

From Incident Records to Conservation Intelligence: A Methodological Assessment of Dolphin Mortality Accounting in Aotearoa New Zealand

Andrew Gibbs, Dmitry Nikolaenko¹

Pollution and Diseases

DOI: <https://doi.org/10.66659/do3rhq84>

Published: July 28, 2026

How to cite: Gibbs Andrew, Nikolaenko Dmitry. From Incident Records to Conservation Intelligence: A Methodological Assessment of Dolphin Mortality Accounting in Aotearoa New Zealand. Pollution and Diseases. 2026. 31 pages.

DOI: <https://doi.org/10.66659/do3rhq84>

Scope statement

This paper evaluates the public evidence architecture used to record, verify, interpret, and learn from deaths of Hector's and Māui dolphins. It does not rank chemical pollution, fishing, disease, or any other factor as the primary cause of population decline. It does not allege deliberate concealment, criminal conduct, or research misconduct.

The analysis is based on publicly available documents, the public incident spreadsheet supplied with this study, and an author-compiled correspondence dossier. Non-public agency files and internal deliberations were not available for review.

¹ Corresponding author dn@pollution-diseases.org

Abstract

Aotearoa New Zealand has developed a visible and substantial conservation architecture for Hector's and Māui dolphins, including statutory reporting duties, a joint Threat Management Plan, public incident tables, a national incident database, necropsy and genetic programmes, fisheries observers, onboard cameras, abundance surveys, and recurring threat-classification assessments. This paper asks a narrower but fundamental question: does that architecture generate mortality information of sufficient completeness, traceability, causal resolution, and correctability to evaluate conservation performance and revise management when evidence changes? We combine a theoretical assessment of effective conservation monitoring with a documentary analysis of the New Zealand Threat Classification System (NZTCS) series, the 2021 research strategy, official incident and fisheries-monitoring materials, a descriptive audit of the public incident spreadsheet, and a bounded examination of disputed historical records raised in correspondence by one of the authors. The public database audit shows major improvement in source retention over time, but also persistent limits. Among mortality records from 2017–2024, 64.3% of linked event locations were still coded as estimated, 62.7% had a complete necropsy, and 31.0% retained an unresolved cause under the conservative coding used here. Historical records contained much larger provenance gaps, and the public export included exact duplicate event rows and candidate legacy discrepancies that cannot be resolved without original source materials and a visible correction history. Fisheries verification has improved markedly with cameras, yet present reporting metrics cannot be treated as a complete estimate of total mortality. NZTCS reports provide useful, standardised extinction-risk classification, but they are not mortality audits or causal syntheses; their change-centred format can render a chronic crisis as “no change.” We conclude that New Zealand's system cannot reasonably be described as inactive or wholly ineffective. However, public evidence does not demonstrate that all reasonably available analytical measures are being used. A separate national dolphin mortality-intelligence framework is needed to estimate unseen mortality, preserve provenance, reconcile institutional data, compare causal hypotheses, publish corrections, and connect monitoring directly to management evaluation.

Keywords: Hector's dolphin; Māui dolphin; mortality surveillance; strandings; bycatch; necropsy; NZTCS; conservation monitoring; data provenance; adaptive management.

Interpretive safeguards

The public incident database is a record of reported events, not a census of all deaths.

Database completeness, cause-of-death resolution, population trend, and effectiveness of conservation action are separate quantities.

The correspondence dossier is used to formulate auditable questions, not to prove manipulation or institutional intent.

Chemical pollution is treated as one possible contributor within a multi-causal system; the paper's conclusions do not depend on accepting it as the dominant cause.

Key Points

1. Aotearoa New Zealand has developed a substantial conservation architecture for Hector's and Māui dolphins, but its public record of reported deaths remains a detection-dependent minimum count rather than an estimate of total mortality.
2. Retention of source information has improved substantially, yet among mortality records from 2017–2024, 64.3% of linked event locations were estimated, 62.7% had a complete necropsy, and 31.0% had an unresolved cause.
3. Onboard cameras have strengthened fisheries verification, but current coverage and review metrics cannot be treated as complete accounting of fishing-related mortality.
4. Candidate legacy discrepancies expose limitations in public provenance, versioning, and correction procedures; they do not establish intentional manipulation.
5. The New Zealand Threat Classification System provides consistent extinction-risk classification, but it is not a mortality audit, causal synthesis, or evaluation of management performance.
6. A separate national dolphin mortality-intelligence framework is needed to estimate unseen mortality, preserve provenance, reconcile institutional data, publish corrections, compare causal hypotheses, and connect monitoring directly to management evaluation.

1. Introduction

Hector's dolphin (*Cephalorhynchus hectori hectori*) and Māui dolphin (*C. h. maui*) are endemic coastal cetaceans whose conservation has received sustained scientific, political, and public attention in New Zealand. The institutional response is not superficial. It includes marine mammal sanctuaries, fishing restrictions, statutory capture reporting, periodic abundance estimation, public incident reporting, pathology, genetic sampling, formal risk assessments (1-3), a Threat Management Plan (TMP), and a dedicated research strategy (4, 5). DOC's current public summary reports that the 2020–2021 survey estimated 54 Māui dolphins over one year of age within the survey area, with a 95% confidence interval of 48–66 (6). Hector's dolphins are distributed among several South Island subpopulations and remain classified as threatened (7, 8).

This article forms part of a continuing research programme on the evidentiary and institutional dimensions of the Hector's and Māui dolphin decline (9-13). The broader programme ultimately includes a detailed examination of chemical contamination and a formal test of the chemical hypothesis. The present paper deliberately precedes that

causal assessment: it asks whether the public mortality-accounting and classification architecture is capable of evaluating any plausible explanation fairly, including explanations that were not prominent in earlier management frameworks. The persistence of very small or declining populations despite sustained institutional activity makes the absence of a transparent, recurring causal synthesis an object of analysis in its own right. The case is therefore relevant not only to cetacean conservation but also to the study of how scientific and administrative systems recognise problems, preserve uncertainty, and revise inherited models.

The existence of many programmes, however, does not by itself establish that the mortality evidence system is effective. Conservation policy may be active, visible, and well intentioned while the underlying analytical chain remains incomplete. An incident must first be detected, then reported, recovered, correctly identified, examined, entered into a database, reconciled with fisheries and museum records, and translated into a demographic or causal model. Failure or uncertainty at any link changes the meaning of the final count. A list of recorded deaths is therefore not equivalent to total mortality, and a threat category is not equivalent to an explanation of population dynamics.

This distinction matters because small populations are highly sensitive to a few unobserved deaths and to modest changes in reproduction or juvenile survival. The challenge is not confined to any one causal hypothesis. Fishing mortality, toxoplasmosis and other diseases, pollution, vessel interactions, habitat change, prey limitation, acoustic disturbance, climate effects, and interactions among these factors all depend on the same evidentiary infrastructure. Weak detection, poor provenance, incomplete pathology, or an inaccessible correction pathway can bias every explanation, not only the explanation favoured by a particular researcher or institution.

The central paradox examined here is therefore institutional rather than rhetorical. New Zealand publicly presents a serious commitment to biodiversity and to the recovery of these dolphins. At the same time, its principal recurring status reports are designed primarily to reproduce standardised classifications, not to audit mortality detection or compare causal models. The public incident database candidly acknowledges uneven historical quality and incomplete representativeness. External challenges to particular legacy records may be difficult to adjudicate because source provenance and revision histories are not fully visible. The result is a gap between policy visibility and analytical accountability.

The paper addresses six questions. First, does the public system distinguish reported incidents from total mortality? Second, can historical and modern records be traced to original evidence and corrections? Third, what proportion of recovered deaths reaches a complete necropsy and a resolved cause? Fourth, how far are fishing-related deaths independently verified? Fifth, do recurring status reports test the completeness of the mortality evidence and remain open to new causal mechanisms? Sixth, is there a visible learning loop connecting identified research gaps, external expert challenges, management action, and subsequent evaluation?

Our argument is deliberately limited. We do not claim that government agencies have done nothing, nor that individual scientists acted improperly. We do not use a disputed-record dossier as proof of intentional alteration. We also do not claim that chemical contamination is the principal cause of decline. The question is whether the public analytical architecture is proportionate to the conservation stakes and whether it is designed to discover its own errors and blind spots. On the evidence reviewed, the answer is mixed: significant capacity exists and some components have improved materially, but the system remains incomplete as a national mortality-intelligence and adaptive-learning framework.

2. Materials and Methods

2.1 Documentary corpus

The analysis draws on six overlapping bodies of evidence. The first is the NZTCS sequence for marine mammals: the 2001 and 2005 national lists, the 2009 audit, and the 2013, 2019, and 2024 assessments (14-19). The second comprises the Hector's and Māui Dolphin TMP, the 2021 research strategy, public research pages, and the October 2022 current-project spreadsheet supplied with this study (4, 5, 20, 21). The third comprises the DOC incident-database page, annual incident tables, the external user guide, and the public spreadsheet export supplied with this analysis (22-24). The fourth includes statutory reporting provisions and official fisheries-verification materials concerning observers, imagery, and onboard cameras (25-28). The fifth comprises relevant peer-reviewed work on carcass recovery, conservation monitoring, pathology, bycatch, and demographic management. The sixth is a 14-page correspondence dossier compiled by A.G. and exchanged with D.N. concerning alleged inconsistencies in historical stranding, sighting, and netting records (29).

The dossier is evidentially asymmetric: it identifies specific records and possible source documents, but it also contains allegations and causal inferences that cannot be tested from the correspondence alone. We therefore extracted only record-level propositions capable of independent audit. Claims about criminal liability, deliberate deception, personal conflicts of interest, or institutional motive were excluded. The dossier is used as a stress test of the public correction architecture: when a citizen researcher identifies a plausible duplicate, relocation, or category error, can the public system show the original record, the chain of revisions, the adjudication, and the final reasoned outcome?

2.2 Evaluation framework

Monitoring effectiveness was assessed against principles from conservation monitoring and adaptive management. Monitoring should not be treated as stand-alone surveillance; it should answer explicit questions, account for imperfect detection, and trigger management learning (30-34). We define an effective mortality-accounting system as one that can distinguish observed events from total deaths, preserve data lineage, quantify uncertainty, resolve causes where possible, compare alternative mechanisms, correct errors transparently, and connect findings to decisions.

Dimension	Operational question	Minimum public evidence
Completeness	What fraction of true deaths is detected, reported, recovered, and entered?	Detection/recovery studies, reporting effort, camera/observer coverage, and explicit correction for unseen mortality.
Accuracy	Are species, date, location, number, and event type correct?	Validation rules, source comparison, imagery or genetic confirmation, and anomaly checks.
Timeliness	How rapidly are incidents verified, examined, and published?	Publication lag, pending-case status, and dates for report, response, necropsy, and final determination.
Provenance	Can every value be traced to an original document or observation?	Persistent source identifiers, archive location, acquisition date, and revision history.
Causal resolution	How often can cause and contributing factors be established?	Necropsy denominator, exclusion reasons, multi-factor coding, and confidence.
Independent verification	How much self-reported mortality is independently checked?	Observers, cameras, imagery, audits, and reconciliation statistics.
Correctability	Can external or internal challenges be adjudicated visibly?	Submission protocol, response deadline, independent review, correction log, and versioning.
Integration	Are agency, pathology, genetics, museum, fisheries, and research records linked?	Common event/individual identifiers and routine cross-system reconciliation.
Management utility	Does information test measures and change decisions?	Predefined indicators, thresholds, counterfactual or before-after analysis, and decision records.
Adaptive learning	Does the system search for new mechanisms and revise its model?	Competing hypotheses, research gap tracking, annual review, and documented uptake of results.

Table 1. Evaluation dimensions used in this study.

2.3 Descriptive audit of the public incident spreadsheet

The supplied public workbook contained an Events sheet with 681 data rows and 32 populated fields, representing 678 unique event identifiers, and an Individuals sheet with 735 data rows and 32 populated fields. We conducted a descriptive audit of record structure, source fields, location method, necropsy status, and cause-of-death coding. Records were retained as mortality records when the individual row represented an explicit death, historic mortality, or an otherwise ambiguous event accompanied by death or necropsy evidence. Explicit live release, refloat, and non-fatal sick or wounded sighting categories were excluded. One mixed entanglement event

described a live-released adult and a dead calf under the same event-level observation type; only the individual with death and necropsy evidence was retained. Because the most recent years contained pending or recently published records, period comparisons were limited to 1921–1989, 1990–2008, 2009–2016, and 2017–2024.

A cause was coded as unresolved when both the suspected primary cause field and the necropsy-result field indicated unknown, none, indeterminable, not applicable, body not recovered, sent for necropsy without a final result, or an equivalent absence of a resolved determination. “Complete necropsy” was taken from the database’s necropsy-status field. “Original report unknown/null” refers to linked event records for which the holder of the original report was not identified. These are metadata indicators, not mortality rates, and they should not be interpreted as proof that any particular death was missed or misclassified. Multiple individuals can occur in one event, and some event records have no complete individual record.

2.4 Analytical limits

This is a public-document and data-architecture analysis, not an audit with compulsory access to agency systems. We did not inspect internal DOC or Fisheries New Zealand databases, original museum accession cards, unpublished field notes, full correspondence files, camera footage, compliance investigations, or panel meeting records. Publicly obscured locations and commercially sensitive fields were not reconstructed. The study therefore evaluates what the public evidence can support and what remains unauditable, rather than making claims about information that may exist internally.

The database user guide requests that DOC be notified before findings derived from the data are submitted for publication and that resulting papers be supplied to the Department (22). Before journal submission, the authors should provide DOC with the coding protocol, candidate-record table, and draft manuscript, invite factual corrections, and retain a transparent record of any resulting amendments. Such consultation should improve accuracy without giving any institution editorial control over the analysis.

3. From event recording to mortality intelligence

The key analytical error in many mortality systems is to treat the visible record as if it were the underlying process. Let D_t be the true number of deaths in period t , and let p_d , p_r , p_c , and p_e denote the probabilities of detection, reporting, carcass recovery, and database entry. A simplified observation model is:

$$\text{Recorded deaths}_t = \text{True deaths}_t \times p_{\text{detection},t} \times p_{\text{reporting},t} \times p_{\text{recovery},t} \times p_{\text{entry},t} + \text{error}_t$$

Each probability varies with coast, season, weather, carcass buoyancy and decomposition, human visitation, reporting awareness, fishing method, observer or camera coverage, and institutional practice. Recorded incidents can therefore rise when true mortality falls if detection improves, and they can fall when true mortality rises if

carcasses are not recovered or reports are not made. A trend in incident counts is uninterpretable unless changes in the observation process are measured or modelled.

The mortality-intelligence chain

True death -> detection -> report -> recovery -> species/individual identification -> necropsy and ancillary testing -> database entry -> cross-system reconciliation -> demographic and causal analysis -> management response -> evaluation.

A conventional incident database begins near the middle of this chain. A conservation intelligence system must make the losses, uncertainties, and decisions at every stage visible.

Two denominators are especially important. The first is the total number of deaths, including carcasses never observed. The second is the number of recovered carcasses eligible for examination. Cause-of-death percentages among necropsied animals cannot be applied directly to total mortality if carcass recovery and necropsy eligibility are selective. Decomposition, location, season, body condition, recovery logistics, and whether a death occurs in a fishery can all affect entry into the examined sample. Effective analysis therefore requires an explicit flow diagram showing how many cases are lost at each stage and why.

The distinction between monitoring and intelligence is not semantic. Monitoring records repeated observations. Intelligence integrates observations, uncertainty, validation, causal interpretation, and decision consequences. For an endangered dolphin, an effective system should be able to say not only how many bodies were recorded, but what fraction of mortality those bodies may represent, how confident the agency is in each record, which causes are directly established or only inferred, what evidence would discriminate among competing explanations, and whether management reduced the relevant vital-rate loss.

4. Institutional architecture and genuine strengths

A balanced assessment must begin with what the New Zealand system does well. The Marine Mammals Protection Act requires accidental or incidental killing or injury to be reported, including location, species or description, and circumstances. DOC publishes annual incident tables, associated pathology reports, and a downloadable national spreadsheet. The website describes a verification delay of up to four weeks, quarterly spreadsheet updates, genetic testing, measurements, photography, and necropsy. It also explicitly warns that causes may be multi-factorial and that decomposition, freezing, loss to the tide, or failure to recover a body can prevent diagnosis (23, 25).

The TMP provides a joint governance framework led by DOC and Fisheries New Zealand. The 2021 research strategy is unusually candid about uncertainty: it lists fishing, toxoplasmosis and other disease, seismic surveying, seabed mining, vessels, oil spills, other pollution and sediment runoff, coastal development, and climate change,

and recognises overlapping direct and indirect mechanisms. Its objectives include better cause-of-death information, disease research, population trajectories, reproductive parameters, and fisheries impacts. It expressly prioritises better bycatch estimates, cryptic mortality, non-commercial bycatch, and seasonal disparities in carcass collection (5). This is not a policy architecture that denies complexity in principle.

There is also evidence that targeted measures can improve demographic parameters. At Banks Peninsula, long-term photo-identification data indicated improved adult survival after sanctuary restrictions, although projected population growth remained close to replacement and required sustained monitoring (35). Genetic mark-recapture has provided repeated Māui abundance estimates using comparable methods (6, 36, 37). Necropsy and tissue programmes have identified disease processes, including toxoplasmosis, and recent surveillance has expanded testing for pathogens (38). These programmes demonstrate real scientific capacity.

Independent fisheries verification is also changing. Historical inshore observer coverage was low, but onboard cameras were rolled out across in-scope fleets and completed in May 2025. In the first quarter of 2026, 220 vessels had cameras, 25% of recorded fishing events were reviewed, and 96% of protected-species capture events detected by reviewers had also been reported by fishers (27, 28). These figures are a substantial governance development. They do not prove complete reporting across all fishing activity, but they provide a stronger verification mechanism than self-reporting alone.

Component	Principal function	Demonstrated strength	Residual limitation
Marine Mammals Protection Act reporting	Creates a legal duty to report accidental death or injury.	Clear statutory obligation and minimum event information.	A duty to report does not itself measure compliance or unseen events.
DOC incident database and annual tables	Publishes reported incidents, metadata, and pathology reports.	Accessible, regularly updated, and increasingly detailed.	Detection-dependent; legacy provenance and revision history are incomplete.
Necropsy and tissue programmes	Determines cause and collects pathology/genetic evidence.	High-value direct evidence and public reports for examined animals.	Selective carcass recovery; decomposition and missing bodies reduce coverage.
Fisheries observers and cameras	Independently verifies capture reporting and fishing activity.	Rapidly improving coverage and measurable reconciliation metrics.	Not all footage is reviewed; historical coverage was uneven; post-release mortality may remain unseen.

Component	Principal function	Demonstrated strength	Residual limitation
Abundance and demographic studies	Estimates population size, survival, connectivity, and trend.	Methodologically strong studies for selected populations.	Intervals between estimates and spatial gaps limit rapid diagnosis.
Threat Management Plan and Research Strategy	Coordinates threats, research, and management.	Recognises multiple threats, cryptic mortality, and knowledge gaps.	Public implementation tracking is not current or comprehensive.
NZTCS assessments	Standardises national threat status.	Consistent categories and periodic reassessment.	Not designed as a mortality audit, causal synthesis, or evaluation of management performance.

Table 2. Main components of the present public architecture.

5. Findings

5.1 Reported incidents are not an estimate of total mortality

DOC accurately describes the incident database as being based on reported events. It also states that the system relies on prompt reports from the public and tends to contain more reports from places with higher visitor numbers (23). The external user guide goes further: national-office data come from different sources and field offices, there is no guarantee that the database is fully representative or accurate, processes have changed over time and have not been followed evenly, and ownership, storage, and formatting have changed (22). These caveats are scientifically important and should be treated as central analytical facts rather than marginal disclaimers.

International studies illustrate why. Modelling and tagging work in France recovered only 8% of released bycaught dolphin carcasses on shore, while a study of California coastal bottlenose dolphins estimated a carcass recovery fraction of 0.25 (95% CI 0.20-0.33) (39, 40). These values cannot be transferred directly to New Zealand because currents, coastline, species behaviour, carcass properties, and search effort differ. Their relevance is methodological: even coastal stranding networks may observe a minority of deaths. Without a New Zealand-specific recovery and detection model, recorded beachcast deaths remain a minimum count rather than a mortality estimate.

The observation process also changes through time. Public awareness campaigns, hotline access, smartphone photographs, GPS, SeaSpotter, improved field response, legal reporting, observers, and cameras all increase detection or verification. A historical time series of incident counts therefore mixes ecological change with surveillance change. Annual reports should publish observation-effort covariates alongside incident counts: camera coverage and review rate, observer effort, number and

distribution of public reports, response time, carcass-search effort, weather-related access, and the proportion of reports that reached a recovered body.

For Māui dolphins, very low abundance makes this problem acute. One unobserved adult female death may have a population consequence far larger than its numerical value suggests. Conversely, a single recovered carcass can dominate a small necropsy sample. The appropriate response is not to abandon incident data, but to embed them in an explicit observation model and to report uncertainty in the total number of deaths. The 2021 research strategy ranks cryptic mortality as a priority, yet the public materials reviewed do not present a national, regularly updated cryptic-mortality estimate for Hector's and Māui dolphins.

5.2 Public database quality has improved, but the time series remains heterogeneous

The spreadsheet audit shows a strong and important improvement in source retention. Among mortality records dated 1921–1989, 72.8% were linked to events for which the original report holder was unknown or blank. This fell to 37.1% in 1990–2008, 10.2% in 2009–2016, and 0.0% in 2017–2024. The result is evidence of institutional improvement, not a reason to dismiss the modern database. It also shows why the full historical series cannot be treated as a homogeneous dataset without period-specific quality modelling.

Period	Mortality records (n)	Estimated linked location (%)	Original report holder unknown/null (%)	Complete necropsy (%)	Unresolved cause (%)
1921–1989	158	98.7	72.8	38.0	50.0
1990–2008	310	70.6	37.1	65.8	43.9
2009–2016	98	58.2	10.2	70.4	45.9
2017–2024	126	64.3	0.0	62.7	31.0

Table 3. Descriptive metadata and cause-resolution indicators in the supplied public incident spreadsheet.

Notes: Percentages are descriptive properties of individual mortality records and their linked event metadata, not estimates of true mortality or agency error. Event rows were de-duplicated by event identifier before linking; the three duplicated identifiers contained exact duplicate rows. “Estimated location” includes the database codes “Estimated location” and “Estimated” and may reflect a legitimate geocoding of a narrative place description. “Unresolved cause” is a conservative composite of unknown, none, indeterminable, not applicable, body not recovered, and pending/no-final-result categories. The most recent 2025–2026 records were excluded from period comparisons because publication and necropsy results may still be pending.

Location coding requires careful interpretation. The persistence of estimated locations - 64.3% in 2017–2024 - does not imply that most modern records are geographically unreliable. A beach name or fishing area can be correctly converted to a point while still being labelled estimated. The problem is analytical precision: an estimated point should not be used as if it were the exact position of death, especially in spatial attribution of fishing, outfalls, currents, or habitat. The database should preserve the original narrative, the geocoding method, an uncertainty radius or polygon, and the identity and date of the geocoder.

The public Events export contained three duplicated event identifiers, each represented by a pair of exact duplicate rows with all 32 populated fields identical. This may be a spreadsheet-generation artefact rather than duplication in the underlying operational database or in any scientific analysis. Nonetheless, a public export intended for research should pass automated uniqueness checks before release. More generally, public users need a version number, schema version, release checksum, and machine-readable change log identifying additions, deletions, merged records, split records, and corrected fields.

DOC's annual tables sometimes explain individual publication corrections. For example, the 2024-2025 page notes one incident missed from the online table but present in the quarterly spreadsheet, and another not publicly reported until a compliance investigation allowed release. This is a useful form of transparency. The next step is to make such explanations systematic through a permanent, searchable correction register rather than leaving them as isolated page notes.

5.3 Cause-of-death information is valuable but affected by strong selection

Necropsy provides the strongest direct evidence in the current system, but its denominator is the set of bodies recovered in examinable condition, not all deaths. DOC recognises that decomposition, freezing, and loss to the tide may prevent a conclusion or prevent examination entirely. The public spreadsheet indicates that complete necropsy coverage among the mortality records used here ranged from 38.0% in the earliest period to 62.7% in 2017–2024, with a peak of 70.4% in 2009–2016. Even in the modern period, therefore, the observed cause distribution excludes a substantial share of mortality records and all deaths that were never detected or recovered.

The toxoplasmosis study by Roe et al. (41) illustrates this attrition. Fifty-four Hector's or Māui dolphins were received over 2007-2011, but 26 were excluded from the relevant analysis, including 23 because of decomposition and three because tissues were incomplete; seven of the 28 examined animals were determined to have died from toxoplasmosis. The study is important evidence that disease can cause death. It is not, by itself, an unbiased estimate of the fraction of all population deaths caused by toxoplasmosis. The same caution applies to any cause identified from a non-random set of carcasses.

A single "primary cause" field also risks compressing multi-factor mortality. DOC's webpage correctly states that several factors may contribute to a death. A dolphin may be chronically diseased or exposed to contaminants, experience nutritional stress, and

then become entangled; alternatively, a carcass may display trauma but have insufficient evidence to determine whether trauma preceded or followed death. Effective causal analysis needs separate fields for immediate cause, underlying condition, contributing factors, evidential basis, confidence, and unresolved alternatives. Population models should then test how different combinations influence survival and reproduction.

The modern unresolved-cause proportion in our coding was 31.0%. This is an improvement over earlier periods, but it remains large enough to affect causal inference, especially when deaths are few. Annual reporting should therefore include a standard case-flow diagram: total reported mortality events; bodies recovered; bodies suitable for necropsy; complete necropsies; cases with immediate cause; cases with contributing factors; indeterminable cases; and cases awaiting results. Reasons for exclusion should be tabulated by season, region, population, and incident source.

5.4 Fisheries reporting is becoming more verifiable, but the historical and statistical gaps remain

Fishing-related mortality has a clearer statutory reporting pathway than many other causes. Section 16 of the Marine Mammals Protection Act requires accidental killing or injury to be logged and reported. Yet self-reporting is not equivalent to independent verification. Inshore observer coverage varied from 0.5% to 6.4% of effort between 2012-2013 and 2021-2022 and declined during the transition to camera monitoring; the official national figures are fleet-wide and should not be read as coverage of the specific Hector's and Māui risk fisheries (28).

The imagery record also shows why verification quality matters. In a 2019–2020 identification project, 114 marine mammal bycatch events were reported; 67 had photographs, three had video, and 44 - 39% - had no imagery available for taxonomic review. Where reasonable-quality photographs existed, observer identification was confirmed (26). The result is encouraging for identification accuracy but also demonstrates that a material share of events could not be independently assessed from imagery.

Onboard cameras are a major improvement. In January-March 2026, Fisheries New Zealand reported cameras on 220 vessels, review of 25% of recorded fishing events, and fisher reporting for 96% of reviewer-detected protected-species capture events. These figures should be interpreted precisely. They indicate strong concordance within the subset of events reviewed and in which reviewers detected a protected-species capture. They do not imply that 96% of all captures in the fleet were detected or reported, because 75% of recorded events were not reviewed and not all fishing activity is necessarily within the same camera scope. The correct policy conclusion is that cameras create a powerful verification platform whose sampling design and review algorithms should be made explicit.

A further problem is cryptic or delayed mortality. Animals may be released alive but subsequently die, or a capture may occur without the body being retained. Work on New Zealand sea lions shows how unobserved or latent bycatch can alter population interpretation even when mitigation devices are used (42). That study is not evidence

that the same magnitude of latent mortality applies to dolphins. It is evidence that post-capture fate and unobserved interactions must be modelled rather than assumed negligible. The Hector's and Māui research strategy already identifies this need; implementation should include tagging or physiological studies where ethical and feasible, gear-specific post-release assessment, and camera review designed around protected-species risk rather than only a generic fraction of fishing events.

5.5 Disputed historical records reveal a correction problem more than they prove a mortality conclusion

The correspondence dossier identifies several candidate discrepancies in historical records. The strongest are propositions that can be matched to entries in the public spreadsheet. Two Māui dolphin records, public event IDs 944 and 946, are dated 28 August 1974 and 24 August 1975, both at Oakura Bay, and have identical coordinates coded as estimated. The 1974 record is coded as female with haplotype G in the public export; the 1.53 m length appears in the DOC 2001 compilation and associated notes identified in the dossier, not in the current spreadsheet. The 1975 record lacks sex and DNA information. The dossier argues that the second entry may duplicate the first. The public export alone cannot establish whether these are two separate animals, a duplicated event, or a later reconstruction from distinct source notes (29).

A second case concerns public event ID 1066, dated 8 April 1989 with haplotype G and located in the export at “Opunaki Beach, New Plymouth,” with estimated coordinates, no detailed description, and an unknown original report holder. The dossier cites contemporary media and museum provenance material said to place the event at Bell Block, approximately 60 km away. Again, the public data cannot adjudicate the conflict. Resolution requires the original museum card or specimen file, contemporaneous media, field notes, the first database entry, and a revision history showing when and why the location changed (29).

The latest dossier material raises a further, specifically testable mtDNA provenance question. It reports that event ID 924 (Waikanae, 1967) is coded as haplotype N in the public export and event ID 1048 (Oakura, 1988) has no haplotype recorded, whereas the 2002 paper cited in the dossier describes the Waikanae specimen (NMNZ 1615) as haplotype J and the Oakura specimen (TM A57.705) as the rare N haplotype. Because the specimen-event crosswalk, original laboratory records, and database revision history were not available, this is a candidate metadata or linkage discrepancy rather than an established error (29, 43).

Other dossier claims concern an apparent additional 1977 netting record and the treatment of a 1995 dolphin neonate reportedly found in a great white shark stomach. An Official Information Act request seeking the underlying 1977 notes and clarification of the disputed haplotype fields was pending at the time of writing; without the response and primary files, it is treated only as a verification step. These matters could not be independently resolved from the supplied public export. They should not be repeated as established errors. Their methodological significance is that the public system does not expose enough lineage to let an external reader determine whether a contested record was copied, split, reclassified, geocoded, or corrected. A high-stakes

conservation database needs a process that converts such challenges into auditable outcomes (29).

Candidate issue	Public-record indication	What is not established	Evidence needed for resolution
Oakura 1974/1975	IDs 944 and 946 share location and estimated coordinates; dates are one year apart; one record is more complete.	The spreadsheet alone does not prove duplication or double counting.	Original stranding notes, specimen/accession numbers, photographs, tissue records, first database versions, and merge/split history.
Bell Block/Opunake 1989	ID 1066 lists Opunaki/Opunake with estimated coordinates and unknown original report holder.	The dossier's Bell Block location has not been independently verified in this study.	Contemporary media, museum provenance card, field notes, specimen identity, geocoding decision, and revision log.
Possible additional 1977 netting	Dossier alleges that a further record appears in later compilations.	No conclusion can be drawn from the public export without a complete source comparison.	Side-by-side tables from the original and later publications, source cards, event identity, and documentary chain.
1995 shark-stomach neonate	Dossier alleges later representation as a dolphin stranding.	Classification and downstream use were not independently verified.	Original predator-capture record, specimen record, publication coding, and exact treatment in later datasets.

Table 4. Candidate legacy discrepancies requiring source-level audit.

Nothing in Table 4 establishes intentional manipulation. Historical databases often accumulate duplicates, shifted coordinates, placeholder dates, and category changes through ordinary transcription and migration. The relevant performance test is how an institution responds after a plausible discrepancy is raised. A robust protocol would acknowledge the challenge, preserve all versions, appoint a reviewer not responsible for the original record, compare primary sources, publish a reasoned determination, and issue a traceable correction if required. We found no public, standardised mechanism of this kind for the incident database or linked museum records. Where a corrected record informed a map, risk assessment, or earlier report, the correction register should identify the affected outputs and, where material, publish an erratum or corrected map.

5.6 NZTCS is useful classification, but weak mortality analysis

The NZTCS series is often the most authoritative public summary of conservation status, but its function is narrower than its authority may suggest. The 2013 report describes itself as a full list, statistical summary, and brief notes on the most important changes. It explicitly repeated the 2009 assessment with identical categories, criteria,

and process, and many of the same panel members (17); both assessments operated under the 2008 NZTCS manual (44). The 2024 assessment likewise used the previous assessment as the starting point and evaluated new information primarily through population size and trend (19), using the revised 2022 manual (45). This is a rational design for standardised classification. It is not an independent reconstruction of mortality evidence or a systematic search for all causal mechanisms.

The series also mixes ecological change with knowledge change. In 2013, only one status change reflected a real monitored population change; four reflected improved knowledge. In 2024, thirteen taxa changed status, but only one was labelled better and one worse; eleven changes were neutral, and six of the thirteen resulted from increased knowledge. The reported improvement for common bottlenose dolphins was explicitly not a genuine improvement in the population but a consequence of new information about offshore abundance (17, 19). A category time series is therefore partly a history of nature and partly a history of data, taxonomy, rules, and expert interpretation.

The term “no change” is especially problematic outside specialist use. In the 2024 table, Māui dolphin is listed as NCu5a with “No change,” and Hector’s dolphin as NVu5d with “No change.” In the same report’s coding system, the trend code 5 denotes a forecast decline of 50-70% over three generations; size code a means fewer than 250 mature individuals and code d means 5,000-20,000 mature individuals (19). There is no formal contradiction: the category did not change. But the wording can communicate administrative stability where the underlying code still contains severe decline. Chronic crisis becomes less narratively visible than reclassification.

The reports devote explanatory space mainly to changed categories, qualifiers, taxonomy, or significant new information. This means a species that remains critically endangered over repeated cycles may receive less analysis than a taxon moved because of new survey data. In 2013, the Hector’s dolphin note said there was “very little new information,” referred to 18 necropsies in 12 months, and inferred continued decline outside Banks Peninsula. The 2024 report reduces the two focal dolphins largely to coded table entries. Neither document is designed to audit the completeness of mortality records, estimate carcass recovery, evaluate reporting compliance, or compare causes across examined and unobserved deaths.

The 2022 revision added the qualifier Conservation Research Needed, defined as a need to understand the cause of decline and/or solution for recovery, and split data-poor qualifiers into recognition, size, and trend. These are constructive changes. However, a qualifier is not a research programme. For Māui dolphin, CR and Data Poor Trend appear in the table, but the report does not specify a time-bound set of causal questions, responsible institutions, funding, completion indicators, or how results will alter the next assessment. A classification system can identify a knowledge gap without closing it.

Assessment cycle	Public form	Analytical emphasis	Implication for this study
2001 and 2005	Marine mammals embedded in national lists of many taxonomic groups.	Assignment of standard threat categories and qualifiers.	Establishes long-term classification lineage, but little species-specific causal analysis.
2009	Dedicated complete audit published as a journal article.	Reassessment under the revised NZTCS manual.	Baseline for the modern, more comparable series.
2013	Dedicated NZTCS report with brief background notes.	Repeat of 2009 categories, criteria, and process; focus on changes.	Strong comparability, but dependence on inherited framing.
2019	Dedicated NZTCS report.	Updated status list and selected explanations.	Continues incremental reassessment.
2024 (published 2025)	Dedicated report, coded criteria, new qualifiers, full database link.	Previous assessment as starting point; size, trend, status transitions.	Improved uncertainty vocabulary but still not a mortality or causal audit.

Table 5. Function of the NZTCS series and limits for mortality accountability

The objective criticism is therefore not that NZTCS incorrectly fails to perform a task outside its mandate. It is that New Zealand lacks an equally visible complementary product that performs the missing task. When the classification report becomes the recurring public synthesis, its narrow structure can be mistaken for the full conservation analysis. The solution is not to turn NZTCS into an unmanageably broad review, but to pair it with an annual or biennial mortality-intelligence report and a periodic causal synthesis.

5.7 The system recognises multiple threats but does not publicly compare them as competing or interacting models

The 2021 research strategy shows that government agencies understand the possibility of multiple and overlapping threats. It lists fishing, disease, seismic and mining activity, vessels, oil spills, pollution and sediment, coastal development, and climate change, and identifies effects on injury, disease, disturbance, prey, foraging, displacement, and habitat fragmentation. This breadth is a strength. Yet the recurring NZTCS reports do not present a structured comparison of causal hypotheses, and the public incident system is organised around observed events and case-level causes rather than population-level mechanisms.

A mature causal framework would distinguish at least four levels. The first is immediate death, such as entanglement, vessel strike, or acute disease. The second is underlying susceptibility, including chronic infection, immune impairment, nutritional stress, or prior injury. The third is vital-rate change not necessarily visible in carcasses,

including reduced conception, fetal loss, calf survival, or recruitment. The fourth is population context, including fragmentation, sex ratio, genetic diversity, habitat displacement, and prey dynamics. A factor may have little representation as an immediate necropsy cause while still altering reproduction or susceptibility.

Chemical contamination belongs in this framework as one hypothesis among many. It may be irrelevant, locally important, synergistic, or historically important; the present paper does not decide among these possibilities. The same restraint applies to bycatch, disease, predation, prey change, and habitat disturbance. The methodological requirement is that the system state what evidence supports each mechanism, what evidence conflicts with it, what is unknown, and what observation would discriminate among alternatives. Absent such a structure, the most visible and easily counted cause may dominate institutional attention regardless of whether it explains all demographic features.

This is where repeated, template-driven assessment can create epistemic inertia. Beginning each cycle from the previous classification is efficient and preserves comparability, but it is better at updating known variables than discovering omitted ones. New causes must arrive through submissions, expert networks, or commissioned work and then be translated into population size, trend, or a recognised qualifier. If a mechanism produces subtle reproductive or interactive effects, it may remain peripheral even when the population fails to recover as expected. Periodic “zero-base” causal reviews are therefore needed alongside incremental updates.

5.8 Public tracking of the research strategy is incomplete

The 2021 research strategy promises that current projects will be updated regularly, that links to outputs will be included, and that research and the strategy will be reviewed annually. The public research page continues to link a current-project spreadsheet supplied here as an October 2022 snapshot. It contains 15 projects across life history, habitat and diet, necropsy, disease, genetics, abundance, distribution, and fishing mitigation. Several report fields are blank or marked as delayed due to COVID. The spreadsheet is useful, but it is not a current implementation matrix for the strategy’s ranked research needs.

We did not identify a public crosswalk showing, for every priority, the funded project, budget, start and completion date, deliverable, peer-review status, management uptake, and remaining uncertainty. This does not demonstrate that annual reviews did not occur internally or that research stopped; the public pages list several newer projects. It does show that an external reader cannot readily audit whether priority 7 on cryptic mortality, priority 10 on recreational/cultural/illegal bycatch, or priority 13 on seasonal carcass collection has been completed, partly addressed, deferred, or superseded.

The gap is particularly relevant to experts who try to contribute evidence. The 2024 NZTCS call received seven submissions covering seven species. The report acknowledges advice from a wider group, but the public document does not provide a submission register, issue-by-issue response, accepted or rejected evidence table, or dissenting opinions. A contributor can therefore see the final category but not

necessarily how a specific concern was evaluated. A transparent evidence-response matrix would reduce perceptions of disregard even where the panel ultimately rejects the submission.

5.9 Fragmented responsibility weakens end-to-end accountability

Mortality information is produced by a distributed institutional network. DOC receives public reports and maintains the incident database; Fisheries New Zealand regulates commercial reporting, observers, and cameras; Massey University and other specialists conduct pathology; universities and archives hold genetic and biological samples; museums and local institutions preserve historical specimens and provenance; researchers build abundance and risk models; independent panels assign NZTCS status. Distribution of expertise is unavoidable and often beneficial. The weakness arises when no public mechanism owns the complete evidence chain.

Interface failures are therefore more plausible than a single, centrally controlled failure. A museum correction may not propagate to a DOC table; a fisheries event may have obscured location detail; a pathology identifier may not be transparently linked to a historical event ID; a genetic sample may clarify species identity without updating the original record; an NZTCS panel may receive a summary rather than the primary data. Effective governance requires common identifiers, interoperability, custody rules, and explicit responsibility for reconciliation.

6. Discussion

6.1 Overall effectiveness: substantial activity, incomplete analytical performance

The evidence does not support a simple verdict of institutional failure. New Zealand has legal reporting, a public incident database, published pathology, repeated abundance work, long-term demographic research, fisheries restrictions, camera deployment, a multi-threat research strategy, and a national classification system. Modern provenance is far better than historical provenance, and camera verification is rapidly improving. These are substantive achievements.

Nor does the evidence support an equally simple claim of effectiveness. The public system cannot presently estimate the relationship between recorded incidents and total mortality; historical source lineage is often absent; complete necropsy and cause resolution remain selective; fisheries verification metrics cover only reviewed activity; public correction and versioning are underdeveloped; research priorities are not tracked through a current public implementation matrix; and the main recurring status reports do not audit the mortality observation process or compare causal models. An active system can still leave decisive questions unanswered.

Function	Assessment	Basis	Priority implication
Statutory and strategic framework	Strong	Clear legal reporting duty, TMP, sanctuaries, research strategy.	Preserve; improve evidence-accountability requirements.

Function	Assessment	Basis	Priority implication
Public incident access	Moderate to strong	Quarterly spreadsheet, annual tables, pathology links, explicit caveats.	Add versioning, provenance, release metadata, and correction log.
Completeness of mortality detection	Unknown/weakly quantified	Reported-event system; public and carcass-recovery biases; no national correction factor identified.	Fund recoverability, drift, detection-effort, and cryptic-mortality models.
Historical data provenance	Weak	Large shares of early records have unknown source holder; candidate inconsistencies cannot be resolved publicly.	Independent legacy audit and source digitisation.
Modern provenance	Improved but incomplete	Source retention greatly improved; many locations remain estimated.	Confidence geometry, source IDs, geocoding record, and automated QA.
Cause-of-death resolution	Moderate for recovered bodies; weak for all deaths	Necropsy programme is valuable; recovery and suitability are selective; 31% unresolved in 2017–2024 coding.	Publish denominators, exclusions, confidence, and multi-factor cause fields.
Fisheries verification	Rapidly improving	Cameras on 220 vessels and 25% event review in Q1 2026; high concordance within reviewed detections.	Risk-based review, post-release studies, and annual cross-system reconciliation.
Correctability and external challenge	Partial/weakly visible	Ad hoc corrections are published, but no standard public adjudication and version-history system identified.	Create an independent challenge and correction protocol.
NZTCS risk classification	Strong for standardisation; insufficient for mortality accountability	Consistent categories and trend codes; change-centred summaries; inherited starting point.	Pair with a separate mortality-intelligence and causal-review product.
Adaptive institutional learning	Partial	Research strategy identifies gaps, but public priority-to-output tracking is incomplete.	Annual public dashboard and management-uptake register.

Table 6. Summary assessment of public-system effectiveness.

The most defensible conclusion is therefore that not everything reasonably possible is being done at the analytical level. This is not a judgement about the effort or motives of individual officials. It is a judgement about system design. The conservation stakes justify a more complete architecture for detecting unseen mortality, documenting uncertainty, auditing legacy records, and demonstrating how new evidence changes policy.

6.2 The visibility paradox

The policy system is highly visible: plans, sanctuaries, consultation, media releases, categories, population estimates, and public databases all signal institutional attention. Visibility can nevertheless create a false sense of evidentiary closure. A category such as Nationally Critical appears precise, while the underlying trend, detection probabilities, causes, and confidence may remain uncertain. Repeated publication of the same category can look like monitoring continuity without showing whether the causes of non-recovery have been understood.

Bureaucratic systems value standardisation because it makes decisions comparable, repeatable, and defensible. Those are legitimate values. The problem arises when reproducibility of the template is confused with adequacy of the model. Ecological processes can change, interacting threats can emerge, and old records can be reinterpreted. An effective public institution must preserve stable measurement while retaining the ability to revise what it measures. Stability without a mechanism for surprise becomes epistemic inertia (31, 46).

A related issue concerns the incentive structure surrounding environmental expertise. Conservation organisations, consultants, universities, and independent specialists may depend on public grants, contracts, permits, data access, or durable working relationships with government agencies. Such dependence does not establish compromised science or institutional loyalty, and the present study did not audit funding relationships. It can, however, increase the perceived professional cost of sustained public criticism and favour technical collaboration over scrutiny of the evidence architecture itself. In the dolphin case, we found extensive debate over particular threats but no recurring independent public audit of the national mortality-accounting system. The scarcity of visible system-level criticism should therefore not be treated as evidence that the system is already adequate.

6.3 Why these gaps affect every causal hypothesis

A weak mortality record does not merely disadvantage alternative explanations. It can also overstate confidence in established explanations. If fishing deaths are underdetected, bycatch risk is underestimated. If recovered carcasses overrepresent particular regions or seasons, disease prevalence may be misestimated. If chronic exposure reduces reproduction rather than causing acute death, pathology records may not capture its population effect. If historical locations are wrong, spatial association with fisheries, pollution sources, or habitat can be distorted. If population structure is incompletely known, local decline may be obscured by national aggregation.

Accordingly, the paper's neutrality on the chemical hypothesis is not evasive. It is the methodological point. A credible system should neither privilege nor exclude a cause because it is institutionally familiar. It should specify evidence pathways for all plausible mechanisms and distinguish direct proof, association, biological plausibility, model-based attribution, and unresolved uncertainty. Better mortality intelligence would strengthen the case for whichever explanations survive comparison.

6.4 Anticipated objections

One objection is that NZTCS was never intended to provide causal analysis. We agree. The criticism is the absence of an equally authoritative companion analysis, not a demand that the classification report become something else. A second objection is that historical data are necessarily imperfect. Again, we agree. Imperfection is precisely why provenance, confidence, and correction history are needed. A third objection is that total mortality can never be observed. That is true, but imperfect detection can be estimated and uncertainty can be propagated rather than left implicit.

A fourth objection is that onboard cameras now solve the bycatch-verification problem. Cameras substantially improve the system, but their evidential value depends on which vessels and events are in scope, the fraction and selection of footage reviewed, image quality, detection performance, compliance with camera operation, and treatment of post-release outcomes. A fifth objection is that the external dossier is unverified. We agree and have treated it accordingly. Its value is to expose whether the institutional system can adjudicate a specific challenge, not to pre-judge the answer.

6.5 A test of institutional learning

The strongest test of a conservation system is not whether it can publish a category, but whether it can learn when expected recovery does not occur. An adaptive loop requires an explicit conceptual model, predictions about the effect of interventions, monitoring of the relevant vital rates, comparison with outcomes, and revision of both management and causal assumptions (30, 31). The present public architecture contains pieces of this loop but not a single auditable chain.

For example, fishing restrictions may be followed by changes in adult survival, as at Banks Peninsula. That is a strong management test. But where a population remains small or declines, the system should publish a structured residual-risk analysis: which predicted improvements occurred, which did not, which vital rates remain limiting, which deaths are unobserved, and which alternative mechanisms gained or lost support. Without that step, repeated classification may document continued risk without explaining policy performance.

7. A national dolphin mortality-intelligence framework

The reforms below are intended to complement, not replace, the present TMP, incident database, fisheries monitoring, necropsy programme, abundance studies, and NZTCS. Their purpose is to connect existing components into a single learning system. Implementation should involve DOC, Fisheries New Zealand, pathology and genetics

providers, museums, independent statisticians, data stewards, and mana whenua/iwi, with clear protections for sensitive information and data sovereignty.

Priority	Reform	Minimum output	Performance indicator
1	Publish an annual Dolphin Mortality Intelligence Report separate from NZTCS.	Reported incidents, estimated unseen mortality, surveillance effort, necropsy flow, cause confidence, bycatch verification, corrections, and management implications.	Report released on schedule with reproducible data and code.
2	Create common persistent event and individual identifiers across DOC, Fisheries NZ, pathology, genetics, museums, and publications.	Crosswalk table and machine-readable links to every derivative record.	Share of records linked end-to-end; unresolved-link backlog.
3	Preserve field-level provenance.	Original source identifier, archive location, acquisition date, transcriber, geocoder, and evidence attachment for each material field.	Percentage of records with auditable source lineage by era.
4	Introduce public versioning and an immutable correction log.	Dataset release ID, schema version, checksums, and field-level changes with reasons.	All corrections traceable; no unexplained record disappearance or mutation.
5	Add uncertainty geometry and confidence.	Date confidence, location polygon/radius, species confidence, cause confidence, and reason codes.	Reduction in unqualified point locations and ambiguous missing values.
6	Automate data-quality checks.	Duplicate-ID tests, near-duplicate detection, impossible dates, geospatial outliers, inconsistent specimen IDs, and cross-tab integrity checks.	QA report published with each release; errors resolved before publication.
7	Estimate cryptic mortality and carcass recoverability in New Zealand waters.	Drift models, tagging/experimental work where ethical, active-search studies, and region/season-specific detection probabilities.	Unseen-mortality intervals incorporated into risk models.

Priority	Reform	Minimum output	Performance indicator
8	Reconcile fisheries evidence annually.	Linked fisher reports, observers, camera review, compliance outcomes, retained bodies, imagery, and pathology.	Mismatch rates by fleet/method; review coverage weighted by protected-species risk.
9	Standardise the necropsy evidence flow.	Case-flow diagram, exclusion reasons, immediate/underlying/contributing causes, ancillary tests, and confidence.	Cause resolution reported against all recovered and all reported deaths.
10	Maintain a multi-causal hypothesis register.	For each mechanism: predictions, supporting and conflicting evidence, missing tests, interactions, and population-model representation.	Periodic zero-base causal review and documented changes in evidential weight.
11	Create an external evidence-challenge protocol.	Acknowledgement, source preservation, independent reviewer, response deadline, reasoned outcome, and correction where needed.	All challenges publicly logged and closed or assigned a status.
12	Publish a live Research Strategy implementation dashboard.	Priority-to-project crosswalk, budgets, dates, outputs, peer review, management uptake, and residual gaps.	Annual review visible; no priority left without status or rationale.

Table 7. Recommended reforms and measurable outputs.

7.1 Independent historical audit

A focused historical audit should begin with the North Island west-coast records used in early distribution and threat analyses, because their small number gives individual entries disproportionate influence. The audit should digitise original museum and agency notes, reconcile specimen numbers, contemporary media, publications, DNA/tissue archives, and database entries, and publish an event-by-event determination. Records should not be silently deleted. Where two entries are merged or a location changes, the old version, reason, evidence, and downstream analyses affected should be preserved.

7.2 Detection and recoverability studies

New Zealand-specific carcass drift and recovery work is essential. A programme could combine oceanographic drift models, historical wind and current data, ethically designed tagged-carcass or surrogate studies, structured shoreline searches, and sensitivity analysis. It should estimate probabilities by region, season, carcass condition, distance offshore, and population. These probabilities would not produce exact

mortality, but they would convert an unqualified minimum count into a range suitable for risk assessment.

7.3 Risk-based electronic monitoring

Camera review should be optimised for conservation questions. Random review is needed for unbiased compliance estimates, while risk-based review can concentrate on vessels, methods, seasons, and places with higher protected-species interaction probability. Automated detection can support, but not replace, expert review. Quarterly public metrics should distinguish footage recorded, footage usable, events selected, selection rule, protected-species events detected, fisher concordance, unreported events, investigation outcomes, and species-identification confidence (47, 48).

7.4 A zero-base causal review

At least once per full NZTCS cycle, an independent multidisciplinary review should rebuild the causal model from the evidence rather than only update the previous assessment. It should include fisheries science, veterinary pathology, toxicology, epidemiology, population ecology, genetics, oceanography, data science, and mātauranga Māori. The review should not attempt to select a single cause prematurely. Its purpose is to identify which combinations of mechanisms can reproduce observed mortality, reproduction, spatial structure, and trend, and which new data would most efficiently discriminate among them.

8. Conclusion

New Zealand has devoted substantial public attention and real scientific resources to Hector's and Māui dolphin conservation. The resulting system contains strong components: legal reporting, public incident publication, pathology and genetics, repeated abundance studies, spatial protections, a broad research strategy, and rapidly expanding electronic monitoring. Any analysis that ignores these achievements would be inaccurate.

The same evidence nevertheless supports serious doubt about analytical effectiveness. The public incident record is detection-dependent and lacks a national estimate of unseen mortality. Legacy provenance is often weak; modern source retention has improved but public versioning and correction history remain limited. Necropsy results are conditioned on recovery and carcass quality. Fisheries verification is improving but cannot yet be treated as complete mortality accounting. External record challenges do not enter a clearly visible adjudication pathway. Research priorities are not linked to a current public implementation dashboard.

Most importantly, the NZTCS series is a classification system, not a causal diagnosis or an evaluation of policy performance. Its standardisation is useful, but its change-centred format can label a continuing severe decline as “no change” and can reproduce inherited analytical frames. The appropriate response is not to abandon classification. It is to build the missing companion system: a transparent, versioned, multi-source mortality-intelligence framework that quantifies observation losses, compares

competing causes, records corrections, and evaluates whether management produces the predicted demographic response.

The answer to the article's central question is therefore qualified but clear. New Zealand's conservation agencies are doing much, and some of the most important tools are improving. Yet the public evidence does not show that everything reasonably possible is being done to account for dolphin deaths and learn from them. For populations near the limits of viability, the distance between a reported incident and a true mortality estimate is not a technical detail. It is part of the conservation problem itself.

Declarations

Data availability

The analysis uses publicly available DOC and Fisheries New Zealand materials and two public spreadsheets supplied with the manuscript: the Hector's and Māui dolphin incident database export (records through April 2026) and the current-research-project snapshot dated October 2022. The correspondence dossier is an author-supplied source document and should be shared with reviewers subject to redaction of personal data. Before submission, the authors should notify DOC in accordance with the database user guide and provide the full coding protocol and candidate-record list.

Author contributions

A.G. assembled the historical correspondence dossier and identified candidate record discrepancies. D.N. developed the comparative analysis of the NZTCS series and the conceptual framing. Both authors contributed to interpretation and manuscript development. The final contribution statement should be confirmed before submission.

Competing interests

A.G. compiled the correspondence dossier examined in this paper. This relationship to one source of evidence is disclosed and is the reason the dossier is treated only as a source of audit hypotheses. The authors declare no financial competing interests. This statement should be reconfirmed before submission.

Funding

No external funding was declared for preparation of this draft. This statement should be confirmed before submission.

Ethics and evidentiary statement

No live animals or human research participants were involved. The manuscript does not make findings of deliberate wrongdoing. Any final published version should offer the relevant institutions an opportunity to correct factual errors and should distinguish clearly between documented records, author allegations, and analytical inference.

References

1. Currey, R.J.C., Boren, L.J., Sharp, B.R. and Peterson, D. (2012). A Risk Assessment of Threats to Māui's Dolphins. Ministry for Primary Industries and Department of Conservation, Wellington.
2. Roberts, J.O., Webber, D.N., Roe, W.D., Edwards, C.T.T. and Doonan, I.J. (2019a). Spatial risk assessment of threats to Hector's and Māui dolphins (*Cephalorhynchus hectori*). New Zealand Aquatic Environment and Biodiversity Report No. 214. Fisheries New Zealand, Wellington. <https://www.mpi.govt.nz/dmsdocument/35007-AEBR-214-Spatial-risk-assessment-of-threats-to-HectorsMaui-dolphins-Cephalorhynchus-hectori>
3. Roberts, J.O., Constantine, R. and Baker, C.S. (2019b). Population effects of commercial fishery and non-fishery threats on Māui dolphin (*Cephalorhynchus hectori maui*). New Zealand Aquatic Environment and Biodiversity Report No. 215. Fisheries New Zealand, Wellington. <https://www.mpi.govt.nz/dmsdocument/35010-AEBR-215-Population-Effects-of-Commercial-Fishery-and-Non-Fishery-Threats-on-Maui-Dolphins-Cephalorhynchus-hectori-maui>
4. Department of Conservation and Fisheries New Zealand (2020). Hector's and Māui Dolphin Threat Management Plan 2020. Wellington. <https://www.doc.govt.nz/our-work/protecting-species/protecting-marine-species/our-work-with-maui-dolphin/hectors-and-maui-dolphin-threat-management-plan/>
5. Department of Conservation (DOC) (2021a). Hector's and Māui Dolphin Research Strategy 2021. Department of Conservation, Wellington. <https://www.doc.govt.nz/globalassets/documents/conservation/native-animals/marine-mammals/maui-hectors-dolphins/2021-hectors-and-maui-5-year-research-plan.pdf>
6. Department of Conservation (DOC) (2026d). Māui dolphin abundance estimate. Accessed 26 July 2026. <https://www.doc.govt.nz/our-work/protecting-species/protecting-marine-species/our-work-with-maui-dolphin/maui-dolphin-abundance-estimate/>
7. MacKenzie, D.I. and Clement, D.M. (2014). Abundance and distribution of ECSI Hector's dolphin. New Zealand Aquatic Environment and Biodiversity Report No. 123. Ministry for Primary Industries, Wellington.
8. MacKenzie, D.I. and Clement, D.M. (2016). Abundance and distribution of WCSI Hector's dolphin. New Zealand Aquatic Environment and Biodiversity Report No. 168. Ministry for Primary Industries, Wellington.
9. Gibbs, A. (2026). Beyond the TCDD Lens in Paritutu New Plymouth, New Zealand: Invisible Phenoxy-Herbicide Co-Contaminants, Visible Developmental Signals, and Decision-Making under Incomplete Evidence. Pollution and Diseases, May. <https://doi.org/10.66659/ar7fqk65>
10. Gibbs, A. and Nikolaenko, D. (2026). Competing Hypotheses and Scientific Non-Recognition: Investigating the Hector's and Māui Dolphin Decline in the Context of Wartime Chemical Legacies. Pollution and Diseases, June. <https://doi.org/10.66659/10csws50>

11. Gibbs, A. and Nikolaenko, D. (2026). Chemical Pollution and the Population Crisis of Hector's and Māui Dolphins: A Critical Research Bibliography, 1970-2026. *Pollution and Diseases*, June, 51 pp. <https://doi.org/10.66659/ybmw2c82>
12. Gibbs, A. and Nikolaenko, D. (2026). Structural Falsification, Scientific Silence, and a New Research Strategy: A Critical Reconstruction of the Hector's and Māui Dolphin Case. *Pollution and Diseases*, June, 23 pp. <https://doi.org/10.66659/tms16797>
13. Gibbs, A. and Nikolaenko, D. (2026). Demographic Precision and Causal Narrowness: A Critical Review of Bennington et al. 2026 on Hector's Dolphin Fecundity and Bycatch Management. *Pollution and Diseases*, July, 21 pp. <https://doi.org/10.66659/ctn59c96>
14. Hitchmough, R. (comp.) (2002). New Zealand Threat Classification System Lists - 2002. Threatened Species Occasional Publication 23. Department of Conservation, Wellington. <https://www.doc.govt.nz/globalassets/documents/science-and-technical/tsop23.pdf>
15. Hitchmough, R., Bull, L. and Cromarty, P. (comps) (2007). New Zealand Threat Classification System Lists - 2005. Department of Conservation, Wellington. <https://www.doc.govt.nz/globalassets/documents/science-and-technical/sap236.pdf>
16. Baker, C.S., Chilvers, B.L., Constantine, R., DuFresne, S., Mattlin, R.H., van Helden, A. and Hitchmough, R. (2010). Conservation status of New Zealand marine mammals (suborders Cetacea and Pinnipedia), 2009. *New Zealand Journal of Marine and Freshwater Research* 44, 101-115. <https://doi.org/10.1080/00288330.2010.482970>
17. Baker, C.S., Chilvers, B.L., Childerhouse, S., Constantine, R., Currey, R., Mattlin, R., van Helden, A., Hitchmough, R. and Rolfe, J. (2016a). Conservation status of New Zealand marine mammals, 2013. New Zealand Threat Classification Series 14. Department of Conservation, Wellington. <https://www.doc.govt.nz/globalassets/documents/science-and-technical/nztcs14entire.pdf>
18. Baker, C.S., Boren, L., Childerhouse, S., Constantine, R., van Helden, A., Lundquist, D., Rayment, W. and Rolfe, J.R. (2019). Conservation status of New Zealand marine mammals, 2019. New Zealand Threat Classification Series 29. Department of Conservation, Wellington. <https://www.doc.govt.nz/globalassets/documents/science-and-technical/nztcs29entire.pdf>
19. Lundquist, D., Boren, L., Childerhouse, S., Constantine, R., van Helden, A., Hitchmough, R., Michel, P., Rayment, W. and Baker, C.S. (2025). Conservation status of marine mammals in Aotearoa New Zealand, 2024. New Zealand Threat Classification Series 48. Department of Conservation, Wellington. <https://www.doc.govt.nz/globalassets/documents/science-and-technical/nztcs48.pdf>
20. Department of Conservation (DOC) (2022). Hector's and Māui dolphin current research projects. Public spreadsheet snapshot dated October 2022; supplied with the manuscript. Current access is through the DOC research page. <https://www.doc.govt.nz/our-work/protecting-species/protecting-marine-species/our-work-with-maui-dolphin/maui-dolphin-research/>

21. Department of Conservation (DOC) (2026b). Research on Hector's and Māui dolphin. Accessed 26 July 2026. <https://www.doc.govt.nz/our-work/protecting-species/protecting-marine-species/our-work-with-maui-dolphin/maui-dolphin-research/>
22. Department of Conservation (DOC) (2021b). Hector's and Māui Dolphin Incident Database - External User Information. Created 21 July 2020; updated 22 February 2021. Department of Conservation, Wellington. <https://www.doc.govt.nz/globalassets/documents/conservation/native-animals/marine-mammals/hectors-maui-incidents/hector-maui-dolphin-incident-database-external-user-guide-2021.pdf>
23. Department of Conservation (DOC) (2026a). Hector's and Māui dolphin incident database. Accessed 26 July 2026. <https://www.doc.govt.nz/dolphinincidents>
24. Department of Conservation (DOC) (2026e). Hector's and Māui Dolphin Incident Database: public spreadsheet export, records through April 2026. Dataset supplied with the manuscript. <https://www.doc.govt.nz/dolphinincidents>
25. Marine Mammals Protection Act 1978, New Zealand Public Act 1978 No. 80, section 16 (latest version accessed 26 July 2026). <https://www.legislation.govt.nz/act/public/1978/0080/latest/DLM25350.html>
26. Childerhouse, S. and Johnston, O. (2021). Identification of marine mammals captured in New Zealand fisheries: 1 July 2019 to 30 June 2020. Cawthron Report No. 3706, prepared for the Department of Conservation. <https://www.doc.govt.nz/our-work/conservation-services-programme/csp-reports/201920/identification-of-marine-mammals-captured-in-nz-fisheries-2019-20/>
27. Fisheries New Zealand (2026a). Accidental captures of seabirds and protected marine species by commercial fishers from 1 January to 31 March 2026. Accessed 26 July 2026. <https://www.mpi.govt.nz/fishing-aquaculture/sustainable-fisheries/managing-the-impact-of-fishing-on-protected-species/accidental-captures-of-seabirds-and-protected-marine-species-by-commercial-fishers-from-1-january-to-31-march-2026>
28. Fisheries New Zealand (2026b). Commercial fishing monitored by fisheries observers. Last reviewed 29 May 2026; accessed 26 July 2026. <https://www.mpi.govt.nz/fishing-aquaculture/sustainable-fisheries/commercial-fishing-monitored-by-fisheries-observers>
29. Gibbs, A. (2026). Correspondence dossier on historical Hector's and Māui dolphin records, supplied to D. Nikolaenko. Unpublished author material, 14 pages.
30. Nichols, J.D. and Williams, B.K. (2006). Monitoring for conservation. Trends in Ecology & Evolution 21, 668-673. <https://doi.org/10.1016/j.tree.2006.08.007>
31. Lindenmayer, D.B. and Likens, G.E. (2009). Adaptive monitoring: a new paradigm for long-term research and monitoring. Trends in Ecology & Evolution 24, 482-486. <https://doi.org/10.1016/j.tree.2009.03.005>
32. Lindenmayer, D.B. and Likens, G.E. (2010). The science and application of ecological monitoring. Biological Conservation 143, 1317-1328. <https://doi.org/10.1016/j.biocon.2010.02.013>
33. Lindenmayer, D.B., Piggott, M.P. and Wintle, B.A. (2013). Counting the books while the library burns: why conservation monitoring programs need a plan for action.

- Frontiers in Ecology and the Environment 11, 549-555.
<https://doi.org/10.1890/120220>
34. Lovett, G.M., Burns, D.A., Driscoll, C.T., Jenkins, J.C., Mitchell, M.J., Rustad, L., Shanley, J.B., Likens, G.E. and Haeuber, R. (2007). Who needs environmental monitoring? *Frontiers in Ecology and the Environment* 5, 253-260
 35. Gormley, A.M., Slooten, E., Dawson, S., Barker, R.J., Rayment, W., du Fresne, S. and Bräger, S. (2012). First evidence that marine protected areas can work for marine mammals. *Journal of Applied Ecology* 49, 474-480.
<https://doi.org/10.1111/j.1365-2664.2012.02121.x>
 36. Baker, C.S., Hamner, R.M., Cooke, J., Heimeier, D., Vant, M., Steel, D. and Constantine, R. (2013). Low abundance and probable decline of the critically endangered Māui dolphin estimated by genotype capture-recapture. *Animal Conservation* 16, 224-233. <https://doi.org/10.1111/j.1469-1795.2012.00590.x>
 37. Baker, C.S., Steel, D., Hamner, R.M., Hickman, G., Boren, L., Arlidge, W. and Constantine, R. (2016b). Estimating the abundance and effective population size of Māui dolphins using microsatellite genotypes in 2015–16, with retrospective matching to 2001–16. Department of Conservation, Auckland.
 38. Department of Conservation (DOC) (2026c). Māui dolphin toxicology and disease research. Accessed 26 July 2026. <https://www.doc.govt.nz/our-work/protecting-species/protecting-marine-species/our-work-with-maui-dolphin/maui-dolphin-toxicology-and-disease-research/>
 39. Peltier, H., Dabin, W., Daniel, P., Van Canneyt, O., Dorémus, G., Huon, M. and Ridoux, V. (2012). The significance of stranding data as indicators of cetacean populations at sea: modelling the drift of cetacean carcasses. *Ecological Indicators* 18, 278-290. <https://doi.org/10.1016/j.ecolind.2011.11.014>
 40. Carretta, J.V., Danil, K., Chivers, S.J., Weller, D.W., Janiger, D.S., Berman-Kowalewski, M., Hernandez, K.M., Harvey, J.T., Dunkin, R.C., Casper, D.R., Stoudt, S., Flannery, M., Wilkinson, K., Huggins, J. and Lambourn, D.M. (2016). Recovery rates of bottlenose dolphin (*Tursiops truncatus*) carcasses estimated from stranding and survival rate data. *Marine Mammal Science* 32, 349-362.
<https://doi.org/10.1111/mms.12264>
 41. Roe, W.D., Howe, L., Baker, E.J., Burrows, L. and Hunter, S.A. (2013). An atypical genotype of *Toxoplasma gondii* as a cause of mortality in Hector's dolphins (*Cephalorhynchus hectori*). *Veterinary Parasitology* 192, 67-74.
<https://doi.org/10.1016/j.vetpar.2012.11.001>
 42. Meyer, S., Robertson, B.C., Chilvers, B.L. and Krkošek, M. (2017). Marine mammal population decline linked to obscured by-catch. *Proceedings of the National Academy of Sciences USA* 114, 11781-11786. <https://doi.org/10.1073/pnas.1703165114>
 43. Baker, A.N., Smith, A.N.H. and Pichler, F.B. (2002). Geographical variation in Hector's dolphin: recognition of new subspecies of *Cephalorhynchus hectori*. *Journal of the Royal Society of New Zealand* 32(4), 713-727
 44. Townsend, A.J., de Lange, P.J., Duffy, C.A.J., Miskelly, C.M., Molloy, J. and Norton, D.A. (2008). New Zealand Threat Classification System Manual. Department of

Conservation, Wellington. <https://www.doc.govt.nz/globalassets/documents/science-and-technical/sap244.pdf>

45. Rolfe, J., Hitchmough, R., Michel, P., Makan, T., Cooper, J.A., de Lange, P.J., Townsend, A.J., Miskelly, C.M. and Molloy, J. (2022). New Zealand Threat Classification System Manual 2022. Part 1: Assessments. Department of Conservation, Wellington. <https://www.doc.govt.nz/globalassets/documents/science-and-technical/new-zealand-threat-classification-system-manual-2022-part-1-assessments.pdf>
46. Lindenmayer, D.B., Likens, G.E., Krebs, C.J. and Hobbs, R.J. (2010). Improved probability of detection of ecological “surprises”. *Proceedings of the National Academy of Sciences USA* 107, 21957-21962. <https://doi.org/10.1073/pnas.1015696107>
47. van Helmond, A.T.M., Mortensen, L.O., Plet-Hansen, K.S., Ulrich, C., Needle, C.L., Oosterwind, D., Kindt-Larsen, L., Catchpole, T., Mangi, S., Zimmermann, C. et al. (2020). Electronic monitoring in fisheries: lessons from global experiences and future opportunities. *Fish and Fisheries* 21, 162-189. <https://doi.org/10.1111/faf.12425>
48. Pierre, J.P., Dunn, A., Snedeker, A., Wealti, M., Cozza, A. and Carovano, K. (2024). Optimising the review of electronic monitoring information for management of commercial fisheries. *Reviews in Fish Biology and Fisheries* 34, 1707-1732. <https://doi.org/10.1007/s11160-024-09895-7>

Appendix A. Reproducible audit protocol for candidate legacy records

For each contested event, the audit should create a bundle containing: (1) the current public event and individual rows; (2) every previous database version available; (3) the original field or stranding form; (4) museum accession and provenance notes; (5) photographs, measurements, sex and age determinations; (6) specimen and tissue identifiers; (7) DNA/haplotype results; (8) contemporaneous media or witness documentation; (9) every publication in which the event was used; (10) the geocoding method and location uncertainty; (11) a chronology of alterations; and (12) a signed adjudication explaining whether the record is confirmed, corrected, merged, split, or left unresolved.

The adjudication should distinguish clerical duplication, legitimate multiple-animal events, repeated reporting of the same event, changed taxonomic identification, corrected date or location, and substantive disagreement about source interpretation. Downstream analyses that used a corrected record should be listed. The old value should remain retrievable, and the public database should expose a stable correction identifier.

Appendix B. Coding notes for Table 3

The descriptive audit was designed to assess public metadata coverage, not to reproduce the agencies’ scientific mortality classifications. The period 2017–2024 was used as the mature modern comparison period because 2025–2026 contained records still subject to quarterly publication, compliance, or necropsy updates. Percentages are calculated over individual mortality records and can therefore give more weight to multi-animal events. A future formal study should publish executable code, use a documented codebook for every observation type, and conduct sensitivity analyses with alternative definitions of mortality and unresolved cause.