine mammal aerial survey image collection in borsele.
- Conkling, T.J., Loss, S.R., Diffendorfer, J.E., Duerr, A.E., Katzner, T.E., 2021. Limitations, lack of standardization, and recommended best practices in studies of renewable energy effects on birds and bats. *Conserv. Biol.* 35 (1), 64–76. <http://dx.doi.org/10.1111/cobi.13555>.
- Courbis, S., Etter, H., Pacini, A., Portland, M.E., Williams, K., Stepanuk, J., 2023. Technology Gaps for Bird Monitoring in Relation to Offshore Wind Energy Development. Tech. rep..
- Courbis, S., Pacini, A., Etter, H., McManus, M., Portland, M.E., Williams, K., Stepanuk, J., 2024. Assessment of technology gaps for statistically robust data and integration of monitoring of birds and marine mammals into equipment and operations of offshore windfarms (year of publication).
- Cubaynes, H.C., Fretwell, P.T., 2022. Author correction: Whales from space dataset, an annotated satellite image dataset of whales for training machine learning models. *Sci. Data* 9 (1), <http://dx.doi.org/10.1038/s41597-022-01472-6>.
- DALSA, T., 2011. CCD vs CMOS. <https://www.teledynedalsa.com/en/learn/knowledge-center/ccd-vs-cmos/>. (Accessed 10 June 2024).
- Department of the Environment, Climate and Communications, 2020. Guidance documents for offshore renewable energy developers. URL <https://www.gov.ie/en/publication/3d66fb-guidance-documents-for-offshore-renewable-energy-developers/>. (Accessed 09 October 2024).
- Department of the Environment, Climate and Communications, 2024. Draft South Coast Designated Maritime Area plan for offshore renewable energy (SC-DMAP): Additional public consultation period (august 2024). URL <https://www.gov.ie/en/consultation/87025-draft-south-coast-designated-maritime-area-plan-for-offshore-renewable-energy-sc-dmap-additional-public-consultation-period-august-2024/>. (Accessed 09 October 2024).
- Development, W., 2022. Growth of onshore to offshore wind: Atlantic region full report. URL <https://westerndevelopment.ie/wp-content/uploads/2022/10/Growth-of-Onshore-to-Offshore-Wind-Atlantic-Region-Full-Report.pdf>.
- Diederichs, A., Nehls, G., Dähne, M., Adler, S., Koschinski, S., Verfuß, U., 2008. Methodologies for Measuring and Assessing Potential Changes in Marine Mammal Behaviour, Abundance or Distribution Arising from the Construction, Operation and Decommissioning of Offshore Windfarms. Tech. rep., BioConsult SH report to COWRIE Ltd.
- Ditria, E.M., Buelow, C.A., Gonzalez-Rivero, M., Connolly, R.M., 2022. Artificial intelligence and automated monitoring for assisting conservation of marine ecosystems: A perspective. *Front. Mar. Sci.* 9, 918104.

- eTenders Government of Ireland, 2024. CFT download notice. URL <https://www.etenders.gov.ie/epps/cft/downloadNoticeForAdvSearch.do?resourceId=2664075>. (Accessed 29 May 2024).
- European Commission, 2024. Offshore renewable energy. URL https://energy.ec.europa.eu/topics/renewable-energy/offshore-renewable-energy_en. (Accessed 28 August 2024).
- Fiori, L., Doshi, A., Martinez, E., Orams, M.B., Bollard-Breen, B., 2017. The use of unmanned aerial systems in marine mammal research. *Remote. Sens.* 9 (6), 543. <http://dx.doi.org/10.3390/rs9060543>.
- Fox, A.D., Petersen, L.K., 2019. Offshore wind farms and their effects on birds. *Dan. Orn. Foren. Tidsskr* 113 (2019), 86–101.
- Garthe, S., Schwemmer, H., Peschko, V., Markones, N., Müller, S., Schwemmer, P., Mercker, M., 2023. Large-scale effects of offshore wind farms on seabirds of high conservation concern. *Sci. Rep.* 13 (1), 4779.
- Government of Ireland, 2024. Offshore renewable electricity support scheme (ORESS). URL <https://www.gov.ie/en/publication/5099a-offshore-renewable-electricity-support-scheme-oress/>. (Accessed 29 May 2024).
- Hatch, S.K., Connelly, E.E., Divoll, T.J., Stenhouse, I.J., Williams, K.A., 2013. Offshore observations of eastern red bats (*Lasiurus borealis*) in the mid-Atlantic United States using multiple survey methods. *PLoS One* 8 (12), e83803.
- Heinänen, S., Žydelis, R., Kleinschmidt, B., Dorsch, M., Burger, C., Morkūnas, J., Quillfeldt, P., Nehls, G., 2020. Satellite telemetry and digital aerial surveys show strong displacement of red-throated divers (*Gavia stellata*) from offshore wind farms. *Mar. Environ. Res.* 160, 104989.
- Hidef Surveying, 2023. Erebus consent. URL <https://www.hidefsurveying.co.uk/hidef-welcome-erebus-consent/>. (Accessed 28 August 2024).
- HiDef Surveying, 2023. Hidef deliver for ossian. URL <https://www.hidefsurveying.co.uk/hidef-deliver-for-ossian/>. (Accessed 28 August 2024).
- HiDef Surveying, 2024. We are recruiting data reviewers. URL <https://web.archive.org/web/20241212191843/https://www.hidefsurveying.co.uk/we-are-recruiting-data-reviewers/>. (Accessed 30 May 2024).
- HiDef Surveying Ltd, 2024. Aerial survey. URL <https://www.hidefsurveying.co.uk/aerial-survey/>.
- High Resolution, 2009. High resolution video survey of seabirds and mammals in the Rhyll Flats Area. <https://tethys.pnnl.gov/publications/high-resolution-video-survey-seabirds-mammals-rhyll-flats-area>. (Accessed 31 May 2024).
- Hong, S.-J., Han, Y., Kim, S.-Y., Lee, A.-Y., Kim, G., 2019. Application of deep-learning methods to bird detection using unmanned aerial vehicle imagery. *Sensors* 19 (7), 1651.
- Humphries, G.R., Fail, T., Watson, M., Houghton, W., Peters-Grundy, R., Scott, M., Thomson, R., Keogan, K., Webb, A., 2023. Aerial photogrammetry of seabirds from digital aerial video images using relative change in size to estimate flight height. *Mar. Biol.* 170 (2), 18.
- Institute for European Environmental Policy (IEEP), 2022. Delivering synergies between renewable energy and nature conservation. URL https://ieep.eu/wp-content/uploads/2022/12/Final_25_Nov_RES.pdf. (Accessed 28 August 2024).
- Irish Government, 2020. Programme for government: Our shared future. URL <https://www.gov.ie/en/publication/7e05d-programme-for-government-our-shared-future/>.
- Jiang, P., Ergu, D., Liu, F., Cai, Y., Ma, B., 2022. A review of yolo algorithm developments. *Procedia Comput. Sci.* 199, 1066–1073.
- Joher, G., 2020. ultralytics/yolov5: v3.0. <http://dx.doi.org/10.5281/zenodo.4154370>.
- Ke, T.-W., Yu, S.X., Koneff, M.D., Fronczak, D.L., Fara, L.J., Harrison, T.J., Landolt, K.L., Hlavacek, E.J., Lubinski, B.R., White, T.P., 2024. Deep learning workflow to support in-flight processing of digital aerial imagery for wildlife population surveys. *PLoS One* 19 (4), e0288121.
- Kellenberger, B., Tuia, D., Morris, D., 2020. AIDE: Accelerating image-based ecological surveys with interactive machine learning. *Methods Ecol. Evol.* 11 (12), 1716–1727.
- Kemper, G., Weidauer, A., Coppack, T., 2016. Monitoring seabirds and marine mammals by georeferenced aerial photography. *Int. Arch. Photogramm. Remote. Sens. Spat. Inf. Sci.* 41, 689–694.
- Kraus, S.D., Kenney, R.D., Thomas, L., 2019. A Framework for Studying the Effects of Offshore Wind Development on Marine Mammals and Turtles. Tech. rep., Report prepared for the Massachusetts Clean Energy Center, Boston, MA 2110.
- Kuru, K., Clough, S., Ansell, D., McCarthy, J., McGovern, S., 2023a. Intelligent airborne monitoring of irregularly shaped man-made marine objects using statistical machine learning techniques. *Ecol. Inform.* 78, 102285.
- Kuru, K., Clough, S., Ansell, D., McCarthy, J., McGovern, S., 2023b. WILDetect: An intelligent platform to perform airborne wildlife census automatically in the marine ecosystem using an ensemble of learning techniques and computer vision. *Expert Syst. Appl.* 231, 120574.
- Largey, N., Cook, A.S., Thaxter, C.B., McCluskie, A., Stokke, B.G., Wilson, B., Masden, E.A., 2021. Methods to quantify avian airspace use in relation to wind energy development. *Ibis* 163 (3), 747–764. <http://dx.doi.org/10.1111/ibi.12934>.
- Marine Data Exchange, 2024. 2001–2015 Robin Rigg offshore wind farm bird and marine mammal surveys. URL <https://www.marinedataexchange.co.uk/details/TCE-1865/2001-2015-robin-rigg-offshore-wind-farm-bird-and-marine-mammal-surveys>. (Accessed 30 May 2024).
- Mitchell, D., Blanche, J., Harper, S., Lim, T., Gupta, R., Zaki, O., Tang, W., Robu, V., Watson, S., Flynn, D., 2022. A review: Challenges and opportunities for artificial intelligence and robotics in the offshore wind sector. *Energy AI* 100146. <http://dx.doi.org/10.1016/j.egyai.2022.100146>.
- NatureScot, 2023. Offshore wind ornithological impact assessment review: Digital aerial survey methods. URL <https://www.nature.scot/doc/offshore-wind-ornithological-impact-assessment-review-digital-aerial-survey-methods>. (Accessed 28 August 2024).
- New York State Energy Research and Development Authority, 2021. Wildlife data standardization and sharing: Environmental data transparency for new york state offshore wind energy. NYSERDA Report 21-11. Prepared by E. Jenkins and K. Williams, Biodiversity Research Institute, Portland ME.
- Norpix, 2024. Airborne imaging solutions. URL <https://www.norpix.com/airborne/>. (Accessed 11 June 2024).
- NYETWG, 2024. Avian displacement guidance. URL https://www.nyetwg.com/avian-displacement-guidance?cid=4a5465fc-2d1e-4e60-8ea5-daf20d4b6174&utm_campaign=b3bc14c4-af9c-4c1c-8043-9f7142df7d3d&utm_medium=mail&utm_source=so. (Accessed 28 August 2024).
- Pavat, D., Harker, J., 2022. Digital video aerial surveys of marine birds and mammals at solway firth SPA: February 2021.
- Sinclair, R., Booth, 2019. Aerial survey data for monitoring harbour porpoise population health edits following external review RRS cgb 2 title: Porpoise aerial photogrammetry. URL https://assets.publishing.service.gov.uk/media/5e60bf4ad3bf7f108624ce76/SMRU_Ltd_2019_SMRU_2019_Aerial_survey_data_for_monitoring_harbour_porpoise_population_health.pdf.
- Studio, L., 2024. Label studio: Open source data labeling tool. <https://labelstud.io/>. (Accessed 13 June 2024).
- Thaxter, C., Ross-Smith, V., Cook, S., 2016. How High Do Birds Fly? A Review of Current Datasets and an Appraisal of Current Methodologies for Collecting Flight Height Data: Literature Review Report of Work Carried out by the British Trust for Ornithology. Tech. rep., URL <https://www.bto.org/sites/default/files/publications/r666.pdf>.
- Vallejo, G.C., et al., 2017. Responses of two marine top predators to an offshore wind farm. *Ecol. Evol.* 7 (21), 8698–8708. <http://dx.doi.org/10.1002/ece3.3389>.
- Vilela, R., Burger, C., Diederichs, A., Nehls, G., Bachl, F., Szostek, K.L., Freund, A., Braasch, A., Bellebaum, J., Beckers, B., Piper, W., 2020. Divers (*gavia* spp.) in the German North Sea: Changes in abundance and effects of offshore wind farms: A study into diver abundance and distribution based on aerial survey data in the German North Sea. <http://dx.doi.org/10.13140/RG.2.2.31427.76321>.
- Webb, A., Nehls, G., 2019. Surveying seabirds. In: *Wildlife and Wind Farms, Conflicts and Solutions. Offshore: Monitoring and Mitigation*. vol. 4, pp. 60–95.
- Wei, V.F., Büttger, H., Baer, J., Welcker, J., Nehls, G., 2016. Erfassung von seevögeln und meeressäugern mit dem HiDef kamerasytem aus der luft.
- Weiser, E.L., Flint, P.L., Marks, D.K., Shults, B.S., Wilson, H.M., Thompson, S.J., Fischer, J.B., 2023. Optimizing surveys of fall-staging geese using aerial imagery and automated counting. *Wildl. Soc. Bull.* 47 (1), e1407.
- Wilding, T.A., Gill, A.B., Boon, A., Sheehan, E., Dauvin, J.-C., Pezy, J.-P., O'beirn, F., Janas, U., Rostin, L., De Mesel, I., 2017. Turning off the DRIP ('data-rich, information-poor')—rationalising monitoring with a focus on marine renewable energy developments and the benthos. *Renew. Sustain. Energy Rev.* 74, 848–859.
- WindEurope, 2022. Ireland makes history with its first offshore wind auction. URL <https://windeurope.org/newsroom/news/ireland-makes-history-with-its-first-offshore-wind-auction/>. (Accessed 29 May 2024).
- WindEurope, 2024a. European offshore wind farms map (public). URL <https://windeurope.org/intelligence-platform/product/european-offshore-wind-farms-map-public/>. (Accessed 28 August 2024).
- WindEurope, 2024b. Wind energy in europe: 2023 statistics and the outlook for 2024–2030. URL <https://windeurope.org/intelligence-platform/product/wind-energy-in-europe-2023-statistics-and-the-outlook-for-2024-2030/>. (Accessed 28 August 2024).
- Xie, L., Hu, H., Wang, J., Zhu, Q., Chen, M., 2016. An asymmetric re-weighting method for the precision combined bundle adjustment of aerial oblique images. *ISPRS J. Photogramm. Remote Sens.* 117, 92–107.
- Xu, Z., Wang, T., Skidmore, A.K., Lamprey, R., 2024. A review of deep learning techniques for detecting animals in aerial and satellite images. *Int. J. Appl. Earth Obs. Geoinf.* 128, 103732.
- Žydelis, R., Dorsch, M., Heinänen, S., Nehls, G., Weiss, F., 2019. Comparison of digital video surveys with visual aerial surveys for bird monitoring at sea. *J. Ornithol.* 160 (2), 567–580. <http://dx.doi.org/10.1007/s10336-018-1622-4>.