RESEARCH PAPER

TurtleGuard: An Edge-Based Sensor-Fusion and Risk-Scoring System for Hatchling Disorientation and Thermal Reproductive Risk at Olive Ridley Rookeries

Suryakanta Lenka | **Tripathy Divyajyoti Senapati**

DAV Public School, Unit-VIII, Bhubaneswar, Odisha

Mentorship
Guided by **Tanmay Kumar Nayak**
ATL Coordinator

Email: contact@turtleguard.in

Abstract

Seven sea turtle species exist worldwide, but synchronized mass nesting, or arribada, is chiefly associated with olive and Kemp's ridleys. Odisha hosts globally important olive ridley (*Lepidochelys olivacea*) arribadas; the Government of India reported more than **6.98 lakh** nesting turtles at the **Rushikulya river mouth (Odisha)** in February 2025. Olive ridleys contribute to marine food webs by consuming jellyfish, crustaceans and other prey, while a persistent temperature-driven imbalance in hatchling sexes could affect future breeding dynamicsShanker et al. (2024). At beaches where nesting occurs at this scale, artificial lighting can disrupt hatchlings' seaward orientation, while elevated incubation temperatures can alter sex ratios and affect developmental outcomes. These risks operate at different stages of the turtle's life but require timely, site-specific environmental monitoring.

TurtleGuard, an edge-based sensor-fusion and risk-scoring system designed to assess both risks without disturbing turtles or nests. Its optical module combines directional horizon brightness, light-source visibility, spectral measurements, and weather conditions to identify settings associated with hatchling disorientation. Where authorised passive footage is available, movement trajectories provide a behavioural indicator for evaluating the optical score. A separate thermal module analyses sand-temperature proxies and historical records to flag elevated incubation risk. Each score includes its contributing measurements and data-quality status for review by conservation personnel.

TurtleGuard treats the first **0–48 hours** after emergence as an operational priority window, rather than a fixed biological energy limit. Its thermal score does not directly determine hatchling sex; that inference requires local calibration. The present contribution is a transparent monitoring framework and a protocol for validating it. A screened, longitudinal dataset spanning nesting, emergence, and later life stages could enable future tests of unresolved questions about arribada timing, hatchling orientation, thermal resilience, and recruitment.

KEYWORDS

Olive ridley, arribada, hatchling orientation, artificial light at night, temperature-dependent sex determination, sensor fusion, edge computing, risk scoring.

1 INTRODUCTION

1.1 Sea turtle diversity, arribadas, and the Odisha context

Sea turtles inhabit expansive, interconnected ecosystems including coastal beach zones, nearshore waters, offshore currents, and reef systems that facilitate their migration, development, and foraging. Globally, seven extant sea turtle species are recognized: the green turtle *(Chelonia mydas)*, hawksbill turtle *(Eretmochelys imbricata)*, loggerhead turtle *(Caretta caretta)*, leatherback turtle *(Dermochelys coriacea)*, flatback turtle *(Natator depressus)*, olive ridley turtle *(Lepidochelys olivacea)*, and Kemp's ridley turtle *(Lepidochelys kempii)*.NOAA Fisheries (2022), NOAA Fishyeries (2025) While all these species require specific nesting beaches, the two ridley species are uniquely associated with "arribada" nesting. As described by NOAA Fisheries (2022), arribadas are large-scale, synchronized group-nesting events in which hundreds or



FIGURE 1 Adult sea turtle photographed on the beach during the authors' Rushikulya field visit. The photograph documents the coastal setting and an adult turtle observed there. Photograph: authors.

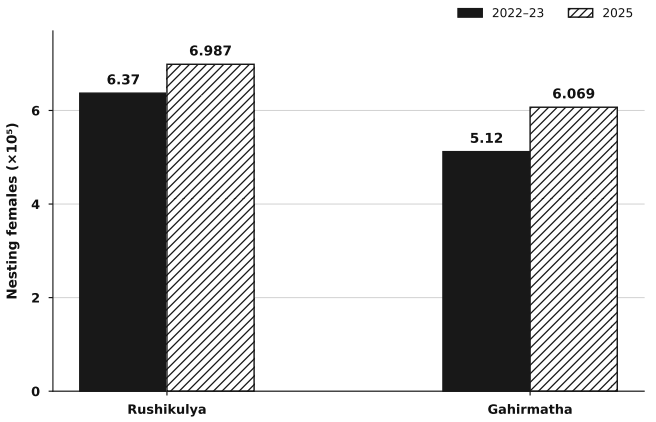


FIGURE 2 Reported numbers of nesting olive ridleys at Rushikulya and Gahirmatha during the 2022–23 and 2025 mass-nesting events. The 2022–23 estimates are approximate; the 2025 report provides exact counts. The bars do not estimate adult population size or identify a cause of year-to-year change.

thousands of individuals may nest within a localized area over a period of hours or days.

The reproductive ecology of the olive ridley distinguishes it from many other solitary nesting species. This synchronization creates both an opportunity and a challenge: while a relatively narrow stretch of coastline can become an exceptionally critical conservation site during an arribada, that same concentration means any local disturbance can impact disproportionately large numbers of adults, nests, hatchlings, and coastal workers in a very short timeframe. In Odisha, the primary current mass-nesting sites are Gahirmatha and RushikulyaShanker et al. (2024). Long-term monitoring reports indicate that while these sites currently host major arribadas, the historically significant Devi rookery has not recorded such events since the late 1990s (Shanker et al., 2024).

The Rushikulya river mouth serves as a particularly salient focal point for this study due to its combination of high nesting density, dynamic beach conditions, proximity to human activity, and an established history of long-term monitoring. The Government of India reported that more than 6.98 lakh olive ridley turtles nested at the Rushikulya river mouth during February 2025Press Information Bureau, Government of India (2025)(Press Information Bureau, 2025). While this figure does not represent a permanent annual population total or serve as evidence that all ecological pressures have been mitigated, it underscores why even localized decisions regarding lighting, road infrastructure, fishing activities, beach access, and environmental monitoring must be managed with precision.

The olive ridley turtle plays different ecological roles depending on its habitat and diet at various stages of development. The feeding behavior of the olive ridleys contributes to the food web of the marine ecosystem, and the nesting process allows the movement of organic matter from the ocean to the sand. These ecological relationships help to understand the need for conservation of a rookery, but these cannot be calculated based on nesting alone. In this current research, the presence of debris, fishing, predators, and habitat conditions is considered separately as the environment.

1.2 Why arribada beaches require a different monitoring approach ?

Conventional wildlife-monitoring systems often record isolated indicators such as nest counts, beach surface temperatures, camera imagery, or meteorological data. While these metrics are useful for specific datasets, they do not fully capture the complex decision environment characteristic of a high-density arribada beach. During nesting seasons, multiple environmental and operational processes occur simultaneously: adults arrive and nest; beach surface conditions fluctuate; sand temperature cycles; cloud cover modifies light reflection; and human personnel require safe pathways. Concurrently, artificial lights may be toggled, hatchlings emerge, and predators may become activeKarnad et al. (2009), Witherington and Bjorndal (1991ba). In this dynamic context, authorized conservation teams must determine which observations necessitate immediate action.

The primary challenge lies not in the sheer number of variables, but in their interdependence. A light source that appears moderate on a clear, moonlit night can become significantly more disruptive when cloud cover reflects skyglow, when the source is directly visible from the beach, or when its position generates a strong landward horizon cue. Similarly, a single measurement of sand-surface temperature at noon does not independently characterize the thermal conditions experienced by embryos during incubation. Factors such as moisture content, depth, shading, beach elevation, rainfall, grain size, metabolic heat within dense nesting clusters, and the timing of developmental stages all interact to alter the actual thermal environment.



FIGURE 3 Sea turtle hatchlings photographed on the sand during the authors' Rushikulya field visit. This illustrates hatchling movement across the beach, it does not establish their heading, destination, or disorientation status. Photograph: authors.

Consequently, TurtleGuard AI is not designed as a standalone "turtle detector." It is engineered as a decision-support framework that addresses four interconnected questions: What is the current environmental condition? Which specific variable is driving the risk? How reliable is the available evidence? What should an authorized person examine or do next? The system is deliberately structured so that an automated score does not translate into an automated ecological intervention.

1.3 Detailed problem statement: artificial-light-induced disorientation

The primary challenge addressed by this study concerns artificial light at night (ALAN). Hatchling sea turtles rely on nuanced visual cues from the terrestrial landscape and horizon to navigate toward the ocean during the critical seafinding phase. Empirical evidence indicates that specific variables including light wavelength, intensity, directionality, and background contrast, significantly influence hatchling phototaxis and subsequent seaward orientation (Witherington and Bjorndal, 1991a; Witherington and Bjorndal, 1991b). Natural coastal environments are rarely devoid of illumination; moonlight, starlight, oceanic reflections, surf luminescence, and structural features such as dune silhouettes, beach gradients, and vegetation architecture all contribute to the prevailing visual environment. The critical factor is often directional contrast: the sea-facing horizon typically provides an expansive, lower, and brighter visual field, whereas dunes and vegetation form a darker landward silhouette.

Anthropogenic lighting can disrupt these vital environmental relationships through several distinct mechanisms. High-intensity road lamps may be directly visible from the beach, while facility lighting may create secondary hazards through reflection off walls, wet sand, cloud bases, vehicles, fishing vessels, signage, or buildings. While cool-white lamps introduce significant short-wavelength energy, even long-wavelength sources can cause significant disturbance if they are excessively bright, poorly shielded, inadequately directed, or visually dominant within the landscape. Consequently, this project does not categorize "amber light" as inherently safeFlorida Fish and Wildlife Conservation Commission (1991), Karnad et al. (2009), nor does it posit that specific wavelength cutoffs, such as 560 nm or 595 nm, provide guaranteed protection. Current sea-turtle lighting guidance emphasizes a multi-faceted design approach: illumination must be minimized, shielded, oriented away from the nesting area, and limited in duration; thus, wavelength is merely one component of a broader, integrated design problem (Florida Fish and Wildlife Conservation Commission, n.d.).

The ecological consequences of hatchling disorientation extend beyond mere deviations in orientation. A hatchling exhibiting landward movement, behavior parallel to the shoreline, navigation toward roads, entry into vegetation, or circling around artificial light sources experiences increased duration of exposure on the beach. This extended period of terrestrial exposure heightens the risk of physiological stress including dehydration and exhaustion, and increases vulnerability to predation, vehicle collisions, physical obstacles, and accidental trampling. Consequently, the project does not define "risk" through the singular metric of directional error; instead, it evaluates a composite of indicators, including deviation from a surveyed seaward heading, cumulative time spent outside the anticipated route, repeated turning, stalled locomotion, and proximity to landward hazards. These indicators are intended strictly for authorized observation and subsequent analytical review, and do not constitute a justification for student intervention, such as approaching, touching, redirecting, or handling hatchlings.

The concept of a "48-hour energy packet," while significant to user requirements, must be articulated with scientific precision. Hatchlings emerge with residual yolk and typically exhibit an initial period of vigorous activity known as the "swimming frenzy."Wyneken and Salmon (1992), Wyneken et al. (2008) While research on loggerheads and other species indicates that early post-hatchling activity is biologically critical, these data do not establish a universal or precisely quantified 48-hour energy reserve specific to the Odisha olive ridley population (Wyneken and Salmon, 1992; Wyneken et al., 2008). TurtleGuard AI therefore employs a 0–48-hour post-emergence timeframe as an operational label rather than a claim that every turtle possesses a precise 48-hour window of usable energy. This label serves a practical function: if seafinding is delayed, the system flags the event as urgent, as the hatchling may lose critical time, remain exposed to hazards, and begin its nearshore transition under sub-optimal conditions.

1.4 Detailed problem statement: thermal reproductive risk

The second primary concern involves thermal reproductive risk associated with temperature-dependent sex determination (TSD)Shanker et al. (2024), Valverde et al. (2010), Jensen et al. (2018). Sea turtles

exhibit TSD, a biological mechanism wherein incubation temperatures influence the probability of embryos developing as male or female offspring. While cooler incubation conditions tend to favor the production of male hatchlings and warmer conditions tend to favor females, fluctuating conditions can result in mixed-sex cohorts. However, this relationship is not governed by a simple, universal rule. Key parameters including the pivotal temperature, the transitional range, the timing of the temperature-sensitive period, and the correlation between heat exposure and embryo survival can vary significantly across different populations, species, nesting sites, nest depths, and annual cycles.

Data from the long-term Rushikulya monitoring program provide critical local context for these complexities. Shanker et al. (2024) report that, across their monitored arribada dataset, the average hatchling sex ratio at Rushikulya was approximately 71% femaleShanker et al. (2024). Their data further indicate substantial interannual variability, with some years exhibiting strong female-biased outcomes while others produced male-biased cohorts. These findings underscore why it is scientifically inappropriate to claim that any single temperature will automatically produce a ”100% female” cohort at Odisha sites. The available local evidence demonstrates significant variance; consequently, a reliable thermal model must be calibrated using population-specific data gathered through authorized scientific methods.

Temperature influences biological outcomes beyond sex determination; extreme or prolonged incubation temperatures may compromise embryo viability, alter development rates, impact hatchling condition, and affect emergence success. While the original TurtleGuard specifications utilized specific parameters including a pivotal temperature of 29.5°C, a slope constant of 1.82, a temperature-sensitive period (TSP) spanning days 18–35, and a critical threshold near 33°C these values are not presented in this paper as established constants for the Odisha regionValverde et al. (2010), Shanker et al. (2024). This is because the project currently lacks a peer-reviewed, site-specific calibration dataset sufficient to confirm their applicability across every nest and season. While studies conducted at other olive ridley rookeries provide valuable context regarding the necessity of thermal monitoring, their specific values cannot be automatically extrapolated to the conditions at Rushikulya or Gahirmatha (Valverde et al., 2010).

Furthermore, the mitigation of thermal risk cannot be achieved through the simple placement of a sensor adjacent to a nest to determine hatchling sex. A temperature probe measures local thermal data but does not directly quantify embryo sex, hatchling survival, humidity, gas exchange, clutch depth, metabolic heat, or microhabitat variation. For ethical reasons, this student project does not involve inserting sensors into active nests or handling eggs. Instead, the proposed system utilizes authorized historical nest-temperature records where available and non-invasive, proximal sand-temperature calibration plots to model environmental context. Any intervention including nest shading, relocation, hatchery decisions, or other management actions remains the sole responsibility of authorized conservation staff and must never be triggered by an automated response from a student-developed device.

1.5 Research gap, aim, and contribution

While existing conservation guidance establishes that inappropriate coastal lighting can disturb hatchlings and that nest temperature significantly influences reproductive outcomes, the primary challenge remains operational integration. Beach managers require a methodology to determine whether an observed risk is driven by direct glare, landward brightness, spectral composition, cloud-amplified skyglow, poor data quality, local thermal conditions, or a combination of these factors. Furthermore, they require outputs that are sufficiently interpretable to facilitate informed discussions with conservation authorities, local communities, facility operators, and safety personnel.

The aim of TurtleGuard AI is to develop a transparent, cost-effective, and edge-capable monitoring architecture that synthesizes multi-source environmental observations into explainable risk assessments. The project pursues five specific objectives: first, to quantify the seaward–landward light relationship rather than relying on single, undirected lux values; second, to record environmental conditions that may alter artificial-light risks; third, to generate non-invasive, permission-dependent behavioral indicators of hatchling orientation; fourth, to monitor thermal context without asserting claims regarding direct sex determination; and fifth, to establish a screened, governed long-term dataset capable of supporting future ecological research.

The central hypothesis is that an explainable combination of directional horizon contrast, landward light exposure, spectral composition, weather context, and behavioral observations will provide a more robust conservation-risk signal than any single measurement alone. A secondary hypothesis posits that seasonally organized thermal and environmental records can identify periods of elevated reproductive risk more transparently than isolated spot-temperature readings. These constitute framework hypotheses; they will be tested exclusively through authorized observation, rigorous data validation, and the explicit reporting of all associated uncertainties.

2 MATERIALS AND METHODS

2.1 Experimental setup

2.1.1 | Study design and evidence boundary

TurtleGuard AI is structured as a multi-stage environmental monitoring and decision-support study; it is not a live-animal manipulation experiment. The first stage utilizes documentary evidence, including historical monitoring reports, environmental records, system simulations, hardware bench testing, and permissioned non-invasive observations. The second stage contingent upon formal approval and supervision would validate whether system outputs correspond with independently recorded environmental or behavioral events. The third stage would evaluate whether authorized lighting modifications produce measurable changes under a properly designed conservation protocol. No

efficacy claims should be made before this staged validation process is completed.

A critical methodological component is the classification of evidence into three distinct categories: verified external evidence, proposed engineering design, and preliminary source-reported prototype values. These categories must not be conflated. Verified evidence is derived from official government releases, established long-term monitoring reports, or peer-reviewed studies. Proposed engineering design describes the parameters TurtleGuard AI is intended to measure or calculate. Preliminary source-reported values originate from existing project specifications; while these may serve as internal test records, they are not treated as independently validated field findings.

2.1.2 | Study area and spatial framework

The primary study focus is the Odisha coast, where Rushikulya and Gahirmatha serve as the principal reference rookeries. It is noted that these sites do not possess identical beach morphology, human infrastructure, weather patterns, nesting density, or thermal regimes; accordingly, each beach is partitioned into distinct sectors rather than treated as a uniform expanse. A sector is defined as a stable segment of coastline characterized by a known seaward bearing, landward background, proximal light sources, beach width, access conditions, and defined observation boundaries.

To mitigate risks associated with potential disturbance or data misuse, each record is associated with a coarse sector identifier rather than a publicly released coordinate. Precise locations including active hatchling routes and sensitive roosting areas are excluded from public datasets to prevent interference. While the public version of the dataset utilizes broader grid cells or sector codes, authorized conservation partners may maintain more granular records under controlled access protocols.

The field model distinguishes four spatial layers. The first layer is the beach surface, where orientation, slope, visible light sources, and obstruction risk are relevant. The second layer is the landward infrastructure zone, comprising roads, buildings, pathways, temporary work lights, signage, fishing activity, and vegetation margins. The third layer is the seaward horizon zone, encompassing the surf, sky, lunar position, cloud cover, and marine-facing light reflections. The fourth layer encompasses the subsurface or near-subsurface thermal context, where sand-temperature proxies are recorded exclusively in authorized non-nesting calibration locations.

2.1.3 | Ethical, legal, and wildlife-safety boundary

All data collection methodologies proposed in this study are strictly non-invasive. Researchers and students are prohibited from physical contact with turtles, eggs, hatchlings, nests, carcasses, or stranded animals. Specific prohibited actions include the excavation of nests, the insertion of probes into active nests, the application of tags or transmitters, the use of flash photography or white-light torches in proximity to nesting or emerging turtles, the operation of drones over restricted wildlife areas, and unauthorized entry into closed beach sectors. In the event that a distressed turtle or hatchling is encountered, individuals must report the situation to the authorized Forest Department, a conservation officer, or the designated field supervisor, rather than attempting an independent rescue.

The installation of cameras, sensors, and other test fixtures is contingent upon prior authorization. Initial deployments are preferred in laboratory bench settings, mock beach environments, or permitted non-nesting buffer zones. Passive imaging is the preferred modality; should near-infrared (NIR) imaging be considered, its specific illumination characteristics must first be reviewed by the supervising authority to ensure it remains undetectable and non-disruptive to the turtles. Consequently, the term ”NIR camera” within the system design refers specifically to a potential low-light imaging option and does not constitute automatic permission to illuminate active nesting beaches.

Lighting interventions are governed by a strict human-in-the-loop protocol. The TurtleGuard AI system may generate alerts recommending that authorized personnel perform specific actions, such as inspecting a light source, lowering a shield, adjusting a beam’s direction, inspecting a temporary work light, or consulting with a facility manager. The system is strictly prohibited from executing autonomous commands, such as deactivating public safety lights, moving nests, deploying shading structures, relocating eggs, or altering coastal operations. This framework ensures that the complexities of conservation and human safety are not reduced to a single, automated control signal.

2.1.4 | Multi-modal sensing architecture

The proposed system integrates five interconnected sensing domains: directional optical sensing, spectral sensing, weather and celestial context, thermal context, and behavioral observation. These streams are synthesized by an edge-computing unit to generate a local score while maintaining a traceable record of the calculation process. Edge computing is utilized to reduce dependence on continuous internet connectivity, minimize latency, and ensure that sensitive imagery remains within a controlled system rather than being automatically uploaded to public cloud services (Shi et al., 2016; Satyanarayanan, 2017).

The directional optical module employs a calibrated low-light camera or paired photometric sensors to estimate the relative brightness of defined seaward and landward sectors. This approach addresses the limitations of single omnidirectional lux readings, which are insufficient for determining whether light originates from the sea, road infrastructure, the sky, a building, or a vehicle. The spectral module utilizes a compact spectrometer or multi-channel light sensor to record relative energy across defined bands. The thermal module incorporates air temperature and permitted sand-temperature proxy sensors at multiple depths within a non-nesting calibration plot. The weather module records environmental parameters including cloud cover, rainfall, humidity, wind, atmospheric pressure, moon phase, and moon elevation where data are available. Finally, the behavioral module is optional and permission-dependent; it analyzes authorized passive imagery to identify movement trajectories without identifying individuals.

TABLE 1 TurtleGuard research questions and evaluation endpoints.

Study question	TurtleGuard measurement	Evaluation endpoint
Does artificial light change the visual route toward the sea?	Seaward and landward sector brightness, direct-source visibility, and horizon contrast	Contrast measured at surveyed beach positions under documented lighting conditions
Which aspects of a light source matter beyond brightness?	Source direction, shielding, operating time, and measured spectral distribution	Contribution of each variable to beach-facing light exposure
Does the optical score anticipate hatchling route disruption?	Environmental score calculated before review of authorised passive footage	Agreement with independently reviewed orientation error, turning, and hazard-sector entry
Does an approved lighting adjustment improve beach conditions?	Paired measurements before and after shielding, re-aiming, dimming, or scheduling	Change in beach-facing light, alongside documented human task-lighting needs
Where does thermal exposure persist during incubation periods?	Repeated sand-temperature proxy measurements by depth and sector	Duration and magnitude of elevated exposure across sectors and seasons
Can proxy temperatures support reproductive interpretation?	Comparison with authorised nest-temperature and biological records, if available	Proxy error and population-specific calibration before any sex-ratio estimate
Does the system remain reliable under field conditions?	Calibration checks, missing-data rate, sensor uptime, and independent review	Performance on nights and sectors excluded from model development



FIGURE 4 Tripod-mounted TurtleGuard field station facing the seaward horizon at the Bateswar Beach olive ridley monitoring hub. The station supports collection of sand temperature, ambient humidity, soil-probe output, and visible-light sensor readings. Photograph: authors.



FIGURE 5 Field-team preparation of the TurtleGuard station at the Bateswar Beach. The photograph shows the tripod-mounted enclosure, sensor wiring, and a depth check for the sand probes. Photograph: authors.

2.1.5 | Hardware design and field robustness

The target edge platform consists of a Raspberry Pi 5 or an equivalent low-power computing unit, integrated with an optional neural-processing accelerator. This hardware selection is based on its relative accessibility for student prototyping, its support for standard camera and sensor interfaces, and its capacity to execute lightweight local models. To ensure that conservation scores remain reliable, the system must record comprehensive diagnostic telemetry including device temperature, power status (battery or mains), storage capacity, sensor heartbeats, time synchronization, and software versions to identify instances where the output may be compromised by hardware malfunction.



FIGURE 6 TurtleGuard field-station setup at New Podampeta. The photograph documents the tripod-mounted enclosure and preparation of sand-probe placement at the shoreline site. Photograph: authors.

Coastal environments present significant challenges for electronic components. Salt spray can lead to corrosion; humidity may cause lens fogging; sand can abrade protective windows; extreme heat can degrade processor performance; rain can impact wiring integrity; insects may penetrate enclosures; and poor grounding can introduce sensor noise. Consequently, all field-ready components must be enclosed, labeled, weather-resistant, and monitored via a maintenance log. Crucially, sensor failures must produce a "data insufficient" status rather than a potentially misleading low-risk score.

The original project specifications included a 180-degree fisheye camera, a 380–780 nm spectral sensor, ambient illuminance sensing, temperature measurements at 0, 10, 30, and 50 cm, and DALI/PWM-compatible lighting interfaces. While these features remain valid design targets, their biological interpretation is refined here: the camera must be calibrated before its pixel values are treated as luminance proxies; the spectral sensor must be validated against a reference source; and the temperature array represents environmental context rather than live-nest conditions. Furthermore, while DALI/PWM compatibility may be utilized in laboratory demonstrations or authorized facility tests, it must never be used as an unsupervised method for controlling beach lighting.

2.1.6 | Dataset organisation and annotation

The primary research output of TurtleGuard AI extends beyond the physical device to encompass a rigorously screened environmental dataset. Each entry within this dataset is structured around a precisely defined observation unit. Depending on the specific data source, these units may manifest as a five-minute environmental window, a lighting event, a sector-night, a temperature time series, a reviewed video segment, or an authorized historical nesting record. To maintain data integrity, records are never merged based solely on shared dates; instead, they must align across four critical dimensions: a known site, sector, timestamp, and provenance level.

2.2 Software implementation and analytical workflow

The software architecture is structured as a sequential pipeline comprising acquisition, quality assessment, feature extraction, risk scoring, explanation, alerting, and storage. The initial layer captures raw data, which the second layer evaluates for completeness, temporal synchronization, and calibration. The third layer extracts interpretable features, including seaward–landward contrast, changes in landward brightness, cloud-adjusted light conditions, thermal anomalies, and trajectory deviations. These features are processed by the fourth layer to generate a provisional risk score. The fifth layer then produces a plain-language explanation for example: "Elevated optical risk: landward sector brighter than seaward sector; source visible; cloud cover high; data quality good." Finally, the sixth layer archives the event and triggers an alert to an authorized reviewer only when specific conditions and data quality metrics satisfy pre-defined rules.

The system is designed to prioritize simple, interpretable models before integrating complex deep learning architectures. In early deployment, rule-based baselines, logistic models, generalized additive models, or decision trees may offer greater scientific utility, as these models allow each feature to be scrutinized individually. While deep-learning models may subsequently assist in tasks such as image classification or source detection, they must not obscure the underlying basis of a conservation recommendation. Consequently, the project employs the term "AI" with precision: AI serves as a supporting analytical layer rather than a replacement for expert ecological judgment.

3 PREPROCESSING AND QUALITY ASSURANCE

3.1 Rationale for preprocessing

Environmental data are inherently susceptible to noise and sensor-induced errors. For instance, a camera may overexpose a high-intensity lamp or underexpose a dark horizon; a spectrometer may drift due to thermal fluctuations; a temperature sensor may be compromised by insufficient burial depth or windblown sand; and a remote weather record may be spatially decoupled from the local site by several kilometers. Furthermore, computer vision models can misidentify a hatchling as a crab, leaf, shoe, shadow, or other moving debris. If raw, unvalidated inputs are integrated into the system without rigorous quality control, even a technically sophisticated dashboard may yield scientifically weak conclusions.

Consequently, preprocessing is not a cosmetic addition but a fundamental requirement to transform raw measurements into traceable observations. Each data point must be accompanied by a timestamp, source identifier, calibration version, and a quality flag. To maintain scientific integrity, a risk score must never be presented with a degree of certainty that exceeds the reliability of its underlying inputs.

TABLE 2 Core TurtleGuard observation record.

Data field	Recorded information	Required check or restriction
Observation identity	Event ID, site sector, date, local time, UTC time, and data source	Unique ID and synchronised timestamps
Optical geometry	Surveyed seaward bearing; camera position; seaward and landward sector boundaries	Record changes in position, obstruction, or camera settings
Directional brightness	Sector brightness and calculated horizon contrast ratio	Label camera-derived brightness as a proxy unless photometrically calibrated
Light source	Source location, direct visibility, shielding, direction, and operating period	Separate fixed fixtures from vehicles and temporary work lights
Spectral measurement	Intensity in sensor-defined wavelength bands	Retain instrument calibration and measurement distance; do not assign a "safe" colour
Weather and sky	Cloud cover, rainfall, lunar conditions, and observation time	Record the source and resolution of weather data
Thermal series	Proxy temperature, depth, sector, shade state, and sampling interval	Identify non-nesting sensor locations; retain gaps and sensor-fault flags
Behavioural record	Reviewed trajectory, seaward bearing, transit time, turns, and final sector	Use only authorised footage; keep reviewer labels separate from predictive inputs
Biological context	Partner-provided aggregate nesting or reproductive records	Preserve provenance, reporting period, and access restrictions
Data quality	Calibration version, missingness, visibility, reviewer status, and device health	Mark the score "insufficient data" when essential inputs fail checks

3.2 Optical calibration and horizon segmentation

Raw image brightness does not inherently equate to physical luminance. Camera output is a product of complex variables including exposure time, gain, aperture, sensor sensitivity, lens vignetting, image compression, white balance, and automatic exposure settings. To mitigate these variables, TurtleGuard AI utilizes fixed exposure where possible, disables automatic image enhancement during measurement, and records calibration images from known reference targets. If a camera has not been specifically calibrated for photometric interpretation, the resulting output is designated as a relative luminance proxy rather than a physical luminance measurement.

The horizon image is partitioned into defined sectors to ensure robust analysis. The seaward sector is a surveyed angular region facing the ocean, excluding direct lamps, boats, or reflective structures where possible. The landward sector is a surveyed angular region facing dunes, roads, buildings, vegetation, and infrastructure. A third exclusion sector is reserved for areas with significant occlusion, temporary equipment, or unreliable visibility. This sector-based methodology provides greater robustness than calculating average brightness over the entire frame, as a single high-intensity lamp or vehicle could otherwise skew the aggregate result.

Furthermore, the system must capture temporal dynamics. While a static image may capture a landward lamp, a time series can distinguish whether that light was continuous, motion-triggered, intermittent, vehicle-related, or activated only during specific work activities. This distinction is critical for management: a permanent, poorly shielded fixture necessitates a structural design response, whereas a temporary vehicle headlamp may only require a behavioral or scheduling intervention.

3.3 Spectral preprocessing

Spectral data are processed as relative bands rather than as simplistic binary safe/unsafe color labels. Candidate bands including short, middle, amber, and red wavelengths must have their boundaries defined within the device specification and maintained consistently throughout the system. The project records both absolute intensity, where calibration data is available, and relative spectral composition. This distinction is critical: for instance, a source with low blue content may still generate a severe landward glare problem if it is sufficiently bright and directly visible. Conversely, a weak source containing shorter wavelengths may not dominate the horizon if it is fully shielded and distant.

The original source specification categorized wavelengths above approximately 560 nm as biologically safe; this paper corrects that assertion. While longer wavelengths may be less disruptive under specific tested conditions, no wavelength renders a poor lighting design inherently harmless. A low-intensity, shielded, downward-directed, and minimally visible long-wavelength source may be preferable to a bright, exposed white source, yet intensity, geometry, shielding, duration, and beach visibility remain essential variables. This nuance is vital to ensure the project does not provide unsafe recommendations, such as the unrestricted use of amber lights.

3.4 Thermal preprocessing

Thermal data are recorded as continuous time series rather than isolated values to preserve temporal context. Each temperature observation is coupled with metadata including depth, sensor identity, local shade conditions, surface wetness, ambient air temperature, and a timestamp. To maintain data integrity, abrupt sensor jumps are flagged as potential hardware faults, and missing intervals are retained as gaps rather than being silently interpolated into a smoothed graph. When environmental data are compared across seasons, measurements are aligned by local time, nesting-stage proxies, and specific weather contexts.

Temperature values derived from non-nesting calibration plots are not treated as direct substitutes for active nest temperatures. Instead, these values serve to characterize broader environmental conditions, such as persistent warming within a beach sector, significant diurnal fluctuations, surface overheating, or cooling following rainfall. Where authorized historical nest-temperature and hatchling-sex data are provided by conservation partners, these datasets may be analyzed separately under a restricted governance protocol. In such cases, the system must explicitly distinguish between proxy thermal contexts and validated nest thermal records.

3.5 Behavioural image processing and human review

Upon approval, the passive video processing pipeline commences by detecting candidate moving objectsRoberts et al. (2017). A lightweight model is employed to classify these candidates into categories such as hatchlings, crabs, humans, moving vegetation, or debris. These candidate tracks are subsequently validated against several criteria: temporal continuity, size range, movement velocity, trajectory morphology, and human review. To prevent data leakage, the model must never be trained and evaluated using randomly mixed frames from the same video; such a methodology allows nearly identical frames to appear in both training and test sets, resulting in artificially inflated accuracy metrics.

Instead, all frames from a single video, a specific night, or a distinct site sector must remain within the same data split. A model evaluated on a novel beach sector, under different weather conditions, or during a separate nesting period provides a more robust measure of performance than one tested on adjacent frames from the same recording. Furthermore, the project should report a comprehensive suite of metrics including false positives, false negatives, precision, recall, F1-score, confusion matrices, and uncertainty rather than relying solely on Sokolova and Lapalme (2009). A system that fails to detect rare hatchling events or consistently misidentifies debris as hatchlings may be deemed unsuitable for deployment, even if its aggregate accuracy remains high.

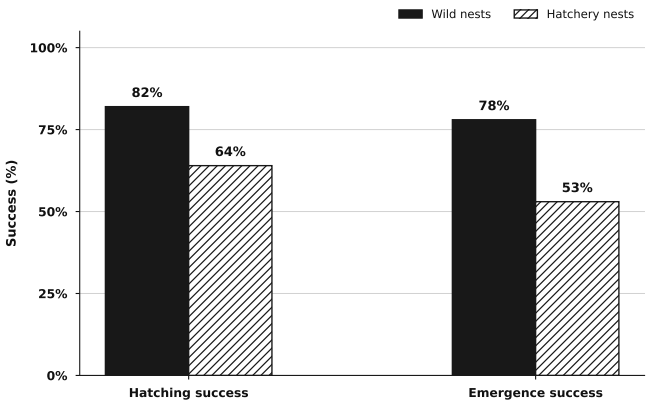


FIGURE 7 Hatching and emergence success in wild nests and nests relocated to hatcheries at Rushikulya. The source does not provide sample sizes or uncertainty intervals beside these summary percentages. The comparison is observational and does not establish a hatchery treatment effect.

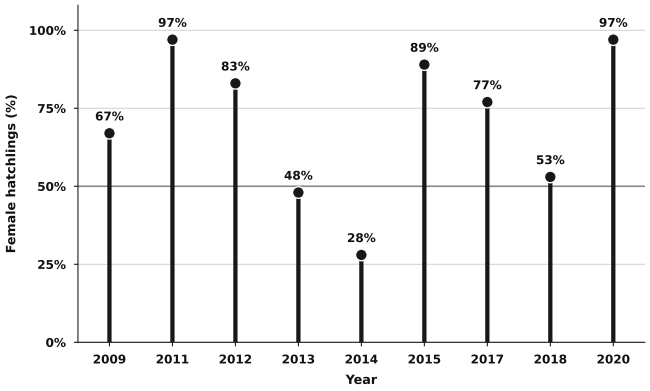


FIGURE 8 Female hatchling percentages in Rushikulya arribada samples for the years displayed. Percentages were transcribed from the published figure; annual sample sizes and uncertainty are not available from these labels. The 50% line is a visual reference and is not a conservation target.

3.6 Data quality score

Prior to the calculation of biological or environmental risk scores, the system computes a distinct data-quality score to evaluate the integrity of the input stream. This metric incorporates several critical variables: sensor calibration status, the presence of missing data, image visibility, time synchronization, weather-data availability, and model confidence levels. If the resulting data-quality score falls below a predefined threshold, the dashboard must indicate that there is insufficient evidence for an automated risk classification. This approach provides a scientifically more robust output than the potential hazard of reporting a misleading low-risk score due to degraded data quality.

A simple form of the data-quality score is:

$$Q_t = \frac{1}{m}\sum_{j=1}^{m} q_{j,t}$$

where $q_{j,t}$ is the quality indicator for sensor or data stream j at time t, scaled between 0 and 1. The score is not a biological measure. It is a transparency measure that prevents low-quality measurements from creating overconfident conclusions.

4 RISK ANALYSIS AND DECISION-SUPPORT FRAMEWORK

4.1 Natural optical horizon gradient model

The principal optical feature in TurtleGuard AI is the Horizon Contrast Ratio, abbreviated HCR. The HCR is an engineering index, not a universal biological lawKarnad et al. (2009), Witherington and Bjorndal (1991a). It compares the relative brightness of the seaward and landward visual sectors:

$$HCR_t = \frac{L_{\mathrm{sea},t} - L_{\mathrm{land},t}}{L_{\mathrm{sea},t} + L_{\mathrm{land},t} + \varepsilon}$$

where $L_{\mathrm{sea},t}$ is the calibrated or relative luminance value for the seaward sector at time t, $L_{\mathrm{land},t}$ is the equivalent value for the landward sector, and ε is a small constant used to prevent division by zero under near-dark conditions.

A positive HCR indicates that the seaward sector is brighter than the landward sector in the measured image or sensor geometry. A negative HCR indicates that the landward sector is brighter. However, this does not mean that all positive values are safe or all negative values are equally dangerous. A single strong direct lamp, a visible illuminated wall, a moving vehicle, or an unusually bright moonlit land surface may affect hatchling decisions differently. HCR is therefore combined with direct-source visibility, spectral distribution, cloud condition, and behavioural indicators.

4.2 Hatchling orientation metrics

Where authorised footage is available, TurtleGuard AI estimates a track heading, θ_i, for each reviewed hatchling trajectory. The expected seaward direction, θ_{sea}, is established by a site survey rather than assumed from the camera frame. The circular orientation error is calculated as:

$$OE_i = \arccos[\cos(\theta_i - \theta_{\mathrm{sea}})]\,.$$

This circular formulation is preferable to simple subtraction because headings wrap around at 0° and 360°; for instance, a heading of 355° and a seaward vector of 5° represent a 10° difference rather than 350°.

The system also records track duration, net seaward displacement, turning frequency, stopped duration, and final sector. A single angular error value may be misleading if a hatchling begins at an extreme angle but corrects its course rapidly. Conversely, a moderate average heading may obscure repetitive looping or a persistent attraction toward a landward source. Consequently, TurtleGuard AI stores the full trajectory summary whenever permitted rather than retaining a single aggregate angle.

The term “disorientation rate” should only be employed once a clearly defined criterion has been established. For example, a study might classify a track as concerning if it terminates landward of the starting point, enters a hazard sector, exceeds a specified delay, or deviates beyond a pre-registered angular threshold. To maintain scientific integrity, these thresholds must be established prior to analysis rather than selected post-hoc to achieve a desired percentage of results.

4.3 Multimodal optical-risk score

The optical-risk score estimates the potential for hatchling disorientation from environmental conditions measured before a behavioural outcome is observedShi et al. (2016), Satyanarayanan (2017). Its inputs are the seaward–landward horizon contrast, visibility and intensity of landward light sources, spectral composition, and weather conditions that may alter light distribution. Hatchling movement is excluded from the predictive inputs so that subsequent trajectory observations can provide an independent test of the score.

For observation window t, the proposed score is

$$R_{\mathrm{opt},t} = 100\,\sigma\Big(\beta_0 + \beta_1 X_{\mathrm{HCR},t} + \beta_2 X_{\mathrm{source},t} + \beta_3 X_{\mathrm{spectral},t} + \beta_4 X_{\mathrm{weather},t}\Big), \tag{1}$$

where $\sigma(z) = 1/(1+e^{-z})$. Each X is a documented feature calculated from measurements available at the time of scoring. Feature direction and scaling must be specified before evaluation; for example, $X_{\mathrm{HCR},t}$ may represent increasing landward dominance rather than the raw horizon contrast ratio. The coefficients $\beta_0, \ldots, \beta_4$ are model parameters, not established biological constants. A rule-based prototype may use provisional weights, but a calibrated predictive model requires labelled training data and evaluation on independent nights or beach sectors.

Authorised passive footage, when available, supplies a separate behavioural endpoint. For example, reviewed trajectories can be summarized by mean orientation error, time to water entry, turning frequency, and entry into predefined hazard sectors. To test predictive performance, the optical score must be fixed before these outcomes are reviewed, and the same trajectory must not appear in both model development and independent evaluation. Agreement between high predicted risk and subsequent behavioural disruption would support the score's usefulness; disagreement would identify conditions under which its measurements or model require revision.

For each observation window, the interface displays the score alongside its contributing features, measurement timestamps, calibration status, and missing-data flags. If essential optical inputs fail quality checks, it reports insufficient data instead of assigning a low-risk value.

4.4 Thermal reproductive-risk framework

The thermal module generates a thermal-context score rather than an automated sex-ratio prediction. The initial iteration of this module evaluates several critical variables: temperature anomalies, temporal persistence, depth gradients, shade conditions, and the reliability of missing data. A generalized cumulative heat-exposure index can be expressed as:

$$E_{\text{heat}} = \sum_{t=1}^{n} \max(0,\, T_t - T_{\text{ref}})\, \Delta t$$

where T_t is the measured proxy temperature at time t, T_{ref} is a documented reference value selected for exploratory analysis, and Δt is the time interval. This equation is useful for comparing relative thermal exposure between sectors or periods. It does not demonstrate embryo mortality or female proportion unless it is calibrated against authorised biological outcome data.

When local calibration data become available, the model may estimate a probability distribution rather than a fixed sex label:

$$\widehat{P}_{\text{female}} = \frac{1}{1 + \exp[-(\alpha + \gamma T_{\text{eff}})]}$$

where T_{eff} is an effective temperature measure appropriate to the validated developmental window. The values of α and γ must come from local authorised records, not copied from another population. The dashboard should report a confidence interval and calibration source. If no local calibration exists, the system must state: "Thermal conditions elevated; direct sex-ratio inference unavailable."

4.5 Integrated TurtleGuard monitoring score

The monitoring interface reports optical and thermal risks separately because the two processes occur on different timescales and require different forms of evidence. An optical alert may describe conditions during a particular night, whereas a thermal assessment summarizes exposure accumulated over days or weeks. If an overall operational priority is needed, it is calculated only after both component scores pass their respective quality checks.

Let $Q_{\text{opt},t}$ and $Q_{\text{therm},t}$ denote the quality scores of the optical and thermal data streams. When both streams meet their predefined quality thresholds, the priority score is

$$\begin{aligned} R_{\text{priority},t} = {} & w_{\text{opt}} R_{\text{opt},t} \\ & + w_{\text{therm}} R_{\text{therm},t}. \end{aligned} \tag{2}$$

The score is displayed only if $Q_{\text{opt},t} \geq q_{\text{opt}}$ and $Q_{\text{therm},t} \geq q_{\text{therm}}$. Otherwise, the system displays "insufficient data" instead of a numerical priority score.

where $w_{\text{opt}} + w_{\text{therm}} = 1$, and the quality thresholds q_{opt} and q_{therm} The individual ratings, quality indicators, timestamps, and associated measurements are all known at the time of assessment. Unavailable or inaccurate measurements will not affect the calculation of the combined rating as opposed to making the numerical score lower through artificial means. The user interface will display the individual ratings and the available measurement.

Algorithm 1 TurtleGuard acquisition and quality-gated monitoring: optical assessment.

Require: Sector data $\mathcal{D}$, surveyed bearings $\theta_{\text{sea},s}$, fixed configuration $\mathcal{C}$
Ensure: Separate scores, optional priority, and audit record
Verify the version of $\mathcal{C}$ and its pre-defined quality rules, feature transforms, coefficients, and thresholds
for each sector s and assessment time t **do**
  Align observations by sector and timestamp
  Retain source IDs, calibration records, and missingness
  Set missing or failed quality indicators to zero
  Compute $Q_{\text{opt},t}$ and $Q_{\text{therm},t}$ from their fixed indicator lists
  $R_{\text{opt},t}, R_{\text{therm},t}, R_{\text{priority},t} \leftarrow \perp$
  Optical window
  $v_o \leftarrow (Q_{\text{opt},t} \geq q_{\text{opt}})$
  $v_o \leftarrow v_o \wedge \text{OpticalInputsPass}(\mathcal{D}, \mathcal{C})$
  if v_o **then**
    Obtain comparable $L_{\text{sea},t} \geq 0$ and $L_{\text{land},t} \geq 0$
    if $L_{\text{sea},t} + L_{\text{land},t} < L_{\text{min}}$ **then**
      $v_o \leftarrow$ false ▷ Below detection floor
    else
      $HCR_t \leftarrow \dfrac{L_{\text{sea},t} - L_{\text{land},t}}{L_{\text{sea},t} + L_{\text{land},t} + \varepsilon}$
      $X_{H,t} \leftarrow (1 - HCR_t)/2$
      $X_{S,t} \leftarrow g_S(\text{source intensity, visibility, glow})$
      $X_{P,t} \leftarrow g_P(\text{measured spectral bands})$
      $X_{W,t} \leftarrow g_W(\text{cloud, rain, lunar context})$
      $\mathbf{x}_{o,t} \leftarrow (1, X_{H,t}, X_{S,t}, X_{P,t}, X_{W,t})$
      $R_{\text{opt},t} \leftarrow 100\, \sigma(\boldsymbol{\beta}^{\top} \mathbf{x}_{o,t})$
    end if
  end if
  if not v_o **then**
    Record optical failure; retain $R_{\text{opt},t} = \perp$
  end if
  Save the optical inputs, output, and quality flags

Algorithm 1 TurtleGuard acquisition and quality-gated monitoring (continued): thermal assessment and validation.

Thermal window
$v_T \leftarrow (Q_{\text{therm},t} \geq q_{\text{therm}})$
$v_T \leftarrow v_T \wedge \text{ThermalInputsPass}(\mathcal{D}, \mathcal{C})$
▷ Check non-nesting location, calibration, depth, coverage, and maximum gap
if v_T **then**
  for each pre-defined sand-proxy depth d **do**
    $E_d, D_d \leftarrow 0$
    for each valid interval k **in** W_{therm} **do**
      $\Delta t_k \leftarrow t_{k+1} - t_k$
      $a_{d,k} \leftarrow \max(0, T_{d,k} - T_{\text{ref},d})$
      $E_d \leftarrow E_d + a_{d,k}\Delta t_k$
      if $a_{d,k} > 0$ **then**
        $D_d \leftarrow D_d + \Delta t_k$
      end if
    end for
    $G_t \leftarrow \text{median}_k\, |T_{d_1,k} - T_{d_2,k}|$ for matched depths and times
    $X_E \leftarrow g_E(\{E_d\})$; $X_D \leftarrow g_D(\{D_d\})$
    $X_G \leftarrow g_G(G_t)$; $X_A \leftarrow g_A(\text{shade record})$
    $\mathbf{x}_{T,t} \leftarrow (1, X_E, X_D, X_G, X_A)$
    $R_{\text{therm},t} \leftarrow 100\,\sigma(\boldsymbol{\gamma}^{\top}\mathbf{x}_{T,t})$
  else
    Record thermal failure; retain $R_{\text{therm},t} = \perp$
  end if
Priority and independent review
  if $v_o \wedge v_T$ **then**
    $R_{\text{priority},t} \leftarrow w_{\text{opt}} R_{\text{opt},t} + w_{\text{therm}} R_{\text{therm},t}$
  else
    $R_{\text{priority},t} \leftarrow \perp$ ▷ Insufficient combined data
  end if
  Store component time windows and configuration version
  Freeze the optical prediction before outcome review
  if authorised, independently reviewed footage exists then
    for each accepted trajectory i **do**
      $OE_i \leftarrow \arccos[\cos(\theta_i - \theta_{\text{sea},s})]$ ▷ Radians
      Record turns, track completeness, and hazard-sector entry as outcomes
    end for
  end if
  Send evidence to an authorised human reviewer
  Perform no automatic wildlife or lighting intervention
end for
Evaluate predictions on held-out nights or sectors; keep all frames from each video in one data split

4.6 Human operational utility and decision constraints

The original TurtleGuard design incorporated a human utility index intended to balance conservation requirements with the illumination needs of ports, roads, villages, or work sites. While this concept is valuable, it necessitates precise definition: human utility must not imply that any degree of wildlife risk is permissible simply because a lamp serves an economic purpose. Instead, it should document specific, legitimate operational constraints, such as pedestrian safety, emergency access, regulated navigation, staff movement, and critical work tasks.

Consequently, a responsible optimization process is governed by two primary constraintsFlorida Fish and Wildlife Conservation Commission (1991). First, it aims to mitigate unnecessary light exposure through technical interventions, including shielding, directional aiming, timed operation, lower mounting heights, reduced spill, task-specific lighting, and minimized spectral impact where appropriate. Second, it must preserve safety requirements established by authorized facility and conservation representatives. The project does not advocate for an uncontrolled lights-off policy; rather, it promotes evidence-based lighting design supported by a formal record justifying the necessity of each specific light.

5 RESULTS AND PERFORMANCE EVALUATION

5.1 Preliminary field-station measurements

The field-station instruments recorded sand temperature, ambient relative humidity, soil-probe output, and visible-light sensor output during 9–10 October 2026. Figures 9–14 show the median of recorded values within each populated three-hour window; the horizontal labels give window start times in Indian Standard Time. The soil-probe output is reported in raw ADC counts, and the visible-light outputs are reported in sensor counts.

5.2 Verified Odisha monitoring context

The preceding figures describe preliminary station measurements.The strongest available ecological context comes from verified Odisha monitoring and official records Press Information Bureau, Government of India (2025).. The Government of India reported that more than 6.98 lakh olive ridleys nested at Rushikulya in February 2025. This figure demonstrates the scale of a recent arribada and justifies the need for careful coastal management, but it does not itself measure hatchling survival, light exposure, sex ratio, or future population trajectory. Long-term monitoring at Rushikulya provides more detailed reproductive context. Shanker et al. (2024) reported that overall hatching success was 82% in wild nests and 64% in hatchery nests, while emergence success was 78% in wild nests and 53% in hatchery nestsShanker et al. (2024). These values are important because they show why conservation reporting should distinguish hatching success from emergence success. A hatchling may develop successfully but fail to emerge from the nest. However, these figures are observational programme results, not a randomised experiment proving that hatcheries always perform worse than wild nests. Beach condition, nest selection, relocation reasons, weather, crowding, and management context may differ between groups.

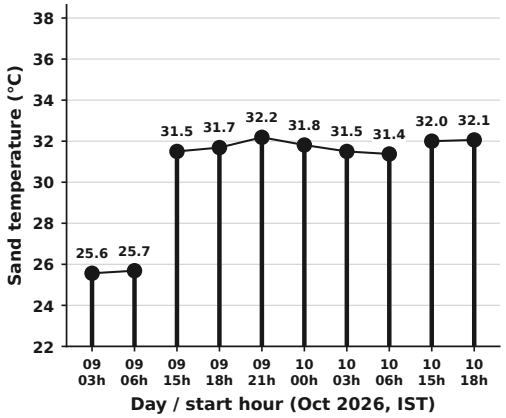


FIGURE 9 Sand temperature at the probe labelled 45 cm (°C); three-hour medians.

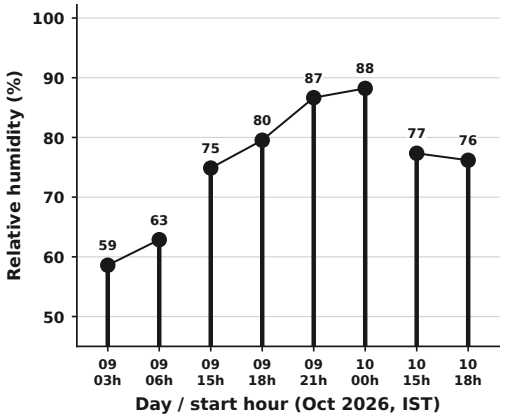


FIGURE 10 Ambient relative humidity (%); three-hour medians.

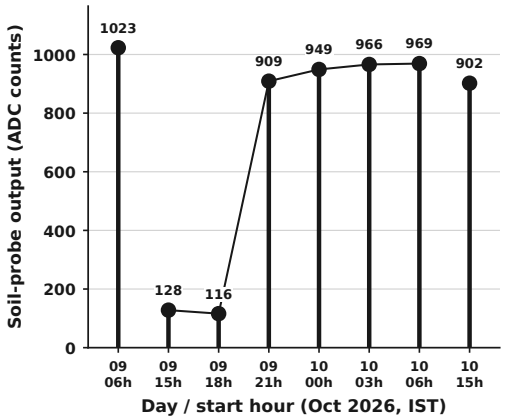


FIGURE 11 Soil-probe output at the labelled 15 cm depth (raw ADC counts); three-hour medians.

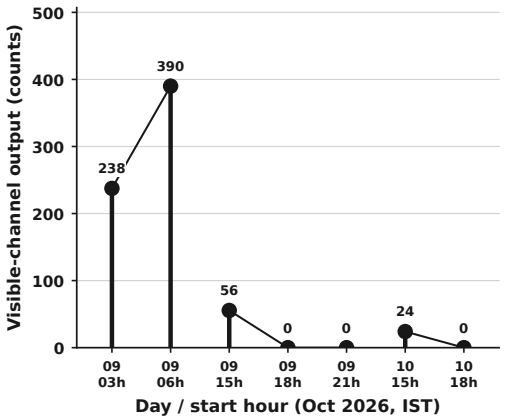


FIGURE 12 TCS3448 visible-channel output (sensor counts); three-hour medians.

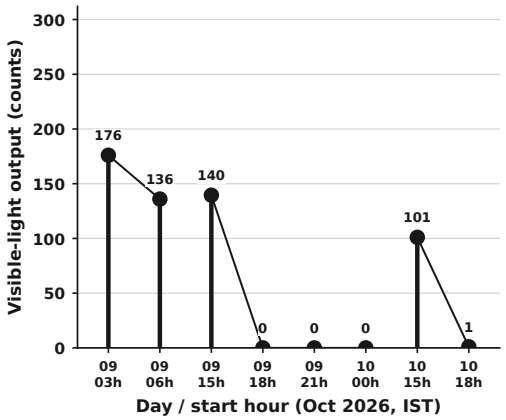


FIGURE 13 Landward-facing TSL2584 visible-channel output (sensor counts); three-hour medians.

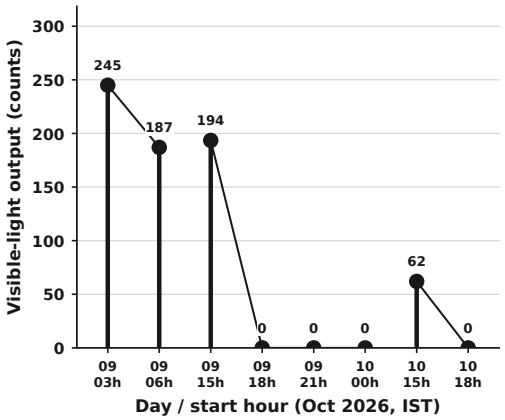


FIGURE 14 Seaward-facing TSL2584 visible-channel output (sensor counts); three-hour medians.

5.3 Source-reported prototype record

The original TurtleGuard project specification incorporated a comparative analysis involving controlled or simulated lighting conditions. Specifically, it recorded data across four distinct scenarios: natural beach environments, unshielded metal-halide lighting, fixed high-pressure sodium lighting, and active TurtleGuard conditions. These figures are retained in the following section as a source-reported prototype record; while they may serve to delineate parameters for future experimental designs, they are not presented as independently confirmed live-field outcomes.

TABLE 3 Optical and thermal alert interpretation.

Displayed status	Condition represented	Information shown to reviewer	Review action
Routine observation	Valid measurements show no predefined elevated indicator	Component scores, contributing measurements, and timestamp	Continue scheduled monitoring
Optical concern	Landward light exposure or source visibility meets a documented alert rule	Affected sector, contrast, visible source, weather, and measurement quality	Inspect the recorded source and conditions
Thermal concern	Sustained proxy heat exposure exceeds a documented site-comparison rule	Temperature series, sensor depth, reference period, and missing intervals	Review the trend against authorised records
Concurrent concerns	Independently valid optical and thermal alerts occur in the same monitored sector	Both component assessments separately, with their different time windows	Prioritise expert review of each concern
Insufficient data	Calibration, visibility, timestamps, or essential sensor data fail quality checks	Failed checks and last valid observation	Repair, recalibrate, or review manually; display no numerical priority score

Alert rules and numerical thresholds must be fixed and validated before field performance is claimed. An optical event and a thermal trend are different biological concerns.

TABLE 4 Published Rushikulya context used in this study.

Indicator	Reported finding	Relevance to TurtleGuard
February 2025 mass nesting	More than 6.98 lakh olive ridley turtles reportedly nested at the Rushikulya river mouth	Establishes the reported scale of this nesting event; it is not a hatchling-survival estimate
Hatchling sex ratio	Approximately 71% female on average in monitored Rushikulya arribada data	Provides population context; substantial variation occurred among years
Hatching success	82% in monitored wild nests and 64% in monitored hatchery nests	Shows why nest setting and hatching outcome must be recorded separately
Emergence success	78% in monitored wild nests and 53% in monitored hatchery nests	Distinguishes emergence success from hatching success
Temperature and sex monitoring	Approximately 15 years of nest-temperature and hatchling-sex-ratio monitoring	Identifies the type of local biological record needed to assess a thermal model

Sources: Government of India, May 2025 release, https://pib.gov.in/PressReleseDetailm.aspx?PRID=2129578; Dakshin Foundation, Monitoring Sea Turtles in India, 2008–2024, https://dakshin.org/wp-content/uploads/2025/08/Dakshin-Turtle-Report-2024_revised.pdf. The wild–hatchery values are observational aggregates and do not establish that nest setting alone caused the difference.

The numerical sequence in Table 5 is biologically plausible as a hypothesis: bright, blue-weighted, unshielded landward lighting would lead to a very high optical conflict, while lighting that is lower and more directed should diminish that optical conflict. However, this table does not provide sample sizes, animal permits, dates, randomisation process, raw trajectories, camera calibration process, arena design, video labels, observer agreement, confidence intervals, or tests. Thus, Table 5 cannot confirm population-level effectiveness.

5.4 Optical Conditions and Hatchling Orientation

The source specification reports substantially different hatchling orientation outcomes across four simulated lighting conditions. Mean orientation error was $8.4^\circ \pm 2.1^\circ$ under natural dark-beach conditions, $138.5^\circ \pm 15.2^\circ$ under unshielded metal-halide lighting, $34.2^\circ \pm 5.8^\circ$ under fixed amber lighting, and $14.1^\circ \pm 3.2^\circ$ under the proposed adaptive-lighting condition. The corresponding landward illuminances were 0.04, 24.80, 8.50, and 2.10 lx, respectively. Reported horizon-contrast ratios were $+0.68$, -0.89, -0.54, and $+0.22$. Together, these figures describe

TABLE 5 Lighting values from the original TurtleGuard specification.

Specification condition	Landward illuminance	Horizon contrast ratio	Reported mean orientation error	Use in this paper
Natural beach / dark-sky reference	0.04 lux	$+0.68$	$8.4^\circ \pm 2.1^\circ$	Proposed reference condition
Unshielded metal-halide simulation	24.80 lux	-0.89	$138.5^\circ \pm 15.2^\circ$	Proposed high-exposure comparison
Fixed high-pressure sodium condition	8.50 lux	-0.54	$34.2^\circ \pm 5.8^\circ$	Proposed fixed-light comparison
TurtleGuard adaptive-lighting condition	2.10 lux	$+0.22$	$14.1^\circ \pm 3.2^\circ$	Proposed system comparison

These are unaudited specification values, not established field results. Sample sizes, raw trajectories, the meaning of "±," and sufficient protocol details are unavailable. The specification also gives a conflicting 114.6° baseline elsewhere; resolve it against raw records before calculating a percentage reduction.

a testable association between the lighting condition, restoration of seaward horizon contrast, and hatchling orientation.

For each recorded trajectory, orientation error can be defined relative to the surveyed seaward bearing as

$$OE_i = \arccos[\cos(\theta_i - \theta_{\text{sea}})], \tag{3}$$

where θ_i is the trajectory heading and θ_{sea} is the seaward bearing, both expressed in radians. The mean error for a condition containing n trajectories is

$$\overline{OE} = \frac{1}{n}\sum_{i=1}^{n} OE_i. \tag{4}$$

A percentage reduction is meaningful only when the baseline and adaptive-lighting measurements come from comparable trials:

$$R_{\text{OE}} = \frac{\overline{OE}_{\text{baseline}} - \overline{OE}_{\text{adaptive}}}{\overline{OE}_{\text{baseline}}} \times 100\%. \tag{5}$$

The specification gives conflicting baseline values: its executive summary uses 114.6°, whereas its results table lists 138.5° for unshielded metal-halide lighting. Substituting these values into the same equation yields approximately 87.7% and 89.8%, respectively. Neither percentage is reported here as a verified treatment effect. Before making an efficacy claim, the underlying trajectory records must establish which baseline was measured, how many trajectories were analysed in each condition, what the reported $\pm$ values represent, and whether lighting conditions were compared under equivalent experimental procedures.

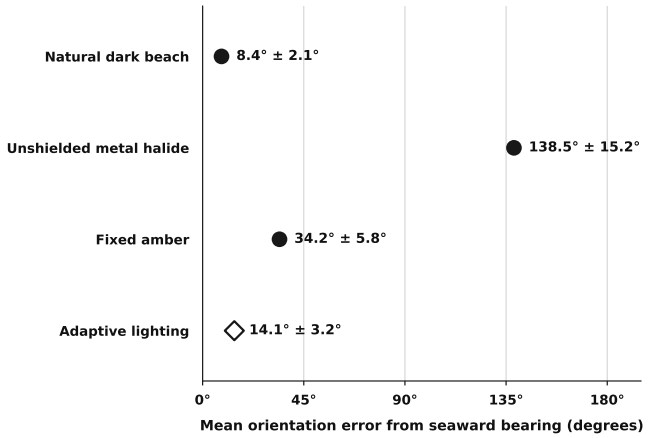


FIGURE 15 Mean hatchling orientation error under four simulated lighting conditions, as reported in the TurtleGuard source text. The source gives additional ± values, reproduced beside the points, but does not define whether they represent standard deviations, standard errors or intervals; they are not plotted as error bars. Sample sizes and raw trajectories are unavailable. No efficacy claim should be inferred.

5.5 Evidence status and validation requirements

The current investigation contains three levels of evidence. Published scientific literature on ecology and official monitoring data provide the conservation context. System architecture and scoring formulas provide the definitions of proposed metrics and analysis procedures. Numerical light comparisons provided along with the initial project specification represent a proto-record that was based on observations not

yet audited. Such separation of the levels is required because a viable method does not guarantee its predictive ability or conservation value in advance.

Validation should follow the chain from the validity of the measurements to their biological meaning. First, optical instruments should be calibrated and measured repeatedly under well-documented lighting conditions. Then, independent evaluators should determine if the system can properly identify visible sources and calculate allowed trajectories of movement. Finally, prospectively collected observations should verify if the scores produced prior to the emergence event can predict independent results of orientation. This chain guarantees that the score is not a self-referential product of the same observations.

5.6 Future evaluation metrics

Future validation of the framework should be evaluated across three distinct dimensions. The first dimension, technical performance, encompasses sensor uptime, calibration drift, data loss, camera reliability, inference latency, and dashboard availability. The second dimension, analytical performance, evaluates the agreement between automated and expert light-source labels, the alignment between automatic and human-reviewed trajectories, score calibration, false-alert rates, and data insufficient behavior. The third dimension, conservation relevance, assesses the association between score conditions and authorized field observations, the efficacy of approved lighting interventions, and the repeatability of results across varying weather conditions and beach sectors.

Reporting should prioritize medians and distributions rather than solely highlighting best-case values. Comprehensive reporting must include sample sizes, dates, the number of nights, sector counts, source-light categories, environmental covariates, and a full accounting of all excluded records. Should an intervention study be approved, the methodology should utilize matched nights or sectors, pre-register outcome definitions, and distinguish between short-term behavioral indicators and long-term ecological outcomes.

6 DISCUSSION

6.1 Why a score-based monitoring system is useful

A score-based system provides utility only when it renders underlying uncertainties more transparent; a solitary traffic-light symbol, devoid of context, is insufficient for nuanced decision-making. TurtleGuard AI addresses this by ensuring every score is accompanied by specific evidence: the measured seaward–landward relationship, visible light-source conditions, spectral profiles, cloud states, thermal trends, available behavioral evidence, and data-quality status. This multi-faceted

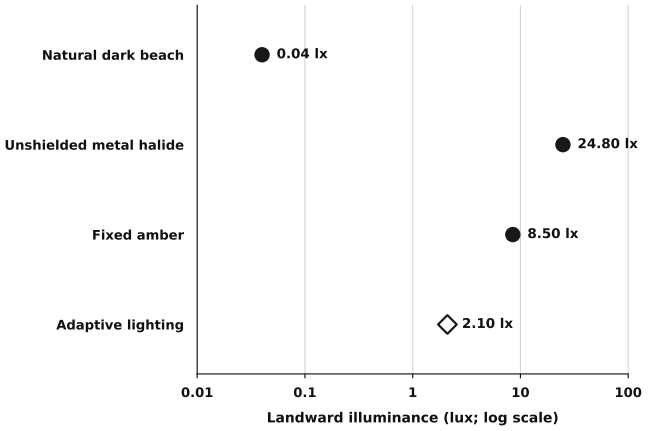


FIGURE 16 Landward illuminance reported for the same four simulated lighting conditions. The horizontal axis is logarithmic to show the full 0.04–24.80 lx range. Calibration and independent field measurement records are unavailable.

output facilitates informed dialogue among diverse stakeholders, including students, mentors, forest staff, conservation organizations, facility managers, and local communities.

The primary value of the score is not to replace expert judgment, but to ensure that critical variables are not overlooked. For instance, a supervisor observing a high score driven primarily by a direct, unshielded landward work lamp can identify and investigate the specific fixture. Conversely, a similar score resulting from poor camera visibility and a missing spectral reading indicates that the necessary response is equipment calibration rather than ecological intervention.

6.2 Verification of lighting modifications

A lighting alteration should be assessed using measurements performed at the beach, and not based simply on the technical details of the fixture itself. For each approved alteration, field observations should record the position, height of mounting, shielding, hours of operation, output level, and visibility from specified sectors on the beach. The directional measurements will be repeated at the exact surveyed points both before and after the modification, while keeping all camera and sensor parameters constantKarnad et al. (2009), Florida Fish and Wildlife Conservation Commission (1991). It is necessary to document weather conditions, phases of the moon, tide, and other sources of light, as they may affect the observed horizon independent of the fixture under assessment.

The immediate engineering consequence is the shift in the amount of light reaching the beach where the alteration took place. Where there are authorized behavioral observations, a second analysis may determine if that corresponds to changes in the hatchling orientation. A comparison of paired locations within the same sector provides better information than one of non-comparable beaches, but still requires consideration of seasonal and nightly differences. The human task-lighting

requirements should be considered with regard to explicitly stated activities for joint review.

6.3 Time-to-water and beach-exposure endpoints

For the emergence event under study, the immediate effects can be measured by the time taken between initial observation and reaching the water, the total distance traveled across the beach surface, and whether the track falls within a pre-designated dangerous sector. The immediate effects reflect different facets of interference as the hatchling might move in the general direction towards the water but take an atypical amount of time or path to do so. If observation was terminated before water entry, it should be deemed incomplete.

For complete tracks, beach-transit time can be defined as

$$\tau_i = t_{\text{water},i} - t_{\text{first},i}, \tag{6}$$

where $t_{\text{first},i}$ is the first reliable detection and $t_{\text{water},i}$ is the observed time of water entry. An exposure summary can, for example, also provide information on the time spent within each beach sector that is being surveyed. Analyses need to consider the distribution of these outcomes along with detection completeness rather than treating transit time as an estimate of residual energy.

6.4 Testing whether sand proxies represent nest conditions

The main issue related to the thermal measurement problem for the thermal module is the discrepancy between temperature in the calibration plot where no nesting occurs and temperature inside the active nest. This discrepancy might depend on the depth, shading, wetness, precipitation, and the phase of the incubation. In case the data on the nest temperature becomes available after authorization, temporally matched measurements of both the proxies and nests could be compared even if the students do not measure the temperatures of nests.

For N authorised, matched observations, proxy error can be summarized as

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}\left|T_{\text{proxy},i} - T_{\text{nest},i}\right|. \tag{7}$$

Error should also be analyzed based on the depth, the weather conditions, the beach segment, and the incubation period. In case there is significant difference, then the proxy may indeed show changes in the environment but it cannot be used in place of nest temperature in the reproductive process model.

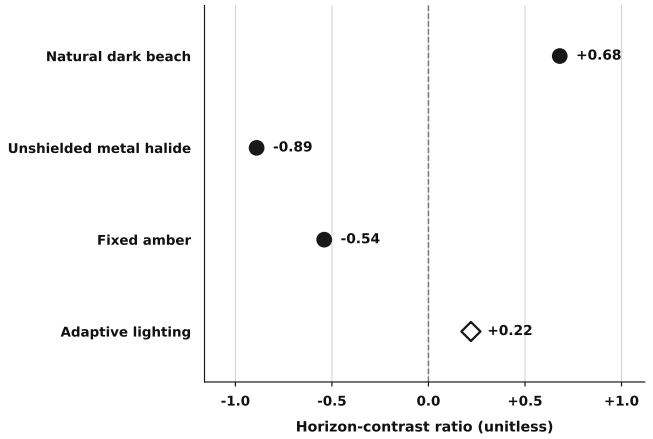


FIGURE 17 Horizon-contrast ratio reported for the four simulated lighting conditions. The zero line identifies the sign change only; it is not a validated ecological safety threshold. Values require calibration and field validation.

6.5 Future work: a screened, organised, life-stage-aligned dataset

The primary future contribution of TurtleGuard AI lies in the construction of a large, screened, and organized dataset that encompasses the turtle's life-history context. This initiative does not presume that students will personally track a single individual from hatching to adulthood; rather, it acknowledges that sea turtles are long-lived, migratory, and late-maturing organisms. Because meaningful population trends often necessitate decades of monitoring, a single student project cannot follow an individual across its entire lifespan. However, a life-stage-aligned dataset can synthesize authorized aggregate records from disparate stages: nesting events, incubation conditions, hatchling emergence, nearshore observations, juvenile records, adult migration studies, fisheries interactions, and return-to-nesting information.

To ensure interoperability, the dataset must adhere to a common data model. Each record should specify its life stage, source, date range, spatial precision, data owner, quality level, and permitted use. For instance, a nesting record can be linked to environmental conditions at a specific beach sector without publicly disclosing exact nest coordinates. Similarly, a nearshore observation can be linked to a date and broad region without releasing sensitive marine-use data. Authorized telemetry data from professional programs may be incorporated only through formal partnerships and at an appropriate level of aggregation.

This longitudinal structure provides a pathway for testing unresolved hypotheses that cannot be addressed by a single season of observation. While it cannot prove every unproven theory, it can transform broad claims into testable scientific questions. For example, researchers can investigate whether arribada onset is more strongly associated with a single environmental trigger or a combination of offshore currents, wind, rainfall, tide, lunar conditions, and prior nesting state. They can evaluate whether landward brightness, spectrum, source visibility, or

cloud-reflected skyglow best predicts orientation risk, or whether thermal exposure patterns interact with beach erosion, moisture, shade, and nesting density. Furthermore, researchers can investigate whether specific environmental signatures predict high hatching success, reduced emergence, or fluctuations in sex-ratio estimates across successive years.

A large dataset must be rigorously screened rather than merely voluminous. Data screening involves removing duplicate records, documenting missingness, retaining failed-sensor periods, recording calibration history, separating direct observations from estimates, reviewing image labels, and preventing data leakage between training and test sets. This process also encompasses governance: sensitive nesting coordinates, active turtle locations, and images featuring local individuals must be restricted. The FAIR principles ensuring data is findable, accessible under appropriate controls, interoperable, and reusable provide a robust framework for this work Wilkinson et al. (2016)

6.5.1 | Proposed longitudinal data pathway

During the initial year, the project must establish foundational infrastructure, including standardized data protocols, a controlled vocabulary, rigorous sensor calibration procedures, a comprehensive sector map, and a repository for historical public and authorized records. The primary output of this phase is a baseline atlas encompassing light sources, beach sectors, weather variables, thermal proxy locations, and identified data gaps; no predictive claims are required at this stage.

In years two and three, the system should facilitate the collection of repeated seasonal environmental observations to be compared against authorized aggregate nesting records. Priority must be placed on methodological consistency: maintaining uniform sector definitions, sensor configurations, time standards, data-quality labels, and documented maintenance logs. A simple, interpretable model should be developed first, followed by a carefully evaluated image-analysis component, provided that sufficient authorized imagery is available.

Over a period of five or more years, the accumulating dataset may support more robust inferences regarding recurring patterns. At this juncture, investigators may evaluate whether specific light configurations consistently correlate with higher concern scores, whether thermal anomalies occur with increasing frequency in particular sectors, whether changes in beach morphology alter environmental exposure, and whether climate-related shifts affect reproductive indicators. These analyses must report associations rather than claim causation unless an appropriate experimental or quasi-experimental design has been employed.

6.5.2 | Future system improvements

Future hardware enhancements may incorporate calibrated high-dynamic-range (HDR) cameras, enhanced spectral resolution, and weather-resistant sensor housings. Additional infrastructure improvements could include solar-backed power systems where applicable, local storage encryption, and a diagnostic dashboard that displays evidence summaries rather than exclusively displaying risk colors. Analytical refinements may include anomaly detection for atypical light events, semantic segmentation to identify visible lamps and structures, site-specific thermal models, uncertainty-aware forecasting, and multilingual alerts tailored for local operational teams.

Biological validation must be conducted by authorized conservation scientists. This process should involve comparing thermal proxy patterns with permitted nest-monitoring records, validating light metrics against supervised behavioral observations, and assessing whether approved lighting modifications successfully reduce measured light spill without compromising defined human-safety requirements. Critically, the project must never utilize AI confidence scores as a substitute for legal approval, ethical review, or professional ecological expertise.

6.6 Limitations

The primary limitation of the current study is its status as a framework and prototype-oriented investigation rather than a completed longitudinal field trial. Consequently, several critical components remain undeveloped: the system lacks an audited, large Odisha-labeled dataset of hatchling trajectories under diverse lighting conditions, a locally calibrated thermal-to-sex-ratio model, and empirical evidence from randomized interventions. Furthermore, the current scope precludes the estimation of adult population size, lifetime survival, or reproductive success from a single season of observations.

A second limitation involves environmental complexityRoberts et al. (2017). Coastal landscapes are dynamic; beach morphology changes, vegetation fluctuates, weather patterns shift rapidly, and nearby infrastructure may be transient, while human activity can unpredictably alter the scene. Because a model trained at a specific site or during a single season may not generalize effectively to other locations, rigorous temporal and spatial holdout testing is essential to address this lack of generalizability.

A third limitation concerns data sensitivity. Increased data volume does not inherently improve outcomes if the data are poorly labeled, legally restricted, biased toward easily observed areas, or released without adequate protection. Accordingly, TurtleGuard AI prioritizes screened, documented, and governed data over the acquisition of uncontrolled data volume.

7 CONCLUSIONS

TurtleGuard AI establishes a robust, evidence-based framework for monitoring two critical risks at olive ridley arribada beaches in Odisha: artificial-light-induced hatchling disorientation and thermal reproductive risk. The project is grounded in the biological reality that olive ridleys are one of the two ridley species associated with arribada nesting, and that Rushikulya can host exceptionally large nesting events. The 2025 official report of more than 6.98 lakh nesting turtles at Rushikulya

underscores the necessity for scalable, systematic monitoringPress Information Bureau, Government of India (2025).

The primary scientific contribution of this work is not an unverified claim that AI has solved turtle conservation, but rather the construction of an explainable system that disentangles directional light, spectrum, weather, behavior, thermal context, data quality, and human operational constraints. The framework actively corrects several common oversimplifications: no single wavelength is universally safe; no universal fixed 48-hour energy reserve can be claimed; non-invasive temperature proxies do not directly reveal hatchling sex; and source-reported prototype values require rigorous audit before they can be presented as indicators of real field efficacy.

The most significant path forward lies in the development of a long-term, screened, governed, and life-stage-aligned dataset. Through structured collaboration, such a dataset could facilitate rigorous testing of unresolved questions regarding arribada timing, beach-scale light exposure, thermal resilience, hatchling orientation, and population recruitment. Consequently, TurtleGuard AI is best understood as a transparent monitoring platform designed to assist authorized conservation teams in transitioning from isolated observations toward evidence-based, accountable decision-making.

ACKNOWLEDGEMENTS

The authors express their gratitude to Er. Tanmay Kumar Nayak, DAV Public School, Unit VIII, Bhubaneswar, local conservation educators, and all authorised agencies and organisations whose long-term sea turtle monitoring efforts facilitate responsible student-led research. The authors acknowledge that all field activities involving sea turtles, nests, hatchlings, restricted beaches, or sensitive location data must be conducted exclusively with formal permission and under professional supervision.

ETHICS AND WILDLIFE-SAFETY STATEMENT

This project is designed as a non-invasive monitoring and data-analysis study. The authors do not handle turtles, eggs, hatchlings, carcasses, or biological samples; furthermore, no active nest is excavated or instrumented by students. The proposed methodology excludes the use of drones, flash photography, white-light surveys, tagging, or any unsupervised activity in restricted areas. All potential field observations, equipment placement, imagery collection, and data access require prior approval from the relevant conservation authorities.

DATA AVAILABILITY STATEMENT

Verified public sources utilized for background analysis are cited in the reference list. Any future TurtleGuard AI dataset should be released in two tiers: a public, coarsened dataset excluding sensitive nesting coordinates and identifiable imagery, and a restricted dataset accessible only to authorized conservation and research partners. Raw prototype data should not be used to support performance claims until comprehensive calibration records, video labels, study protocols, and audit documentation are available.

REFERENCES

Florida Fish and Wildlife Conservation Commission (1991) *Fwc sea turtle lighting guidelines*. Undated guidance; accessed September 25, 2026.
URL https://myfwc.com/media/18511/seaturtle-lightingguidelines.pdf

Jensen, M.P., Allen, C.D., Eguchi, T., Bell, I.P., LaCasella, E.L., Hilton, W.A. et al. (2018) Environmental warming and feminization of one of the largest sea turtle populations in the world. *Current Biology*, 28(1), 154–159.e4. doi:10.1016/j.cub.2017.11.057.

Karnad, D., Isvaran, K., Kar, C.S. & Shanker, K. (2009) Lighting the way: Towards reducing misorientation of olive ridley hatchlings due to artificial lighting at rushikulya, india. *Biological Conservation*, 142(10), 2083–2088. doi:10.1016/j.biocon.2009.04.004.

NOAA Fisheries (2022) *7 sea turtle facts for the ocean lover*. Published June 13, 2022.
URL https://www.fisheries.noaa.gov/feature-story/7-sea-turtle-facts-ocean-lover

NOAA Fishyeries (2025) *Kemp's ridley turtle*. Species account; accessed September 25, 2026.
URL https://www.fisheries.noaa.gov/species/kemps-ridley-turtle

Press Information Bureau, Government of India (2025) *Operation olivia: Icg protects over 6.98 lakh olive ridley turtles along the odisha coastline*. Ministry of Defence press release, May 19, 2025.
URL https://www.pib.gov.in/PressReleasePage.aspx?PRID=2129578

Roberts, D.R., Bahn, V., Ciuti, S., Boyce, M.S., Elith, J., Guillera-Arroita, G. et al. (2017) Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, 40, 913–929. doi:10.1111/ecog.02881.

Satyanarayanan, M. (2017) The emergence of edge computing. *Computer*, 50(1), 30–39. doi:10.1109/MC.2017.9.

Shanker, K., Swaminathan, A., Pusapati, C., Ramesh, H., George, R. & Manoharakrishnan, M. (2024) *Monitoring sea turtles in india, 2008–2024*. Dakshin Foundation, Bangalore, India.
URL https://dakshin.org/wp-content/uploads/2025/08/Dakshin-Turtle-Report-2024_revised.pdf

Shi, W., Cao, J., Zhang, Q., Li, Y. & Xu, L. (2016) Edge computing: Vision and challenges. *IEEE Internet of Things Journal*, 3(5), 637–646. doi:10.1109/JIOT.2016.2579198.

Sokolova, M. & Lapalme, G. (2009) A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, 45(4), 427–437. doi:10.1016/j.ipm.2009.03.002.

Valverde, R.A., Wingard, S., Gómez, F., Tordoir, M.T. & Orrego, C.M. (2010) Field lethal incubation temperature of olive ridley sea turtle Lepidochelys olivacea embryos at a mass nesting rookery. *Endangered Species Research*, 12, 77–86. doi:10.3354/esr00296.

Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J. et al. (2016) The FAIR guiding principles for scientific data management and stewardship. *Scientific Data*, 3, 160018. doi:10.1038/sdata.2016.18.

Witherington, B.E. & Bjorndal, K.A. (1991) Influences of artificial lighting on the seaward orientation of hatchling loggerhead turtles Caretta caretta. *Biological Conservation*, 55(2), 139–149. doi:10.1016/0006-3207(91)90053-C.

Witherington, B.E. & Bjorndal, K.A. (1991) Influences of wavelength and intensity on hatchling sea turtle phototaxis: Implications for sea-finding behavior. *Copeia*, 1991(4), 1060–1069. doi:10.2307/1446101.

Wyneken, J., Madrak, S.V., Salmon, M. & Foote, J. (2008) Migratory activity by hatchling loggerhead sea turtles (Caretta caretta l.): Evidence for divergence between nesting groups. *Marine Biology*, 156, 171–178. doi:10.1007/s00227-008-1074-0.

Wyneken, J. & Salmon, M. (1992) Frenzy and postfrenzy swimming activity in loggerhead, green, and leatherback hatchling sea turtles. *Copeia*, 1992(2), 478–484. doi:10.2307/1446208.