

Aligning Research Statistical Modelling with Official Statistical Methods

Provides guidance on aligning research-derived statistical models with the concepts and classifications used in official ocean accounts.

Version 0.6.1 **Status** Draft **Stability** Applied

Last updated September 2026 **Responsible officer** Dr Jordan Gacutan

Drafting AI-assisted drafting from GOAP Secretariat source materials; source list to be published at the close of consultation.

DRAFT

1. Outcome

This Circular provides guidance on aligning research-derived statistical models with the concepts, classifications, and estimation standards used in official statistical production for ocean accounts.

3. Guidance Material

"Alignment" is understood across three operational dimensions:

3.1 Conceptual Framework

3.2 Reconciling Research and Official Methods

- Conceptual harmonisation – ensuring that the variables, spatial units, and time periods used in a

3.3 Model Validation and Uncertainty

research model correspond to the definitions and classifications of SEEA EA (e.g., ecosystem

3.4 Data Requirements and Sources

asset boundaries, condition variable definitions, service-flow categories as set out in SEEA EA

3.5 Reporting and Integration

Chapter 5).

4. Acknowledgements

5. References

- Classification mapping – establishing explicit correspondence between research-model outputs and the account rows, column headings, and measurement units used in extent, condition, and ecosystem services supply-and-use tables.

ABOUT THIS CIRCULAR

- Estimator reconciliation – determining whether, and under what conditions, a research-derived point estimate or distribution can substitute for, supplement, or be benchmarked against an

Aligning Research Statistical Models with Official Design-based Estimates (2026)

(GOAP Technical Guidance on Ocean Accounting, TG-4.11, v0.6.1), GOAP

Secretariat. [https://guidance.oceanaccounts.org/circulars/section-4/4-11-research-](https://guidance.oceanaccounts.org/circulars/section-4/4-11-research-statistical-modeling)

statistical-modeling

Compilers who complete this Circular will be able to (a) categorise a candidate research model by type and purpose, (b) apply a structured decision framework to determine how the model's outputs may enter official accounts, (c) specify a validation and uncertainty protocol consistent with the United Nations National Quality Assurance Frameworks (UN NQAF), and (d) document provenance in a way that meets international metadata standards.

© 2026 Global Ocean Accounts Partnership. Licensed under CC BY-NC 4.0.

📌 Emerging Methods within an Applied Circular

TG-4.11 is classified Applied: it provides practical implementation guidance grounded in established statistical frameworks, whilst several subsections address methods at the methodological frontier where consolidated international guidance is still developing. Specifically, Section 3.2 (reconciliation involving model-assisted estimation), Section 3.3 (Bayesian uncertainty quantification), and Section 3.4 (data requirements for machine learning models) cover Emerging approaches. Compilers applying these methods should treat associated guidance as indicative rather than prescriptive, and should monitor evolving UNECE HLG-MOS and Eurostat methodological publications for updates.

¶1 ⓘ Framework Position

¶2 TG-4.11 sits at the boundary between research science and official statistical production. It provides the methodological bridge by which NSO compilers assess, validate, and integrate research-derived statistical models into ocean accounts compiled in accordance with the *System of Environmental-Economic Accounting – Ecosystem Accounting* (SEEA EA).¹ Readers should be familiar with quality assurance principles (TG-0.7 Quality Assurance Principles) and research data integration practices (TG-4.5 Integrating Research Data into Official Statistics) before proceeding.

1. Outcome

¶1 This Circular provides guidance on aligning research-derived statistical models with the concepts, classifications, and estimation standards used in official statistical production for ocean accounts. “Alignment” is understood across three operational dimensions:

- ¶2 **Conceptual harmonisation** – ensuring that the variables, spatial units, and time periods used in a research model correspond to the definitions and classifications of SEEA EA (e.g., ecosystem asset boundaries, condition variable definitions, service-flow categories as set out in SEEA EA Chapter 5¹).
- ¶3 **Classification mapping** – establishing explicit correspondence between research-model outputs and the account rows, column headings, and measurement units used in extent, condition, and ecosystem services supply-and-use tables.
- ¶4 **Estimator reconciliation** – determining whether, and under what conditions, a research-derived point estimate or distribution can substitute for, supplement, or be benchmarked against an official survey-based or design-based estimate.

¶5 Compilers who complete this Circular will be able to (a) categorise a candidate research model by type and purpose, (b) apply a structured decision framework to determine how the model’s outputs may enter official accounts, (c) specify a validation and uncertainty protocol consistent with the United Nations National Quality Assurance Frameworks (UN NQAF)², and (d) document provenance in a way that meets international metadata standards.

¶6 ⓘ Emerging Methods within an Applied Circular

¶7 TG-4.11 is classified Applied: it provides practical implementation guidance grounded in established statistical frameworks, whilst several subsections address methods at the methodological frontier where consolidated international guidance is still developing. Specifically, Section 3.2 (reconciliation involving model-assisted estimation), Section 3.3 (Bayesian uncertainty quantification), and Section 3.4 (data requirements for machine learning models) cover Emerging approaches. Compilers applying these methods should treat associated guidance as indicative rather than prescriptive, and should monitor evolving UNECE HLG-MOS and Eurostat methodological publications for updates.

2. Requirements

- ¶1 **TG-0.1 General Introduction to Ocean Accounts** – foundational concepts and scope of ocean accounts
- ¶2 **TG-0.7 Quality Assurance Principles** – UN NQAF quality dimensions and QA protocols applied throughout this Circular
- ¶3 **TG-4.1 Remote Sensing and Geospatial Data** – geospatial and remote sensing inputs that frequently serve as primary data for research models
- ¶4 **TG-4.5 Integrating Research Data into Official Statistics** – broader data integration framework within which research modelling sits

3. Guidance Material

3.1 Conceptual Framework

3.1.1 Research Models in Ocean Accounting

¶1 As used in this Circular, research statistical models are quantitative models developed outside the official statistical production process, typically in academic, government research agency, or international organisation settings, whose outputs are candidates for incorporation into official ocean accounts. They fall into four broad classes:

Research model class	Examples in ocean context	Primary purpose
Ecological process models	InVEST (coastal protection, fishery), Atlantis (ecosystem dynamics), EwE (trophic structure)	Predict ecosystem service flows under scenarios
Spatial-statistical models	Kriging, species distribution models (MaxEnt, BRT), habitat suitability indices	Estimate spatial distribution or abundance from point observations
Econometric models	Hedonic pricing (coastal property values), travel cost (recreation), production function (fishery rent)	Estimate monetary values of ecosystem services; see TG-1.9 Safe Usage of Monetary Valuation
Machine-learning and AI models	Random forests for habitat mapping, neural networks for ocean colour classification	Classification or regression from high-dimensional remote sensing inputs

¶2 The last class, machine-learning (ML) and artificial intelligence (AI) models, is still developing as a basis for official statistics. Where ML methods are applied, compilers should consult the UNECE HLG-MOS guidance on machine learning for official statistics³, together with the parallel guidance issued by Eurostat⁴, FAO⁵ and the OECD⁶, and treat associated results with hedged uncertainty language pending further methodological consolidation.

3.1.2 Official Statistical Estimation Paradigms

¶1 Official statistical production relies on established estimation paradigms that carry known statistical properties. Compilers should understand how research model outputs relate to each paradigm:

Official paradigm	Description	Research model role
Design-based	Estimates derived from probability samples; variance from sampling design	Research model can serve as auxiliary variable or domain post-stratifier
Model-assisted	Design-based framework augmented by a working model to improve precision (e.g., GREG estimator)	Research model can supply the working model used for calibration
Model-based (small-area estimation)	Borrows strength across domains using mixed models; Rao-Molina framework ⁷	Research model can contribute covariates or specify the linking model
Synthetic and composite	Applies higher-level model to lower-level domains	Research model provides the synthetic component

¶2 The crosswalk between research model class and official paradigm determines the methodological requirements for integration, particularly the extent to which design-based uncertainty estimates remain available to bound the research-model contribution. Anchoring the crosswalk to

SEEA EA Chapter 5¹ and the UN-GGIM Integrated Geospatial Information Framework (IGIF)^{8, 9} keeps spatial and temporal resolution decisions coherent with account structure.

¶3 Key Takeaway

¶4 Not all research models can enter official accounts directly. The decision depends on which official estimation paradigm is available and on the bias, variance, and coherence properties of the research model relative to it.

3.2 Reconciling Research and Official Methods

3.2.1 Three-Mode Decision Framework

¶1 When a research-derived estimate is a candidate for inclusion in an official ocean account, compilers should classify the intended mode of use before proceeding with technical validation. Three modes are recognised. The full decision sequence, from mode classification through the four admissibility and governance gates, is summarised in Figure 4.11.1.

DRAFT

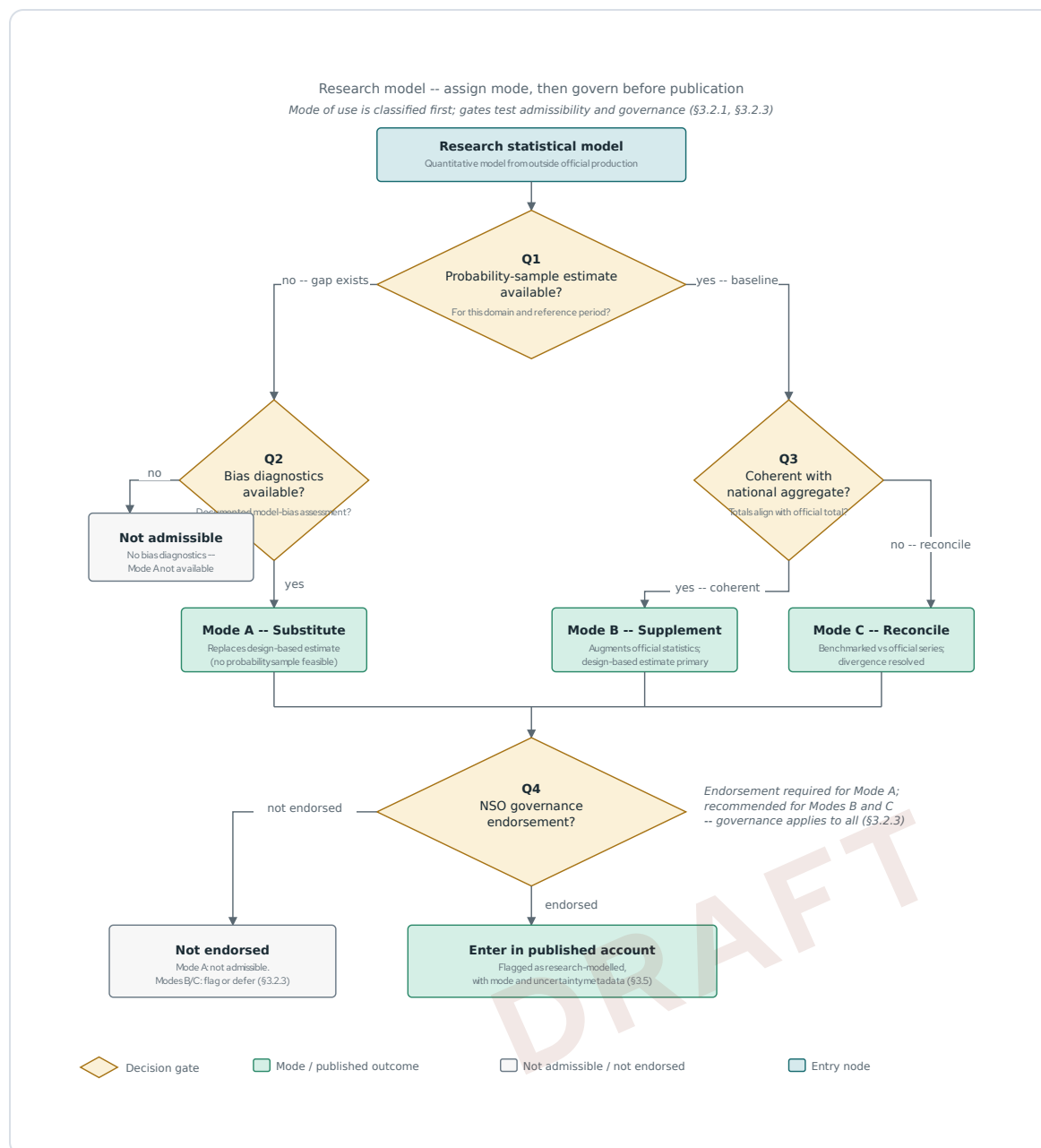


Figure 4.11.1 Four sequential gates route an external research statistical model to Mode A, B, or C use, or to rejection, before NSO endorsement.

Source: TG-4.11 §3.2.1 (three-mode decision framework, steps 1--4 and modes A--C) and §3.2.3 (governance acceptance).

¶2 **Mode A – Substitute.** The research estimate replaces a survey-based or design-based estimate. This is appropriate only when: (i) no probability-sample design is operationally feasible for the domain (e.g., deep-sea extent estimates); (ii) the research estimator has demonstrated negligible bias relative to an independent benchmark or external validator; and (iii) uncertainty is explicitly quantified and disclosed. Substitution requires the highest level of governance endorsement (see Section 3.2.3 below).

¶3 **Mode B – Supplement.** The research estimate is used as an auxiliary variable or covariate within an official design-based or model-assisted estimation framework, improving precision without replacing the probability-sample estimator. The design-based estimate remains the primary figure, whilst the research model improves its precision. This is the most common and least restrictive mode of integration, consistent with established model-assisted survey estimation practice.¹⁰

¶4 **Mode C – Reconcile.** An independently produced research estimate is benchmarked against an existing official total using temporal or spatial benchmarking techniques. Discrepancies are investigated and resolved, either by adjusting the research series or by flagging the official series for revision. This mode is appropriate when both a research-derived time series and an official survey series exist but diverge, and the compiler must determine which is more reliable for a given period or domain.

¶5 Compilers should apply the following decision sequence:

1. ¶6 Is a probability-sample-based (design-based) estimate available for this domain and reference period? If yes, consider Mode B first. Consider Mode A only where design-based estimates are demonstrably infeasible.
2. ¶7 Does the research estimator have documented bias diagnostics (e.g., cross-validation RMSE, comparison against holdout samples or external benchmarks)? If no, Mode A is not admissible.
3. ¶8 Is the research estimate coherent with any established national or regional aggregate (e.g., does the spatially disaggregated estimate sum to an independently known total)? Where no such aggregate exists, this coherence check should be conducted against the most closely related official or internationally reported figure available (e.g., FAO fishery statistics, regional remote sensing baselines). Document the reference used. Incoherence must be resolved before any mode is finalised.
4. ¶9 Has the NSO methodology committee or equivalent governance body endorsed the use of this model? Endorsement is required for Mode A, and recommended for Modes B and C.

3.2.2 Classification Mapping

¶1 Before entering an account, the research model's output variable must be mapped explicitly to an account row. This mapping should state: (a) the SEEA EA account type (extent, condition, ecosystem services supply, or monetary); (b) the measurement unit and conversion applied; (c) the spatial and temporal scope; and (d) any aggregation or disaggregation step performed. Misaligned classification mapping is a common source of double-counting or scope errors, for example including both a research-derived "coastal protection service flow" and a separately compiled defensive expenditure figure without netting.

3.2.3 Governance Acceptance

¶1 ⓘ Governance Requirement

¶2 A technically valid research model is not automatically eligible for official statistical use. Before a research model's outputs are incorporated into published ocean accounts, the following governance steps should be completed:

1. ¶3 **Peer review** – the model and its ocean-account application should have been reviewed by at least two independent domain experts not involved in the original modelling, with review documented and accessible.
2. ¶4 **Methodology committee endorsement** – the NSO or equivalent statistical authority should record formal endorsement of the model for the specific application and reference period, consistent with the UN Fundamental Principles of Official Statistics (Principle 2: professional standards and scientific principles).¹¹
3. ¶5 **Replication check** – a second analyst should be able to reproduce the model outputs from the stated inputs and code (see Section 3.4.3 on reproducibility).
4. ¶6 **Revision protocol** – a documented plan for updating or replacing the model estimate in future revision cycles, including triggers for re-validation.

3.3 Model Validation and Uncertainty

3.3.1 Validation Framework

¶1 Model validation should be structured around the UN NQAF quality dimensions as the umbrella framework,² applied as set out in Table 3.3.1.1 below for research models entering official accounts.

Quality dimension	Application to research models
Relevance	The model's output variable corresponds to the intended SEEA EA account variable (see Section 3.2.2).
Accuracy and reliability	Demonstrated through a documented validation protocol (see Section 3.3.2).
Timeliness	The model reference period aligns with the account reference period; lag between data collection and model output is documented.
Accessibility and clarity	Model documentation, code, and outputs are accessible to account users and auditors (see Section 3.4.3).
Coherence and comparability	Estimates are consistent with related account aggregates and comparable across time and space.
Completeness	Coverage of the domain is stated; gaps are flagged and quantified where possible.

3.3.2 Validation Protocol Requirements

¶1 Every research model submitted for inclusion in official accounts must have a documented validation protocol that addresses, at minimum:

- ¶2 **Internal validation** – at least one form of resubstitution, cross-validation (k-fold), or bootstrap validation on the training dataset, with performance metrics (e.g., RMSE, MAE, R^2) reported.
- ¶3 **Holdout validation** – performance metrics on a withheld test dataset not used in model fitting or tuning. For spatial models, holdout samples should be spatially blocked to avoid spatial autocorrelation inflating apparent accuracy.
- ¶4 **External benchmark** – where an independent dataset or survey estimate exists for any portion of the domain, model predictions must be compared against it. Systematic divergence exceeding an agreed tolerance threshold (e.g., $\pm 10\%$ of the benchmark estimate, or a threshold agreed in the model's validation protocol and recorded in its metadata – see Section 3.5.2) must be investigated before integration.

¶5 Compilers should document the validation protocol in the model's metadata record (see Section 3.5).

3.3.3 Tiered Uncertainty Reporting

¶1 Uncertainty in research-model estimates must be disclosed in the account record. The following three-tier framework provides minimum standards scaled to NSO analytical capacity:

¶2 **Tier 1 – Expert range (minimum requirement).** The compiler documents a plausible lower and upper bound derived from sensitivity analysis over key model parameters or input assumptions. The range should reflect at least the 10th–90th percentile of expert elicitation or scenario variation. This tier is accessible to all NSOs and constitutes the floor for uncertainty disclosure.

¶3 *Example:* A species distribution model produces a mean coral reef extent estimate of 12,400 ha. Sensitivity runs varying the habitat-suitability threshold between 0.35 and 0.65 produce a range of 10,800–14,100 ha. The account record states: "Estimated extent 12,400 ha (expert range 10,800–14,100 ha; threshold sensitivity)."

¶4 **Tier 2 – Monte Carlo confidence interval (recommended).** The compiler propagates parametric uncertainty through the model using Monte Carlo simulation (minimum 1,000 iterations), producing an empirical distribution of estimates. The 95% confidence interval (2.5th–97.5th percentile) is reported alongside the point estimate. Tier 2 is recommended whenever computational resources permit.

¶5 *Example:* Monte Carlo propagation of parameter uncertainty across 5,000 runs yields a 95% CI of 11,200–13,700 ha. The account record states: "Estimated extent 12,400 ha (95% CI 11,200–13,700 ha; Monte Carlo, 5,000 iterations)." (The Monte Carlo CI is narrower than the Tier 1 expert

range because it propagates only parametric uncertainty around a fitted model, whereas the expert range also captures structural uncertainty from threshold choice.)

¶6 **Tier 3 – Bayesian credible interval (advanced).** For models specified in a Bayesian framework, the posterior predictive distribution is used directly to report credible intervals. Where small-area estimation methods are applied (e.g., Rao–Molina mixed models⁷), the empirical best linear unbiased predictor (EBLUP) and its mean squared error estimate should be reported. Note that Bayesian approaches in marine small-area estimation are still developing – results should be labelled as experimental pending broader methodological consensus.

¶7 Key Takeaway

¶8 Tier 1 uncertainty disclosure is the minimum for any research-model estimate entering an official ocean account. NSOs should progress to Tier 2 as computational capacity allows. Tier 3 is appropriate only where the model is Bayesian by specification or small-area estimation methods are applied.

3.4 Data Requirements and Sources

3.4.1 Mapping Research Data Inputs to Account Rows

¶1 Research models used in ocean accounting draw on a range of input data types. The following table maps common input types to SEEA EA account targets, suitable model classes, and key quality caveats. Compilers should treat this as a starting checklist rather than an exhaustive specification, and should consult [TG-4.1 Remote Sensing and Geospatial Data](#) and [TG-4.4 Citizen Science and Community-Based Monitoring](#) for source-specific data quality guidance.

Input data type	SEEA EA account type	Suitable model class	Quality caveat
Satellite imagery (optical, SAR)	Extent (habitat mapping); Condition (spectral indices)	Spatial-statistical; ML/AI	Atmospheric correction; cloud cover; sensor drift across time series
Airborne LiDAR / acoustic bathymetry	Extent (reef, seagrass); Condition (structural complexity)	Spatial-statistical	Coverage gaps; temporal mismatch with account period
Ecological field survey (transects, trawls)	Condition (biotic variables); Extent (fine-scale)	Ecological process; spatial-statistical	Survey design may not be probability-based; spatial coverage limited
Citizen science observations	Condition (species presence/absence); Extent (intertidal)	Spatial-statistical (occupancy models)	Detection bias; spatial clustering near access points – see TG-4.4
Oceanographic sensors / Argo floats	Condition (temperature, DO, pH, salinity)	Ecological process; spatial-statistical	Calibration drift; spatial sparsity in coastal zones
Catch and effort logbooks	Flows from environment to economy (fish biomass removal)	Ecological process (stock assessment)	Reporting compliance; misidentification
Stock assessment model outputs (VPA, surplus production, integrated models such as SS3 or MULTIFAN-CL)	Asset accounts (fish biomass stock); Flows from environment to economy (sustainable yield proxy)	Ecological process (stock assessment)	Model uncertainty rarely reported as CI; point estimates common – apply Tier 1 uncertainty disclosure at minimum (see Section 3.3.3)
Household / firm survey data	Ecosystem services (recreation, subsistence, cultural)	Econometric	Recall bias; incomplete market coverage – see TG-4.2

Input data type	SEEA EA account type	Suitable model class	Quality caveat
Land-use / land-cover change data	Extent change (mangrove, seagrass loss); Condition (disturbance)	Spatial-statistical	Classification accuracy; minimum mapping unit

3.4.2 Minimum Data Quality Standards

¶1 Input datasets must meet minimum quality standards before being used to estimate account values. As a minimum, compilers should document: (a) spatial and temporal resolution and coverage; (b) known biases and their estimated magnitude; (c) calibration and cross-validation status; and (d) licence and access conditions. Where input datasets are derived from citizen science or community monitoring programmes, the additional quality considerations in [TG-4.4](#) apply.

3.4.3 Reproducibility Requirements

¶1 Research models feeding official accounts must be reproducible: a second analyst, given the same inputs and following the same documented steps, should be able to replicate the model outputs within numerical precision. Table 3.4.3.1 below summarises the requirements that apply.

Requirement	Description
Versioned code repository	Model code must be deposited in a publicly or institutionally accessible repository (e.g., GitHub, GitLab, national data repository) and assigned a persistent identifier (DOI or equivalent) that is recorded in the account metadata.
Model card	A structured summary document following the model card template (Mitchell et al. 2019 ¹²) or an equivalent agreed format should accompany the model, covering: intended use, model inputs, key assumptions, known limitations, and validation results.
Environment specification	The computational environment (software versions, dependencies) should be specified via a container image (e.g., Docker), environment lockfile (e.g., <code>requirements.txt</code> , <code>renv.lock</code>), or equivalent, so that the model can be re-run in a future revision cycle without dependency conflicts.

¶2 The reproducibility requirements in Table 3.4.3.1 align with the FAIR data principles (Findable, Accessible, Interoperable, Reusable) and with the broader open-statistics agenda reflected in UNECE HLG-MOS guidance.³

3.5 Reporting and Integration

3.5.1 Embedding Research-Modelled Values in Official Accounts

¶1 When research-modelled values are incorporated into published ocean accounts, the account record must allow users to identify and trace those values. Transparency serves two purposes. It preserves statistical integrity, since users need to understand what is officially surveyed and what is modelled. It also supports revision management, since modelled values may need updating when a better model or a new survey becomes available.

3.5.2 Minimum Metadata Fields

¶1 Every account cell or table entry that derives partly or wholly from a research model must be accompanied by, or linked to, a metadata record containing at minimum the following fields:

Metadata field	Content required
Source type	"Research model" (distinguishing from "official survey", "administrative data", "expert estimate")
Model identifier	Name, version, and persistent identifier (DOI or URL) of the model and code repository

Metadata field	Content required
Model card reference	Persistent identifier or location of the model card document
Input datasets	List of primary input datasets with version/date and source identifier
Validation summary	Validation protocol components applied (internal, holdout, external benchmark; Section 3.3.2) and key performance metric(s)
Uncertainty statement	Tier (1/2/3), method, and numerical range or interval (Section 3.3.3)
Reference period	Start and end date of the model's reference period
Spatial scope	Geographic extent and coordinate reference system
Governance endorsement	Date and authority of methodology committee or equivalent endorsement
Revision trigger	Condition under which the estimate will be updated in the next revision cycle

3.5.3 Dissemination Standards

¶1 Account dissemination should apply the international metadata and exchange standards covered in TG-4.6 Data harmonisation.¹³ In particular, quality flags should distinguish research-modelled cells from survey-based cells, for example using the SDMX Observation Status code “E” (estimated) or a user-defined code agreed within the national statistical system. Geographic metadata lineage should be completed for all spatially referenced research-model outputs, recording the source datasets, processing steps, and transformation methods applied.

¶2 ⚠ Common Misconception

¶3 Embedding a research-modelled value in an account table without a quality flag or metadata link gives it the same apparent status as a probability-sample-based estimate. This is misleading to users and obscures the revision risk associated with modelled values. All research-model-sourced cells must be flagged.

4. Acknowledgements

¶1 **Authors:** [To be confirmed]

¶2 **Reviewers:** [To be confirmed]

5. References

- ¶1 United Nations, European Commission, Food and Agriculture Organization of the United Nations, Organisation for Economic Co-operation and Development, & World Bank Group. (2021). *System of Environmental-Economic Accounting – Ecosystem Accounting (SEEA EA)*. United Nations. (Methodological anchor: Chapter 5, Ecosystem condition; Chapter 3–4, Account structure). <https://seea.un.org/ecosystem-accounting> (<https://seea.un.org/ecosystem-accounting>)
- ¶2 United Nations Statistics Division (UNSD). (2019). *United Nations National Quality Assurance Frameworks Manual for Official Statistics*. United Nations. <https://unstats.un.org/unsd/methodology/dataquality/> (<https://unstats.un.org/unsd/methodology/dataquality/>)
- ¶3 United Nations Economic Commission for Europe (UNECE), High-Level Group for the Modernisation of Official Statistics (HLG-MOS). (2021). *Machine Learning for Official Statistics*. UNECE.
- ¶4 European Commission, Eurostat. (2021). *Methodological guidance for machine learning in official statistics*. Publications Office of the European Union.
- ¶5 Food and Agriculture Organization of the United Nations (FAO). (n.d.). *Statistical Working Paper series on Machine Learning for Agricultural and Fisheries Statistics*. FAO.
- ¶6 Organisation for Economic Co-operation and Development (OECD). (n.d.). *Methodological guidelines and work on Machine Learning for Official Statistics*. OECD.
- ¶7 Rao, J. N. K., & Molina, I. (2015). *Small Area Estimation* (2nd ed.). Wiley.
- ¶8 United Nations Committee of Experts on Global Geospatial Information Management (UN-GGIM) & World Bank. (2018). *Integrated Geospatial Information Framework (IGIF) Part 1: Overarching Strategic Framework*. United Nations. <https://ggim.un.org/IGIF/> (<https://ggim.un.org/IGIF/>)
- ¶9 United Nations Committee of Experts on Global Geospatial Information Management (UN-GGIM) & World Bank. (2020). *Integrated Geospatial Information Framework (IGIF) Part 2: Implementation Guide*. United Nations. <https://ggim.un.org/IGIF/> (<https://ggim.un.org/IGIF/>)

0. ¶10 European Commission, Eurostat (European Statistical System). (2020). *Methodological Manual/Handbook on Small Area Estimation*. Publications Office of the European Union.
1. ¶11 United Nations Statistics Division (UNSD). (2014). *Fundamental Principles of Official Statistics*. United Nations. https://unstats.un.org/UNSDWebsite/statcom/session_45/documents/statcom-2014-45-fundamental-principles-of-official-statistics-E.pdf (https://unstats.un.org/UNSDWebsite/statcom/session_45/documents/statcom-2014-45-fundamental-principles-official-statistics-E.pdf)
2. ¶12 Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I. D., & Gebru, T. (2019). Model Cards for Model Reporting. In *Proceedings of the Conference on Fairness, Accountability, and Transparency (FAccT)*. ACM. <https://doi.org/10.1145/3287560.3287596> (<https://doi.org/10.1145/3287560.3287596>)
3. ¶13 Relevant dissemination and metadata standards include SDMX 3.0 (Statistical Data and Metadata eXchange), ISO 19115-1:2014 (geographic metadata), and the Data Documentation Initiative (DDI) Codebook – see [TG-4.6](#) for full treatment.

DRAFT