https://doi.org/10.5194/egusphere-2026-4208
Preprint. Discussion started: 21 August 2026

Knowing the ocean: an epistemological and communication framework for Physical Oceanography

Alvise Benetazzo[1], Tiziano Minuzzo[2]

[1] Consiglio Nazionale delle Ricerche, Istituto di Scienze Marine, CNR-ISMAR, Venice, Italy.
[2] Section Développement et Maintenance ICT, Service ICT, Direction Technologies Policières, Police Grand-Ducale, Luxembourg.

Correspondence to: Alvise Benetazzo (alvise.benetazzo@cnr.it)

Abstract. Physical Oceanography is the branch of marine sciences concerned with understanding the oceans as a geophysical fluid system, focusing on the processes that govern the spatial organisation and temporal variability of momentum, heat, and mass across a wide range of scales. Knowledge in Physical Oceanography is produced through a combination of observations, theory, numerical models, and statistical inference; yet its epistemological structure remains largely implicit. This paper develops a framework to clarify how such knowledge is generated, validated, and constrained by the ocean's vast spatial extent, continuous variability, and multiscale dynamics. We propose a two-dimensional epistemic map that classifies knowledge claims by their aim and representational form, providing a structured vocabulary for interpreting the diversity of oceanographic research. We illustrate the map with an exploratory classification of scientific synopses, and use it to read recent tendencies in the discipline. We further examine the epistemic status of oceanic numerical models and reanalysis systems, and present a case study of wind-wave spectral representation that exposes the underlying assumptions and limits of common practices. Building on these elements, we propose a diagnostic reporting template aimed at making explicit the assumptions, validation criteria, and epistemic payoff of individual studies. By rendering the epistemic structure of a result legible, the framework supports clearer interpretation, comparison, and communication of physical-oceanographic knowledge, both within the discipline and across the wider Earth-system sciences.

1 Introduction

Physical Oceanography (hereafter PO) is a branch of marine sciences concerned with the quantitative description and theoretical understanding of the ocean as a geophysical fluid system. PO addresses the governing conservation laws of mass, momentum, and energy, with particular focus on the mechanisms of energy input, transfer, dissipation, and the coupling between the ocean and external forcings such as the atmosphere, cryosphere, and solid Earth. By combining in situ and remote sensing measurements with dynamical frameworks and high-resolution numerical simulations, PO seeks to resolve the variability of oceanic phenomena across a broad range of scales, from bursting bubbles at timescales of a tenth of a second to the Milankovich timescales of 100,000 years and spatially from the Kolmogorov microscale (a tenth of a millimeter) to basin-wide gyres (thousands of kilometers), hereby elucidating the underlying mechanisms governing ocean dynamics and its variability (Cushman-Roisin and Beckers, 2011; Gill, 1982; Holthuijsen, 2007; Janssen, 2009; McPhee, 2008; Pedlosky, 2013; Talley et al., 2011; Vallis, 2017).

Given the complexity of the marine environment, PO knowledge cannot generally be obtained as a direct empirical read-off from observations alone; rather, it emerges from the synthesis of measurements, theoretical constructs, and numerical modelling, through which oceanic processes are

https://doi.org/10.5194/egusphere-2026-4208
Preprint. Discussion started: 21 August 2026

interpreted and reconciled across a wide spectrum of interacting spatial and temporal scales (Cushman-Roisin and Beckers, 2011; Emery and Thomson, 2004; Griffies, 2004; Oreskes et al., 1994; Pugh and Woodworth, 2014; Roemmich et al., 2009; Stewart, 2008; Thorpe, 2007). The inference of unobservable interior states from sparse data is itself a long-standing methodological concern of the field (Wunsch, 1996).

This paper argues it is necessary to interrogate how such knowledge is generated, justified, and limited, thereby motivating an explicit theory of knowledge for PO. In doing so, we draw on two distinct but related bodies of philosophy: *epistemology* proper, the study of the nature, justification, and limits of knowledge (Audi, 1997; BonJour, 2010), and the *philosophy of science*, which studies how scientific theories are tested, change, and progress (Kuhn, 1962; Lakatos, 1978; Popper, 1959). Together, these help us state what our PO claims mean, how they are warranted, and where they are valid and limited. We propose that such considerations are central in PO, because knowledge is generated under intrinsic constraints imposed by sparse and heterogeneous observations, strong nonlinear dynamics, and pervasive multiscale interactions. Consequently, understanding relies on integrating theoretical frameworks and numerical models, which mediate the interpretation of incomplete data and enable inference of underlying physical processes (Frigg and Hartmann, 2025; Oreskes et al., 1994; Winsberg, 2010)(Frigg and Hartmann, 2025; Oreskes et al., 1994; Winsberg, 2010).

This perspective motivates foundational questions for PO, which structure the analysis developed in this paper:

- What constitutes scientific knowledge in PO?
- What are the respective roles and mutual relationships among theory, numerical models, and observations?
- Through which mechanisms does the discipline achieve progress?
- On what grounds are scientific claims established and validated?

Since PO sits at the intersection of other disciplines, not only in the body of marine sciences, such as meteorology, climate science, biogeochemistry, and geophysics (e.g., Bigg, 1996; Karnauskas, 2020; Pereira and Pardal, 2024; Stewart, 2008), epistemology also helps clarify demarcation and unity, to judge if a small set of governing principles unifies the field, or it is best understood as an interdisciplinary patchwork tied together by shared modelling practices and observing infrastructures. Finally, the objectivity of PO results must be examined in context: societal needs (climate risk, coastal hazards, resources) shape funding, priorities, and even what counts as an "important" result in Science (Douglas, 2009; Edwards, 2022; Stanev, 2023). This does not automatically undermine objectivity, but it motivates explicit standards for transparency, uncertainty, and evidential strength. The PO's future then hinges on limits to predictability (chaos, turbulence, forcing uncertainty) and on whether current computational–observational paradigms are sufficient, or whether new representations and inference strategies are required.

Similarly to what is proposed here for PO, over the past three decades, climate science has increasingly engaged with the philosophy of science to clarify the epistemic status of its claims, the role of models, and the nature of uncertainty and consensus in complex Earth systems (Dethier, 2022; Kawamleh, 2022; Lenhard and Winsberg, 2010; Lloyd, 2010, 2015; Parker, 2006). In particular, the recognition that climate models function as mediating instruments rather than direct representations of reality has led to a richer understanding of robustness, validation, and the pluralistic nature of scientific knowledge (e.g., Edwards, 2022; Katzav, 2014; Lloyd, 2010; Oreskes, 2004; Oreskes et al., 1994; Parker,

https://doi.org/10.5194/egusphere-2026-4208
Preprint. Discussion started: 21 August 2026

2024; Winsberg, 2018). A key outcome of this philosophical engagement has been the development of concepts such as robustness analysis, model hierarchies, and ensembles as epistemic tools, which allow scientists to make reliable claims despite structural uncertainty and limited observational constraints (e.g., Parker, 2010). Similarly, debates on confirmation and underdetermination have clarified how multiple lines of evidence can jointly support causal claims such as anthropogenic climate change (Oreskes, 2004). Beyond theory, philosophy of science has also contributed to methodological standards in climate science, including transparency in model assumptions, reproducibility of simulations, calibrated language, managing the risk of "explaining away", and the role of expert judgement in uncertainty assessment (Intergovernmental Panel on Climate Change (IPCC), 2021; Parker, 2024; Shackley et al., 1998). These advances are highly relevant to PO, which faces analogous epistemic challenges: sparse and heterogeneous observations, strong nonlinearity, multi-scale interactions, and the pervasive use of numerical models as primary investigative tools. Yet, despite these similarities, explicit philosophical reflection remains relatively underdeveloped in oceanography.

Yet, we do not claim it is absent: the history and epistemology of oceanography have been examined, notably in the historiography of how observing missions and funding shaped what the field comes to know (e.g., Oreskes, 2022); but a systematic, practice-facing epistemology aimed at working physical oceanographers is still lacking. As in climate science, oceanographic knowledge is increasingly model-mediated, requiring careful consideration of how simulations, parameterisations, and data assimilation schemes contribute to inference about ocean dynamics. In this sense, PO stands to benefit from adopting the epistemological tools already developed in other fields. The novelty of this work lies not in introducing new epistemic categories, but in (i) operationalising them for PO, (ii) embedding them in a unified diagnostic framework, and (iii) applying them to concrete oceanographic practices. Doing so would not only improve methodological rigour but also strengthen the interpretability and credibility of oceanographic knowledge in interdisciplinary contexts, particularly within the broader Earth system sciences.

The paper is organised as follows. Section 2 defines the nature of knowledge in PO and introduces the epistemic foundations of the study. Section 3 develops a two-dimensional map to classify knowledge claims according to their aim and representational form and illustrates it with a classification of abstracts from the *Journal of Physical Oceanography*. Section 4 examines the epistemic status of ocean models and observational systems, with a focus on model-data interactions. Section 5 discusses scientific progress and the role of exploratory observation, culminating in an analysis of induction, uncertainty, and the limits of empirical knowledge. Section 6 provides a detailed case study of the wind-wave spectrum, highlighting its assumptions, validation issues, and epistemic limits. Section 7 introduces a diagnostic reporting template for communicating oceanographic knowledge claims. Section 8 summarises the main findings and outlines implications for future research.

2 The nature of knowledge in Physical Oceanography

The ocean's vast spatial extent, its continuous variability across a wide range of temporal scales, and the intrinsic complexity arising from turbulent, nonlinear, and tightly coupled processes give rise to a fundamentally partial and mediated form of scientific knowledge. Understanding is therefore inherently provisional, model-dependent, and constrained by what can be observed, inferred, and rigorously justified.

Within this context, knowledge in PO can be defined as the justified, theory-informed, model-mediated, and uncertainty-quantified understanding of ocean dynamics as a fluid system. The overarching goal of PO is to transform complex and often inaccessible oceanic phenomena into measurable and

https://doi.org/10.5194/egusphere-2026-4208
Preprint. Discussion started: 21 August 2026

modelable forms, enabling the representation of otherwise unobservable processes, such as deep mixing, mesoscale energetics, submesoscale fronts, and wave energy, in scientifically meaningful terms (e.g., Ferrari and Wunsch, 2009; McWilliams, 2016). Knowledge in PO can thus be viewed as emerging through multiple epistemic pathways, in which independent but complementary lines of evidence jointly support robust scientific claims (e.g., Schupbach, 2018). These pathways can be broadly categorised as follows:

- *Empirical knowledge*: derived from observations and measurements collected in situ or via remote sensing.
- *Theoretical knowledge*: grounded in first-principles physics, analytical formulations, and conceptual models that explain underlying processes.
- *Model-based knowledge*: obtained from numerical simulations and data-assimilation systems that integrate theory, observations, and parameterizations to reconstruct past states or predict future ocean states.
- *Statistical/inferential knowledge*: emerging from statistical analysis and pattern extraction, including climatologies, stochastic models, ensemble approaches, and machine learning methods.
- *Practical* (*tacit*) *knowledge* (Polanyi, 1966): arising from hands-on expertise, such as instrument deployment, data interpretation, recognition of artefacts, and the intuitive assessment of data quality.

Knowledge in PO is further shaped by its foundational role within the broader Earth-system sciences, as physical processes constrain what other ocean disciplines can observe and explain. For example, marine biology depends on physical structure: stratification regulates nutrient supply, turbulence affects encounter rates, and currents influence population connectivity (e.g., Cowen and Sponaugle, 2009; Mann and Lazier, 2006; Rothschild and Osborn, 1988). Biogeochemistry relies on physical fluxes, as ventilation, mixing, circulation, and wave-driven gas exchange control the rates and pathways of chemical transformations (e.g., Deike, 2022; Emerson and Hedges, 2010; Sarmiento and Gruber, 2006; Wanninkhof, 1992). Similarly, coastal and offshore engineering directly applies physical oceanographic knowledge of waves, currents, and extreme events for design, risk assessment, and operational decision-making (e.g., Dean and Dalrymple, 2002; Molin, 2023). At the same time, PO itself relies on external domains of knowledge to interpret and formalize its processes. It draws on fluid mechanics and thermodynamics for governing principles (e.g., Vallis, 2017); engineering for measurement systems and instrumentation (e.g., Emery and Thomson, 2004); applied mathematics and statistics for modeling and data assimilation (Evensen, 2009); atmospheric dynamics for external forcing (Holton and Hakim, 2013); and biogeochemistry for the interpretation of tracers (Sarmiento and Gruber, 2006).

In summary, knowledge in PO is best understood not as a fixed or complete representation of the ocean, but as a dynamic and evolving construct that arises from the integration of heterogeneous sources of evidence, theoretical frameworks, and methodological constraints. It is inherently interdisciplinary, relying on the continuous interplay between observation, abstraction, and modelling. As a result, oceanographic knowledge is unavoidably partial, model-mediated, and context-dependent. This pluralistic and conditional character is not a limitation but a defining strength: it enables the discipline to construct robust, testable, and progressively more informative representations of ocean dynamics, despite the system's inaccessibility, multiscale complexity, and limited observability. In this sense, knowledge advances not through convergence to a complete description, but through the coordinated refinement of multiple lines of inquiry. Recognising and explicitly articulating this epistemic structure is essential for interpreting results, assessing their scope and limitations, and ensuring their appropriate use. More

https://doi.org/10.5194/egusphere-2026-4208
Preprint. Discussion started: 21 August 2026

broadly, it supports clearer communication across disciplines and strengthens the integration of PO within the wider Earth system sciences and their societal applications.

3 Mapping knowledge gains across two epistemic axes

Knowledge gains refer to the advancement and refinement of understanding that arise from new observations, improved methodologies, and the integration of theory with empirical evidence (Bogen and Woodward, 1988; Kuhn, 1962); they capture both incremental progress, such as increased measurement accuracy or model skill, and transformative insights that reshape how systems are conceptualised. In PO, knowledge gains are particularly significant because the oceans are a complex, multiscale system where processes ranging from small-scale turbulence to basin-wide circulation interact nonlinearly. Over the decades, advances in observational technologies, numerical modelling, and theoretical frameworks have enabled an increasingly detailed decomposition of ocean dynamics into identifiable components, thereby improving our ability to attribute variability to specific forcings such as wind, atmospheric pressure, gravitation, Earth's rotation, and climate change. As a result, knowledge gains not only deepen scientific understanding but also enhance predictive capability, with direct implications for coastal risk assessment, navigation, and climate projections.

3.1 The two epistemic axes

We propose that knowledge gains in PO can be meaningfully located along two epistemic axes. These axes are not intended as a rigid classification of research outputs, nor as sequential stages in a linear trajectory of scientific maturity. Instead, they serve as a diagnostic two-dimensional map: a way to articulate the diverse modes through which PO generates understanding of oceanic phenomena despite observational limits, model idealizations, and theoretical uncertainties. The map also functions as a shared language for communicating the nature of scientific claims. The two axes are:

Axis 1 - *Aim of the knowledge* (what question is being asked): *Descriptive* ("What is happening or happened?"); *Explanatory* ("Why/how does it happen?"); *Predictive* ("What will happen?").

Axis 2 - *Form of the knowledge* (how the answer is constructed): *Statistical* (knowledge constructed through the identification of patterns and probabilistic relationships in data); *Causal* (understanding of what factors produce or change an outcome); *Mechanistic* (insight into detailed processes).

We do not claim these distinctions are new. Analogous epistemic structures recur across a wide range of scientific disciplines (Craver, 2007; Pearl, 2009; Salmon, 2005), where research is often implicitly organized along two intersecting dimensions: an axis of epistemic aim and an axis of the form of account. In medicine and epidemiology, this distinction underpins the progression from descriptive epidemiology to causal analysis and mechanistic understanding of disease (e.g., Rothman et al., 2008); in economics, it is reflected in the contrast between reduced-form (statistical) and structural (mechanistic) models (e.g., Angrist and Pischke, 2009); and in climate science, it emerges in the interplay between empirical analyses, causal attribution studies, and process-based numerical models (Intergovernmental Panel on Climate Change (IPCC) 2021). More broadly, these distinctions resonate with broader methodological divides identified in statistics and machine learning (Breiman, 2001), the formal theory of causal reasoning (Pearl, 2009), and the accounts emphasising the central role of mechanistic explanation (Craver, 2006). Our contribution is therefore not the axes themselves but their operationalisation for PO: the worked oceanographic examples below, the reading of disciplinary tendencies in Sect. 3.2, and the reporting template of Sect. 7.

We see value in adding a qualified hierarchy in PO, but it should be framed as a partial order rather than a strict ranking. The two axes are conceptually orthogonal, yet certain aims and forms

https://doi.org/10.5194/egusphere-2026-4208
Preprint. Discussion started: 21 August 2026

often precede or contain others. For instance, there is a procedural precedence across aims, since *Descriptive* → *Explanatory* → *Predictive*, where the arrows indicate common dependencies and increasing claim strength: reliable description typically lays the evidential ground for explanation, and robust explanations often improve prediction. There are exceptions; for instance, one can predict without a deep explanation. Along the form axis, a typical nesting is *Mechanistic* ⊇ *Causal* ⊇ *Statistical*, again with cavetas: a statistical pattern can be highly predictive without causal content; a causal claim can be identified without a full mechanism; a mechanistic model can be correct yet too simplified to yield system-wide predictions. The axes should not be read as strictly categorical: claims can exhibit low, moderate, or high degrees along each dimension, so the map accommodates continuous positioning and captures the hybrid nature of oceanographic knowledge. Examples of PO studies that intersect the two axes are as follows:

- *Descriptive* and *Statistical*. Wave climate summaries of, e.g., significant wave height, wave periods, and directions from altimeters or wave buoys, providing long-term distributions, trends, and extremes of wave conditions at both local and global scales.
- *Descriptive* and *Mechanistic*. Internal-wave generation over topography (e.g., from moorings, gliders, or altimetry), providing observation of phase propagation and modal structure interpreted in light of internal-wave theory .
- *Predictive* and *Statistical*. Seasonal sea-surface temperature or sea-level forecasts using empirical models, where statistical relations provide lead-time skill for regional anomalies.
- *Predictive* and *Mechanistic*. Coupled surge–wave–tide setups in complex coastal and lagoon environments, where process-based, mechanistic coupling predicts compound water levels for a better reproduction of peak timing and amplitude under multi-driver events.
- *Explanatory* and *Causal*. Sea-state-dependent air–sea drag modifying storm surge: causally linking roughness and wave age to the drag coefficient (young, steep seas increase wind stress for the same wind speed), thereby improving surge forecasts and reducing bias in peak water levels when drag is coupled to the wave model rather than kept constant.

3.2 An exploratory reading of epistemic tendencies: the *Journal of Physical Oceanography*

To illustrate how the two axes can be put to work, we apply them to a corpus of papers from the *Journal of Physical Oceanography* (*JPO*; https://journals.ametsoc.org/view/journals/phoc/phoc-overview.xml). *JPO* is widely recognised as a reference journal in PO, in terms of both scientific influence and authority. Since its establishment in 1971 by the American Meteorological Society (https://www.ametsoc.org/ams/), *JPO* has served as a primary venue for the dissemination of advances in geophysical fluid dynamics, ocean circulation, air–sea interaction, and mesoscale and submesoscale processes; its corpus reflects the theoretical and methodological foundations of the discipline. We point out that the outset that what follows is a demonstration of the coding procedure, not an empirical characterisation of the discipline: it is intended to show what the map makes visible, and to raise questions, not to settle them.

The classification is performed on article abstracts, which are treated as compact representations of scientific claims, for two benchmark years, 1995 (Volume 25; n = 207 papers) and 2025 (Volume 55; n = 140 papers). The three-decade gap allows a first look at long-term tendencies. To process the corpus at scale, we use a large language model (LLM; Microsoft Copilot, based on GPT-5) to assign each abstract an epistemic aim (*Descriptive, Explanatory, Predictive*) and a form of knowledge (*Statistical, Causal, Mechanistic*). To keep coding reproducible, we adopt the following priority rules:

https://doi.org/10.5194/egusphere-2026-4208
Preprint. Discussion started: 21 August 2026

- If a paper combines description and explanation → classified as *Explanatory*.
- If a paper uses statistical methods but seeks process understanding → classified as *Mechanistic*.
- If attribution or sensitivity experiments are present → classified as *Causal*.

These rules are themselves consequential: the second rule, in particular, moves studies off the statistical form whenever a process aim is detectable, and therefore partly shapes the marginal distribution along the form axis. Results should be read with that dependence in mind, and the rules treated as one codebook among possible alternatives.

For the years 1995 and 2025, the summary of results is presented in Tables 1 and 2, respectively. The quantitative values reported in the matrices may depend on sample size and classification choices; however, the main structural trends we identified persist across different subsets of the corpus.

Aim ↓ / Form →	*Statistical*	*Causal*	*Mechanistic*	**Total**
Descriptive	12%	5%	7%	**24%**
Explanatory	10%	22%	34%	**66%**
Predictive	2%	3%	5%	**10%**
Total	**24%**	**30%**	**46%**	

Table 1: Distribution of abstracts published in *JPO*, Volume 25 (1995), across the two epistemic axes: aim of knowledge (*Descriptive, Explanatory, Predictive*) and form of knowledge (*statistical, causal, mechanistic*). Attribution by the LLM Microsoft Copilot (GPT-5).

Aim ↓ / Form →	*Statistical*	*Causal*	*Mechanistic*	**Total**
Descriptive	25%	7%	7%	**39%**
Explanatory	14%	14%	22%	**50%**
Predictive	6%	2%	3%	**11%**
Total	**45%**	**23%**	**32%**	

Table 2: Distribution of abstracts published in *JPO*, Volume 55 (2025), across the two epistemic axes: aim of knowledge (*Descriptive, Explanatory, Predictive*) and form of knowledge (*statistical, causal, mechanistic*). Attribution by the LLM Microsoft Copilot (GPT-5).

Read as a first-order map, the comparison between 1995 and 2025 is suggestive of three tendencies, all to be treated as hypotheses for further study rather than established results:

- a rise in the share of descriptive and statistically framed work;
- a relative decline in the share of mechanistically framed work;
- the persistence of explanatory aims as the most common.

A plausible reading is that the growth of global observing systems, satellite altimetry, and multi-platform datasets has made description a primary layer of knowledge production rather than a merely preliminary one, while explanatory aims remain dominant, which is consistent with *JPO*'s orientation toward process understanding rather than operational forecasting (predictive work stays a small minority in both years). The apparent emergence of hybrid claims that combine statistical inference with descriptive reasoning is, if real, the most interesting feature; but with the present sample, it cannot be distinguished from coding noise or from drift in scientific writing style.

https://doi.org/10.5194/egusphere-2026-4208
Preprint. Discussion started: 21 August 2026

The present analysis should be understood as an epistemic experiment, rather than a definitive characterisation of the discipline. By projecting the corpus of *JPO* papers (1995 and 2025) onto two epistemic axes, we test the usefulness of a simplified conceptual framework for describing the structure of scientific activity. In particular, it relies on the analysis of abstracts as epistemic artefacts and on categorical distinctions that compress inherently hybrid and multi-layered practices. A more comprehensive understanding of the evolution of the discipline would require deeper analyses, including the examination of full papers, models, equations, and experimental practices, as well as comparisons across subfields and methodological traditions. The present experiment thus constitutes a first-order mapping of epistemic tendencies, whose main value lies in identifying trends and raising questions for further investigation, rather than providing a complete account of the dynamics of knowledge in PO.

We also note that using an LLM for abstract classification introduces potential sources of bias related to linguistic patterns, training data, and the implicit interpretation of epistemic categories. Importantly, such biases may not be invariant across time, as the style and vocabulary of scientific writing have evolved between 1995 and 2025. While applying the same LLM to both corpora ensures methodological consistency and partially mitigates systematic bias, it does not guarantee that observed differences reflect only underlying epistemic changes, rather than shifts in scientific discourse. Consequently, the results should be interpreted as capturing the evolution of expressed epistemic structure, that is, how knowledge is articulated in the literature, which is itself an integral component of scientific practice (Hacking, 1983; Kuhn, 1962; Lenhard and Winsberg, 2010).

4 The status of ocean models and model-data hybrids

Having defined how knowledge claims are classified in PO, we now examine how a primary carrier, model systems, acquire epistemic status within this space. In this field, epistemic pay-offs increasingly depend on numerical models that, although grounded in primitive physical principles and implemented through numerical schemes, are extensively mediated by parameterisations, calibration against observations, discretisation choices, data assimilation, and, more recently, data-driven components. Further, statistical emulators, machine-learning surrogates, and highly parameterised dynamical models may exhibit forms of "opacity", insofar as the internal pathways by which they generate results are not fully accessible, interpretable, or traceable to underlying physical principles (e.g., Burrell, 2016; Rudin, 2019). This underscores the hybrid and constructed character of knowledge in the field, and motivates care to avoid over-interpreting model outputs.

Following works on model-based explanation (e.g., Bokulich, 2017), we distinguish two independent dimensions of ocean-model judgment:

- *Empirical Adequacy*: a model's ability to match observations (considered the sea truth), including predictive skill, pattern replication, and performance out of sample. This dimension is necessary for any predictive use.
- *Explanatory Depth*: the extent to which a model supports causal or mechanistic interpretation. This does not simply reflect internal complexity, but the degree to which the model's structure corresponds to physically meaningful mechanisms.

These dimensions are independent. An ocean model may score high in empirical adequacy while remaining shallow in explanatory depth, as in the case of a purely data-driven prediction. Conversely, a mechanistically structured model may provide insight into causal architecture while displaying only moderate predictive performance, not merely because of resolution limits or incomplete data, but because deliberate idealisations isolate essential processes at the cost of empirical completeness. This trade-off should be treated as an epistemic choice, not a technical deficiency (Weisberg, 2013).

https://doi.org/10.5194/egusphere-2026-4208
Preprint. Discussion started: 21 August 2026

In terms of the two axes introduced in Sect. 3.1, models can occupy different positions depending on how they are used and interpreted. A machine-learning surrogate typically serves a *Predictive* aim and adopts a statistical representational form. A physics-based model used operationally may function primarily as a *Predictive* tool, even though its internal structure is *Mechanistic*. A theory-testing simulation, by contrast, aims at explanation and engages directly with the *Mechanistic* or *Causal* structure. Hybrid physics–Machine Learning systems may combine *Statistical* components with *Mechanistic* constraints, resulting in mixed representational forms. Recognizing ocean models as epistemic surrogates (i.e., tools that stand in for the ocean for specific purposes) helps prevent the conflation of predictive success with explanatory truth (Parker, 2009)

Furher, PO increasingly operates through integrated model-data systems: data assimilation frameworks, reanalysis products, coupled atmosphere–ocean configurations, and multi-source observational synthesis (e.g., Carrassi et al., 2018). These systems do not merely compare model with observation; they recursively combine observational inputs, model dynamics, statistical adjustments, and parameterized processes into unified computational architectures (e.g., Parker, 2017). While this integration enhances internal consistency and model skill, it simultaneously generates epistemic dependencies between data, models, and inferential procedures that weaken the independence of evidence and complicate the interpretation of model success, thereby requiring explicit methodological scrutiny (e.g., Winsberg, 2010).

We consider the hybrid nature of reanalysis products. Indeed, observations embedded in the model are rarely "raw givens": they are processed through inversion schemes, filtering procedures, calibration pipelines, and theoretical assumptions (theory-ladenness). The reliability of these observational inputs ultimately depends not only on formal calibration procedures but also on tacit experimental expertise embedded in measurement practices (Polanyi, 1966). Assimilation systems then incorporate these processed products into dynamical models, which themselves embed parameterizations and structural idealizations (e.g., Storto et al., 2019). Reanalysis fields are therefore hybrid constructs: neither purely empirical (the observations) nor purely theoretical (the model). For this reason, in terms of the epistemic axes defined above, a reanalysis should typically occupy a hybrid position: it serves a *Descriptive* aim (i.e., reconstructing past ocean states) but relies on a representational form that blends *Statistical* components (the assimilation schemes) with a *Mechanistic* core (the dynamical model).

An important implication is the need to shift from field-based validation to process-oriented evaluation. Agreement between reanalysis outputs and observational fields may not necessarily reflect the accuracy of the underlying physical representation, but rather the constraints imposed by data assimilation (e.g., Carrassi et al., 2018; Talagrand, 1997). As a result, empirical correspondence alone is insufficient to establish model adequacy: data assimilation can compensate for model deficiencies by continuously correcting state variables toward observations, producing empirically adequate fields even when the underlying physical representation remains incomplete or partially incorrect (Storto et al., 2019). A more robust evaluation complements traditional state-based metrics by assessing the consistency of simulated processes with theoretical and physical expectations. While reanalysis products are typically validated through agreement with observations, such agreement may be influenced by data assimilation, which constrains model states toward available measurements. A process-oriented evaluation (e.g., energy budget, momentum balance), therefore, provides an additional and partially independent line of evidence, focusing on whether the simulated ocean reproduces the governing dynamical relationships and balances, rather than only matching observed fields. Such an approach does not rely solely on agreement with assimilated data, but evaluates whether the underlying dynamical structure of the solution is physically meaningful and internally coherent. This aligns with the view that scientific understanding

https://doi.org/10.5194/egusphere-2026-4208
Preprint. Discussion started: 21 August 2026

depends on capturing relevant mechanisms, not merely reproducing observed states (e.g., Weisberg 2013; Potochnik 2017).

5 Evaluating scientific progress

Scientific progress in PO cannot be understood as a steady refinement of models or an accumulation of more accurate observations alone; it unfolds within broader research programmes that organise problems, methods, and standards of evidence. Because marine sciences are a constellation of interacting disciplines, each with its own commitments and assumptions, evaluating PO progress requires tools that capture this diversity while clarifying how physical oceanographers generate, stabilise, and transform knowledge. We use two classical accounts of scientific change (Kuhn's and Lakatos's) to characterise the dynamics of theory, modelling, observation, and anomaly-handling in PO. We use them as complementary lenses, not as independent confirmations of one another.

5.1 Kuhnian reading

Kuhn's account (Kuhn, 1962) rests on the idea that scientific knowledge develops within *paradigms* (coherent sets of exemplars, methods, instruments, and values) that structure a community's puzzle-solving activity and define what counts as a meaningful problem, a valid result, and a form of progress. A crisis arises when persistent anomalies accumulate until the existing paradigm can no longer absorb them, creating pressure for conceptual change. Kuhn stresses that accurate, reproducible measurement is central to puzzle-solving and that observations are theory-laden.

On this reading, PO is in a mature phase of normal science: a community working within shared exemplars (e.g., rotating, stratified fluid dynamics; numerical models; data-assimilation practice; observational strategies) and a disciplinary matrix of symbols, methods, tools, and values that guide puzzle-solving rather than paradigm fights. PO has passed through paradigm-like episodes; for instance, the adoption of the geostrophic balance (e.g., Sverdrup et al., 1942), the rise of numerical modelling, now coupled within Earth system models (e.g., Bryan, 1969), or the use of spectral models for wind-wave studies (Gelci et al., 1957).

The map of Sect. 3.2 is consistent with a broadly normal science: sustained puzzle-solving, increasing precision of representations, and systematic extension of established theories to new scales and regimes, with theory choice governed by shared epistemic values rather than radical discontinuity. Anomalies do emerge, for instance, in unresolved multiscale processes (e.g., submesoscale–mesoscale coupling, wave–current interactions, air–sea flux parameterisations) and in the rising prominence of data-rich observing systems, but they remain local and tractable within the existing framework. The picture is one of evolution through consolidation and hybridisation, rather than revolutionary change.

5.2 Lakatosian reading

Lakatos (Lakatos, 1978) would frame PO as a research programme with a hard core (e.g., conservation laws, rotating stratified fluid dynamics, continuum and hydrostatic approximations) and a protective belt of auxiliary hypotheses and methods (e.g., parameterisations, boundary conditions, numerics, data-assimilation and ensemble systems, physics-informed machine-learning methods), guided by a negative heuristic ("do not attack the core") and a positive heuristic ("how to extend models and tests"). One can read PO as several interacting programmes: a geophysical-fluid-dynamics core, an empirical/statistical programme, and a model–data integration programme. A programme is progressive when changes in the protective belt increase theoretical content and yield novel predictions that are subsequently corroborated (e.g., out-of-sample skill on extremes, cross-regime portability, new observable diagnostics); it is

https://doi.org/10.5194/egusphere-2026-4208
Preprint. Discussion started: 21 August 2026

degenerative when updates are largely ad-hoc patches to fit known data (e.g., hindcast tuning, site-specific knobs) without new, independently corroborated success.

On this reading, PO appears as a robust programme with an intact hard core, within which most innovation occurs in the protective belt. The core commitments remain stable. The tendencies shown in Sect. 3.2 are naturally interpreted not as challenges to the hard core but as adaptations within the belt: refinements of parameterisations, boundary conditions, and data-assimilation schemes. PO thus reads as a progressive but internally differentiated programme, in which innovation occurs mainly through the modification and extension of auxiliary structures rather than the revision of foundational principles.

What the two readings share and what they do not deserve explicit clarification. The Kuhnian and Lakatosian interpretations arrive at broadly similar conclusions, namely a stable core with innovation at the periphery, because they are applied to the same disciplinary material; their agreement should therefore not be mistaken for independent corroboration. Their value lies less in converging on a single diagnosis than in articulating different dimensions of knowledge production in PO. The Lakatosian reading can be made to do more than run in parallel with the map introduced in Sect. 3.1: the two structures are plausibly coupled. The hard core anchors the Aim axis; because these commitments remain unrevised, *Explanatory* claims retain their epistemic grounding, and the precedence *Descriptive* → *Explanatory* → *Predictive* holds as a stable dependency. The protective belt, by contrast, is where most of the movement occurs, and that movement is legible along the Form axis: refined parameterisations, assimilation schemes, and machine-learning surrogates shift a study's representational form toward the *Statistical* pole without touching the hard core's *Mechanistic* content. On this reading, the plausible rise in *Descriptive–Statistical* work and the relative decline in *Mechanistic* framing are not symptoms of a weakening core, but exactly what a progressive programme looks like when its innovation is concentrated in the belt (that's classic belt material: it's revised constantly, that's what "protective belt" means).

5.3 Beyond research programmes: the epistemic role of exploratory observation

While Kuhnian and Lakatosian frameworks capture the structured evolution of theory and models, they only partially account for phases in which the objects of inquiry themselves are not yet stabilised. In PO, these correspond to exploratory observational practices that precede and enable the formation of a research program (Feest and Steinle, 2012; Steinle, 1997). Exploratory experimentation is used to describe research that probes a domain before stable theories, or even settled concepts, are in place.

The oceans, vast, sparsely sampled, intermittently forced, and highly non-linear environment, naturally lean on this style. Historically, this has corresponded to early stages where researchers are not yet sure what the dominant scales, governing balances, or relevant descriptors of a system might be (e.g., Stewart, 2008). This emphasis on exploratory practice is consistent with the long-standing observation that technological innovation often precedes conceptual consolidation in oceanography. As Walter Munk put it in a recorded interview, "I have the basic conviction that most of the progress that's been done in oceanography was the result of applying some new technology, not because somebody had some great new ideas. A new technology almost always leads to a new understanding, and I think it's a perfectly good way to go" (Munk, 2008). This suggests that new observational capabilities frequently shape the very questions that can be asked. Observational programs often begin with broad, instrument-led surveys (e.g., hydrographic sections, glider transects, satellite-derived surface fields, moored arrays) that iterate on calibration, metadata standards, and field protocols to render phenomena tractable (see, e.g., the European https://marine.copernicus.eu/access-data/). Such descriptive campaigns help identify regimes (e.g., wind-driven vs buoyancy-driven flows), scales (e.g., synoptic vs mesoscale vs submesoscale variability), and invariants (e.g., potential vorticity structure, water-mass fingerprints, spectral slopes) before moving into explanatory or predictive phases.

https://doi.org/10.5194/egusphere-2026-4208
Preprint. Discussion started: 21 August 2026

The present day for PO is a period where observational practice, instrumentation, and nascent theory co-evolve (exploratory work in PO is often theory-informed without being theory-driven), generating new epistemic objects and establishing evidential standards appropriate to a fluid environment shaped by intermittency and multi-scale coupling. Yet the enterprise remains open-ended and receptive to the unanticipated: for example, the discovery of unforecast internal tide beams, unexpected cross-shelf exchanges during storm events, or abrupt regime shifts in the mixed layer revealed by sustained mooring records. Many such findings arise from platforms (e.g., ARGO floats, biogeochemical floats, HF radar fields, turbulence microstructure profilers) designed not only to test hypotheses but to broaden the space of observable behaviours.

5.4 Induction, uncertainty, and the limits of empirical knowledge

PO is fundamentally empirical. Climatologies, spectral descriptions, return periods of extreme events, and parameterisations used in numerical models are all constructed by generalising from observations: buoy records, satellite altimetry, radar measurements, and field campaigns. From a finite time series, scientists infer statistical regularities and project them into the future or into unobserved regions. Inductive reasoning is, therefore, unavoidable.

The problem of induction arises because there is no purely logical guarantee that patterns observed in the past will continue to hold (Hume, 1748). For instance, the fact that the significant wave height has followed a certain distribution over the last decades does not logically entail that it will do so tomorrow or under altered climatic conditions. In PO, this difficulty is exacerbated by the fact that the ocean is not a closed, stationary system: it is influenced by atmospheric forcing, boundary conditions, and long-term climate variability. Thus, inductive knowledge in PO is always provisional, resting on the assumption that the processes are sufficiently stable for past regularities to remain informative. This assumption is methodological rather than logically demonstrable. The problem becomes particularly acute in the study of extreme and rare events (Gumbel, 1958), such as rogue waves or centennial storms: by definition, these events lie in the tails of probability distributions where observational data are sparse, and inductive inference is stretched far beyond the range of direct experience.

The implication is not that inductive inference should be abandoned, but that its epistemic status must be made explicit. Many central constructs in PO are best understood as conditional summaries of past observations, whose validity depends on assumptions of stability, representativeness, and scale separation. These assumptions are rarely testable in a strict sense and are often violated in a system characterised by non-stationarity, intermittency, and regime transitions; apparent precision in statistical descriptors can therefore mask deeper epistemic fragility.

Within our framework, inductively derived knowledge typically occupies the *Descriptive–Statistical* domain but is frequently extended toward *Predictive* claims without a corresponding increase in *Explanatory* or *Mechanistic* grounding. Recognising this mismatch shifts the evaluation of oceanographic knowledge from the accuracy of numerical outputs alone to the conditions under which generalisations are warranted. Inductive reasoning remains unavoidable, but it must be accompanied by explicit statements of scope, assumptions, and limits, particularly when extrapolating beyond observed regimes.

6 Case study: the wind-wave spectrum as an epistemic construct

The purpose of this case study is not to refine theory or improve existing formulations, but to make explicit the epistemic commitments embedded in a representation that PO treats as routine. The wind-wave spectrum is a paradigmatic case: it is widely used as a primary descriptor of sea-state (i.e., the ocean

https://doi.org/10.5194/egusphere-2026-4208
Preprint. Discussion started: 21 August 2026

surface elevation field under the forcing of winds) variability and as a bridge between observations, theory, and models, yet its status as knowledge is rarely examined beyond its operational success. In this section, we analyse the wave spectrum as an epistemic construct. The goal is to clarify the conditions under which the wave spectrum constitutes reliable knowledge.

6.1 Definition

In PO, the wave spectrum can be understood as a transformation of the space-time sea-surface elevation field $\eta(x, y, t)$ into a representation that describes how variance (or energy) is distributed across frequencies and spatial scales, denoted as $S(\omega, k_x, k_y)$. The concept of the wave spectrum was introduced in wave studies to transform the intrinsically irregular, stochastic 3D sea surface (an example is given in Figure 1) into a formal and communicable representation of the physical sea state as an energy distribution, thereby enabling statistically grounded inference, intercomparison, and prediction (Gelci et al., 1957; Longuet-Higgins, 1952).

Figure 1: View of an irregular sea-surface elevation field $\eta(x, y)$ at a given time (photo credit: Alvise Benetazzo).

Without loss of generality, we consider the frequency spectrum $S(\omega)$ obtained from the analysis of the time record $\eta(t)$ measured at a fixed sea-surface location (x_0, y_0), for instance, as it is operational by traditional wave gauges. Mathematically, the wave spectrum is defined by the Fourier transform of $\eta(t)$:

$$X(\omega) = \int_{-\infty}^{\infty} \eta(t)\ e^{-i\omega t}\, \mathrm{d}t, \qquad (1)$$

or for a finite record (of duration T)

$$X_T(\omega) = \int_{-T/2}^{T/2} \eta(t)\ e^{-i\omega t}\, \mathrm{d}t, \qquad (2)$$

through which the spectral density function is given by:

$$S(\omega) = \lim_{T\to\infty} \frac{1}{2\pi T} \mid X_T(\omega) \mid^2. \qquad (3)$$

The function $S(\omega)$ assigns to each angular frequency ω (units in rad/s) the proportion of total variance (units in m^2) carried by waves at that frequency (see the example in Figure 2 for a finite-time T wave record), and it satisfies the equality:

$$\int_0^{\infty} S(\omega)\, d\omega = \sigma_\eta^2, \qquad (4)$$

meaning that its integral equals the total variance $\sigma_\eta^2 = \langle \eta^2 \rangle$ of the sea surface elevation (square brackets denote ensemble average).

https://doi.org/10.5194/egusphere-2026-4208
Preprint. Discussion started: 21 August 2026

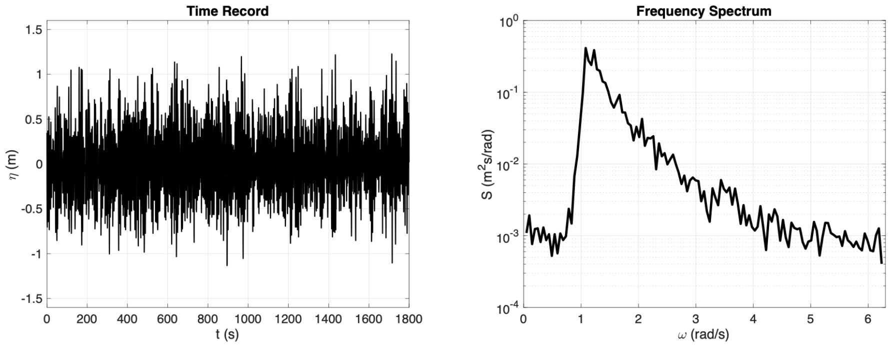

Figure 2: (left) Example of 30-minute wave time record $\eta(t)$ and (right) estimate of the variance spectrum S (logarithmic scale). For reference to the data source and analysis, the reader is referred to Benetazzo et al. (2025).

The use of the spectrum for ocean wind waves is legitimate, to leading (linear) order, because the linearised free-surface equations admit sinusoidal solutions (Airy, 1845; Stokes, 1847), and superposition makes this representation legitimate, since the sum is still a solution of the linearised free-surface equations. An irregular ocean surface can thus be represented as a superposition of harmonic components, and the spectrum is the natural way to describe how the wave variance/energy is distributed among those frequencies (e.g., Cavaleri et al., 2007; Holthuijsen, 2007; Ochi, 1998).

6.2 A view along the two axes

The wave spectrum is statistical by construction, yet PO often treats it as if it could support causal or even mechanistic explanations. Viewed along the statistical axis, the spectrum is a distribution of variance: it records how energy is arranged across frequency (and direction, in its more general form). Although constraints imposed by physical processes are indirectly reflected in its shape, it does not, by itself, encode physical mechanisms, temporal ordering, or causal relations. When the spectrum is embedded along the mechanistic axis, however, through source terms and the wave-energy balance equation (Gelci et al., 1957), it becomes linked to dynamical processes such as wind input, nonlinear transfers, and dissipation (Cavaleri et al., 2007; Komen et al., 1994). In this sense, the spectrum can participate in causal explanation, not because it is itself causal, but because a mechanistic model makes it the state variable through which causal dynamics are represented.

6.3 Does the Spectrum-Truth Exist?

In PO, the spectrum is often treated as a fundamental descriptor of variability, from turbulence to internal waves. Yet, it is not directly measured: instruments record finite time series or spatial samples, while spectra arise only through transformations and assumptions (Bendat and Piersol, 2010; Percival and Walden, 1993; Thomson and Emery, 2014). The spectrum is therefore a model-mediated inference rather than a raw observable. This raises a question central to the philosophy of science: does the spectrum reflect an intrinsic property of the ocean, or does it primarily express the analytical framework imposed upon the data (Bokulich and Parker, 2021; Hacking, 1983; Kuhn, 1962)?

The derivation of the frequency spectrum S relies on the following steps:

Measurement. A time series $\eta(t)$ or spatial maps $\eta(x, y, t)$ of sea surface elevation is obtained through buoys, radars, lidars, or optical methods. These instruments do not measure the spectrum; they provide samples of η.

Assumptions. Spectral estimation requires:

https://doi.org/10.5194/egusphere-2026-4208
Preprint. Discussion started: 21 August 2026

- *Ergodicity* (time averages equal ensemble averages). No finite realisation of sea-surface elevation can demonstrate that its time averages truly converge to ensemble averages; ergodicity is thus a methodological idealisation imposed to justify treating a single record as if it represented the whole statistical ensemble.
- *Stationarity* (statistical properties invariant over the measurement window). We can detect deviations from stationarity but cannot strictly verify it for finite records, and approximate stationarity can be assessed empirically; the assumption of stationarity is not regarded here as anything more than a pragmatic construct.
- *Linear superposition* (the sea state decomposable into independent harmonic components). This decomposition is not directly observable and rests on the linear approximation, an idealisation justified under weakly nonlinear conditions but not universally valid, which renders nonlinear, phase-coupled dynamics artificially treatable within a spectral framework.
- *Random, uniformly distributed phases*. Assuming the phases are independent and uniformly distributed on $[0,2\pi)$ makes the cross-terms in expectation $\langle \eta^2 \rangle$ vanish. It is important to separate two distinct justifications here: the partition of variance into a sum of component variances follows already from the orthogonality of distinct Fourier frequencies over a finite record and does not, by itself, require any phase assumption; the random-phase assumption is the additional hypothesis that yields Gaussianity of $\eta(t)$ and the familiar wave-height statistics. Epistemologically, the random-phase assumption plays a dual role: (i) it is a structured claim about ignorance, since we do not model phases deterministically, but claim they are effectively decorrelated; and (ii) it yields testable consequences, under linear superposition with random phases, $\eta(t)$ tends toward Gaussianity, and wave-height statistics follow known laws, which are empirically confrontable even when phase-resolved prediction is not.

Mathematical transformation. The spectrum is estimated through Fourier-based methods (Cooley and Tukey, 1965; Welch, 1967). Because the Fourier transform is formally defined for signals of infinite duration (i.e., the observation interval $T \rightarrow \infty$), any spectrum derived from a finite and discrete marine record inevitably reflects not only the ocean dynamics but also the limitations introduced by truncation, sampling, and the implicit periodic extension associated with practical Fourier methods.

Under these assumptions, the estimated spectrum is a conditional, model-dependent representation rather than a directly observable quantity. This does not render the target unreal: the second-order statistical structure that the spectrum encodes is a genuine property of a (statistically idealised) sea state, and the random-wave model yields empirically testable predictions. What is model-mediated is the estimator, not the existence of an underlying second-order structure. Accordingly, the routine phrase "observed wave spectrum" (Collins et al., 2025) is more accurately rendered as “wave spectrum inferred from observations”: the change is terminological, but disciplines the claim being made.

6.4 What the spectrum captures and what it omits

Despite its inferential nature, the wave spectrum $S(\omega)$ contains a concentration of information about the sea state: statistically persistent scales (e.g., Ochi, 1998); characteristic time scales, for instance, such as peak period or spectral width (e.g., Young, 1999); wind-wave growth stage (young, developing, or fully developed) and fetch dependence (Hasselmann et al., 1973); presence of multiple wave systems (swell–wind sea superposition; e.g., Portilla et al., 2009); evidence of nonlinear energy transfer (e.g., Komen et al., 1994). When the full space-time sea-surface elevation field $\eta(x, y, t)$ is considered, the statistical

https://doi.org/10.5194/egusphere-2026-4208
Preprint. Discussion started: 21 August 2026

spatio-temporal energy distribution across frequencies and wavenumbers is similarly informative (e.g., Leckler et al., 2015).

Epistemologically, the spectrum embodies the belief, physically grounded in linear wave theory, that much of the dynamically significant information in the sea state is captured by the distribution of energy across harmonic modes. It therefore functions as a compressed representation of ocean surface variability. However, this compression is selective: it prioritises quantities that are predictable under linear superposition and treats others as negligible or inaccessible. Consequently, the wave spectrum $S(\omega)$ necessarily omits several aspects of real wave fields:

- *Local phase information*. The Fourier wave spectrum contains only variance and autocorrelation (second-order statistics), so it cannot describe the specific arrangement of waves in time, wave groups, or coherent structures (e.g., Longuet-Higgins, 1984).
- *Strong wave–wave nonlinear interactions*. Phenomena such as transfer of energy between waves, modulational instability, and bound harmonic generation arise from nonlinear interactions among wave components that depend on phase coherence and higher-order statistical properties (e.g., Maestrini et al., 2026). The wave spectrum cannot represent phase correlations or higher-order moments and therefore cannot describe or diagnose these nonlinear processes.
- *Wave breaking.* Wave breaking is a strongly nonlinear and intermittent process tied to local crest instability, turbulence, and dissipation. As it depends on phase-resolved dynamics, wave grouping, and local steepness exceeding a threshold, it cannot be determined from the wave spectrum alone (e.g., Malila et al., 2022). While spectral properties may provide statistical estimates of breaking probability, the occurrence and structure of individual breaking events are not directly inferable from the spectral shape.

The implication is that the wave spectrum is not a complete description but a deliberate abstraction, designed around what is most stably measurable and predictively useful under linear theory; knowledge extracted from it is therefore partial and theory-dependent. Contemporary spectral-wave models built on the Fourier spectrum are a pragmatic compromise: they retain sufficient physical grounding and computational tractability while providing a useful, if inherently selective, description of wave dynamics.

Alternative representations are designed to capture physical features that the standard spectral framework tends to suppress or average out. For instance, time–frequency methods, such as wavelets (Daubechies, 1990), relax the assumption of stationarity and reveal the temporal evolution of sea states and transient wave groups. Higher-order measures, such as bispectra, capture phase relationships and nonlinear interactions invisible to variance-based spectra (Kim and Powers, 1979). Fully adaptive methods such as empirical mode decomposition (Huang et al., 1998) relax both stationarity and linearity by decomposing the signal into intrinsic mode functions based on local time scales. These approaches do not invalidate the common spectral framework; rather, they highlight that it is a selective projection of the sea state, privileging certain aspects of variability while neglecting others. From an epistemological standpoint, the choice of representation determines which features of ocean dynamics become visible, measurable, and ultimately knowable.

6.5 Dependence and partial circularity in wave-model validation

A central issue arises when, as a common practice, ocean observations and spectral representations are used as the basis for wave-model validation (Komen et al., 1994); indeed, the comparison between model output and observational estimates is not fully independent. While both aim to describe the same physical object (i.e., the ocean surface elevation spatio-temporal field), they rely on a shared representational

https://doi.org/10.5194/egusphere-2026-4208
Preprint. Discussion started: 21 August 2026

framework rooted in common idealisations. We have highlighted that observational spectra are not direct measurements but are inferred from finite records through transformations that already embed these assumptions. Similarly, spectral-wave models are formulated in terms of energy distributions over frequency space, adopting the same conceptual structure. As a result, the agreement between modelled and observed spectra (or their integral parameters) does provide evidence of consistency and model skill, but primarily within the confines of this shared framework. It does not, by itself, constitute an independent test of the validity of the underlying assumptions that define the representation. This introduces a form of partial circularity: model validation assesses the coherence of two implementations of the same conceptual lens, rather than directly confronting that lens with assumption-free observations. The epistemic risk is not that model validation becomes meaningless, but that it may overstate the extent to which spectral agreement confirms the underlying physical description of the ocean.

7 A diagnostic reporting template for knowledge claims

Operationalising the perspectives described above allows a physical oceanographer to ask a more demanding question than whether a numerical model's correlation coefficient or bias improved. The relevant issue becomes whether a framework exhibits heuristic power: whether it guides inquiry toward new explanatory ocean structures and previously unobserved regimes, rather than merely optimising agreement within known data. Progress, in this view, is measured not only by predictive accuracy but by the capacity of a framework to enlarge the space of testable possibilities.

To make this assessment communicable, we propose an epistemic reporting protocol that classifies a PO study along the two epistemic axes and lists its primary validation criterion and payoffs. The diagnostic tool can be used to classify a research study to tell what kind of knowledge it produces (*Descriptive*, *Predictive*, *Explanatory*; *Statistical*, *Causal*, *Mechanistic*) and where it sits in the conceptual space of PO knowledge. The matrix is not a classificatory grid with rigid boundaries, but a heuristic device. Its purpose is to clarify the nature of the claim being made and the type of justification supporting it. In practice, most PO studies occupy intermediate or hybrid positions, whose precise location depends on how results are constructed, interpreted, and applied. The aim of the framework is thus not to police scientific language, but rather to discipline epistemic claims by rendering their underlying structure explicit. A possible structure is (multiple selections allowed per block):

- *Methods*: Observations; Numerical Models; Data Assimilation; Machine Learning; Laboratory Experiments; Other.
- *Epistemic Aim*: Descriptive (low/moderate/high); Predictive (low/moderate/high); Explanatory (low/moderate/high).
- *Structural Form*: Statistical (low/moderate/high); Causal (low/moderate/high); Mechanistic (low/moderate/high).
- *Payoff / Outcome*: New descriptive regularity; Improved prediction skill; Identified causal pathway; Mechanistic understanding; Anomaly discovery; Generalizability/scale-awareness; Decision-support relevance; Tacit/procedural expertise codified; Other.
- *Fit-for-Purpose (Adequacy-for-Purpose) Assessment*: Operational forecasting; Long-term projection; Mechanistic inference; Theory development; Policy/decision support; Exploratory science; Other.
- *Validation Criteria*: Consistency (multi-sensor agreement, reprocessing stability); Out-of-sample predictive tests; Process/energy/budget closure; Counterfactual model experiments; Cross-regime or cross-region transferability; External/independent datasets; Internal vs. external model consistency; Other.

https://doi.org/10.5194/egusphere-2026-4208
Preprint. Discussion started: 21 August 2026

- *Limits*: Sampling limitation (finite, sparse, uneven data); Representation dependence (results depend on chosen framework, e.g. spectrum); Assumption sensitivity (stationarity, linearity, ergodicity); Model structural uncertainty (parameterisation, closure choices); Scale limitation (lack of cross-scale validity); Nonstationarity (changing forcing / climate regime); Extrapolation risk (beyond observed range, especially extremes); Methodological dependence (filtering, processing choices); Mechanistic incompleteness (no causal/process resolution); Interpretability limitation (opaque or hybrid models); Context dependence (valid only for specific regions/regimes); Other.

7.1 Worked examples

Placed side by side, the two fillings shown here occupy nearly opposite regions of the map, which is exactly the discrimination the template is meant to make legible to a reader, a reviewer, or a decision-maker. We propose that such a diagnosis be appended to scientific papers to communicate the epistemic status of their results, the assumptions on which they rely, and the scope within which their claims can be considered reliable, by promoting transparency, comparability, and a more critical evaluation of knowledge production in PO. It is in this communicative function that the framework is intended to be useful.

Maestrini, D., Dematteis, G., Benetazzo, A.: *Revealing wave-wave resonant interactions in ocean wind waves. Communication Physics* 9, 193 (2026). (Maestrini et al., 2026).	**Moldovan, G., Pinnington, E., Prieto Nemesio, A., Lang, S., et al.: *AIFS Single 1.1.0: an update to ECMWF's machine-learned weather forecast model AIFS. Geoscience Model Development* 19, 4703–4724 (2026). (Moldovan et al., 2026).**
Methods: Observations (primary: in situ / stereoscopic measurements); Data processing and spectral analysis techniques.	*Methods*: Machine Learning (data-driven model); Observations (training and validation data); Numerical models; Statistical learning and optimisation techniques.
Epistemic Aim: Explanatory (high): mechanism governing nonlinear energy transfer across scales; Descriptive (moderate); Predictive (low).	*Epistemic Aim*: Predictive (high): accurate mapping between inputs and outputs; Descriptive (moderate); Explanatory (low).
Structural Form: Mechanistic (high): identification of nonlinear interaction processes; Statistical (moderate); Causal (moderate).	*Structural Form*: Statistical (high): learned functional relationships; Causal (low); Mechanistic (low–moderate).
Payoff / Outcome: Mechanistic understanding; Identification of a causal pathway for energy transfer; Resolution of a long-standing theoretical problem	*Payoff / Outcome*: Improved prediction skill (primary); New descriptive regularities; Possible anomaly detection; Tacit knowledge embedded in trained model parameters.
Fit for Purpose: Mechanistic inference (primary); Theory development.	*Fit for Purpose*: Operational forecasting (strong, if generalisation holds); Exploratory data-driven science; Decision support.
Validation Criteria: Process/energy consistency; Confrontation with independent observational evidence.	*Validation Criteria*: Out-of-sample predictive tests; External/independent datasets; Consistency across datasets; Cross-regime generalisation.
Limits: Sampling limitation (finite, context-specific observations); Representation dependence (spectral framework); Assumption sensitivity (e.g. weak nonlinearity, resonance conditions); Scale limitation (restricted range of resolved scales); Methodological dependence (reconstruction and processing techniques);	*Limits*: Sampling limitation (training data constraints); Representation dependence (model architecture, feature space); Assumption sensitivity (training distribution, stationarity); Generalisation failure (out-of-distribution / regime shifts); Extrapolation risk (extremes, climate change); Methodological dependence (training procedure,);

https://doi.org/10.5194/egusphere-2026-4208
Preprint. Discussion started: 21 August 2026

Context dependence (specific sea states and forcing conditions); Mechanistic incompleteness (focus on a subset of processes).	Interpretability limitation (opaque models); Weak causal identifiability; Context dependence (valid only within training domain).

Table 3: Worked examples of the proposed epistemic reporting template. The table compares a mechanistic/explanatory oceanographic study and a statistical/predictive machine-learning forecasting study, illustrating how different research approaches occupy distinct positions within the proposed epistemic framework.

8 Conclusions

Physical Oceanography operates under intrinsic epistemic constraints: sparse and heterogeneous observations, strong nonlinearity, multiscale interactions, and pervasive model mediation. In this context, knowledge cannot be understood as a direct representation of the ocean, but as a constructed, conditional, and evolving synthesis of measurements, theory, and numerical modelling. We have argued that making these conditions explicit is necessary for clarifying what oceanographic knowledge means, how it is justified, and where its limits lie.

We proposed a two-dimensional map to characterise knowledge in Physical Oceanography along the axes of epistemic aim (*Descriptive, Explanatory, Predictive*) and representational form (*Statistical, Causal, Mechanistic*). The map situates results within a structured epistemic space, rather than treating all forms of knowledge as equivalent. Our exploratory reading of *Journal of Physical Oceanography* abstracts (1995 vs 2025) is consistent with a field whose aims remain predominantly explanatory while its representational form has diversified toward statistical and hybrid approaches, with a relative decline in the share of mechanistically framed work in the sampled corpus. We stress that this reading is illustrative, not a measured trend: it is offered as a hypothesis-generating use of the map, subject to the validation caveats. We further examined the epistemic status of models and observational systems, highlighting their interdependence and the role of assumptions in shaping both data and interpretation. The case study on wind-wave spectral representation illustrated how widely used constructs embed specific idealisations and may introduce partial epistemic circularity in validation, insofar as models and observations are compared within a shared representational framework. This does not invalidate such comparisons, but it constrains their interpretive power: agreement demonstrates consistency within a framework, rather than independent confirmation of its assumptions.

Against this background, we introduced a diagnostic reporting template designed to make explicit the epistemic structure of individual studies. By systematically identifying methods, aims, representational forms, epistemic payoffs, fit for purpose, validation criteria, and limits, the template shifts the evaluation of scientific work from a narrow focus on performance metrics toward a broader assessment of heuristic power, scope, and robustness. In this view, scientific progress in Physical Oceanography is not measured only by improved prediction or reduced error, but by the ability of a framework to generate new testable hypotheses, reveal previously unobserved regimes, and deepen mechanistic understanding. The implications of this approach are both methodological and practical. Explicitly stating assumptions, limits, and the scope of applicability can improve transparency, facilitate comparison between studies, and reduce the risk of overinterpretation, particularly in contexts involving extrapolation or decision support. More broadly, recognising the inductive and model-dependent nature of oceanographic knowledge provides a more realistic and robust foundation for communicating uncertainty and assessing the reliability of scientific claims.

Finally, while developed here for Physical Oceanography, the framework may be transferable to other domains within the marine sciences, where similar challenges arise from complexity, limited observability, and model-mediated inference. By treating epistemology not as an abstract reflection but as an operational component of scientific practice, this work aims to contribute to a more explicit, disciplined, and practically useful understanding of knowledge production in a complex environmental system.

Data and Code Availability

The study does not involve the generation of original datasets or research software. The work is based on published literature and on a corpus of *Journal of Physical Oceanography* abstracts used for the exploratory classification presented in Sect. 3.2. The classification was performed using Microsoft Copilot (GPT-5) following the decision rules described in the manuscript. Because no dedicated software package or reusable code was developed for this analysis, no code repository is available. Information regarding the corpus and classification procedure can be obtained from the corresponding author upon reasonable request.

Author contributions

AB: Conceptualization, Investigation, Formal analysis, Methodology, Writing original draft; TM: Investigation, Formal analysis, Methodology, Writing original draft.

Competing interests

The authors declare that they have no conflict of interest.

Ethical statement

This study is a theoretical and methodological investigation in the field of Physical Oceanography and the philosophy of science. The research did not involve human participants, animals, biological materials, personal data, interviews, surveys, experiments requiring ethical oversight, or any other activities subject to formal ethical review. Consequently, approval from an institutional ethics committee was not required under the regulations of the authors' institutions.

Acknowledgments

This work has benefited from ongoing interactions and discussions with numerous colleagues in the physical oceanography community, whose insights and perspectives on Physical Oceanography knowledge and its meaning have contributed to shaping the ideas developed in this paper. The authors gratefully acknowledge Luigi Cavaleri for insightful discussions, which have significantly contributed to the conceptualization of this work. The authors thank Øyvind Breivik for carefully reading the manuscript and for comments that contributed to clarifying and strengthening the arguments developed in this work. Alessio Langiu is thanked for LLM insights and for providing the *Journal of Physical Oceanography* database used for the analysis in Sect. 3.2. The authors acknowledge the use of Microsoft Copilot (GPT-5) for assistance in improving the clarity of English and refining the expression of some arguments. All scientific content, interpretations, and conclusions remain the direct responsibility of the authors.

Financial support

The authors declare that no specific funding was received for this work from any funding agency in the public, commercial, or non-profit sectors.

https://doi.org/10.5194/egusphere-2026-4208
Preprint. Discussion started: 21 August 2026

References

Airy, G. B.: Tides and Waves, in: Encyclopedia Metropolitana, 241–396, 1845.

Angrist, J. David. and Pischke, J.-Steffen.: Mostly harmless econometrics : an empiricist's companion, Princeton University Press, 373 pp., 2009.

Audi, Robert.: Epistemology, Taylor & Francis, 352 pp., 1997.

Bendat, J. S. and Piersol, A. G. .: Random data : analysis and measurement procedures, Wiley, 604 pp., 2010.

Benetazzo, A., Bergamasco, F., Barbariol, F., Ferla, M., Nardone, G., Orasi, A., Picone, M., Stevanin, A., Marchitiello, A., Pavan, T., Bastianini, M., Pistellato, M., and Cavaleri, L.: Evaluation of the 3D response and performance of an operational wave buoy for coastal wave monitoring, Coastal Engineering, 200, 104756, https://doi.org/10.1016/j.coastaleng.2025.104756, 2025.

Bigg, G. R.: The oceans and climate, Cambridge University Press, 266 pp., 1996.

Bogen, J. and Woodward, J.: Saving the Phenomena, Philos. Rev., 97, 303, https://doi.org/10.2307/2185445, 1988.

Bokulich, A.: Models and Explanation, in: Springer Handbook of Model-Based Science, Springer International Publishing, Cham, 103–118, https://doi.org/10.1007/978-3-319-30526-4_4, 2017.

Bokulich, A. and Parker, W.: Data models, representation and adequacy-for-purpose, Eur. J. Philos. Sci., 11, 31, https://doi.org/10.1007/s13194-020-00345-2, 2021.

BonJour, Laurence.: Epistemology : classic problems and contemporary responses, Rowman & Littlefield Publishers, 2010.

Breiman, L.: Statistical Modeling: The Two Cultures, Statistical Science, 16, https://doi.org/10.1214/ss/1009213726, 2001.

Bryan, K.: A numerical method for the study of the circulation of the world ocean, J. Comput. Phys., 4, 347–376, https://doi.org/10.1016/0021-9991(69)90004-7, 1969.

Burrell, J.: How the machine 'thinks': Understanding opacity in machine learning algorithms, Big Data Soc., 3, https://doi.org/10.1177/2053951715622512, 2016.

Carrassi, A., Bocquet, M., Bertino, L., and Evensen, G.: Data assimilation in the geosciences: An overview of methods, issues, and perspectives, WIREs Climate Change, 9, https://doi.org/10.1002/wcc.535, 2018.

Cavaleri, L., Alves, J. H. G. M., Ardhuin, F., Babanin, A., Banner, M., Belibassakis, K., Benoit, M., Donelan, M., Groeneweg, J., Herbers, T. H. C., Hwang, P., Janssen, P. A. E. M., Janssen, T., Lavrenov, I. V., Magne, R., Monbaliu, J., Onorato, M., Polnikov, V., Resio, D., Rogers, W. E., Sheremet, A., McKee Smith, J., Tolman, H. L., van Vledder, G., Wolf, J., and Young, I.: Wave modelling - The state of the art, https://doi.org/10.1016/j.pocean.2007.05.005, 2007.

Collins, C. O., Amador, A., Babanin, A., Behrens, J., Benetazzo, A., Bergamasco, F., Blenkinsopp, C., Bonneton, P., Breivik, Ø., Christensen, K. H., Colosi, L. V., Ewans, K., Gemmrich, J., Glover, H. E., Grare, L., Grigorieva, V., Gulev, S. K., Hauser, D., Hole, L. R., Hope, G., Houghton, I., Hsu, J.-Y., Laxague, N. J. M., Lenain, L., Lund, B., O'Dea, A., Pistellato, M., Magnusson, A. K., Martins, K., Matsuba, Y., McAllister, M., Müller, M., Passaro, M., Rabault, J., Roarty, H., Smit, P. B., Smith, M. M., Tamura, H., Tilinina, N., Timmermans, B., Thomson, J., Voermans, J. J., Wengrove, M. E., Wyatt, L., Yoo, J., Young, I., Zappa, C. J., and Zhang, D.: Measuring Ocean Surface Waves, https://doi.org/10.22541/essoar.175130482.28547248/v1, 2025.

Cooley, J. W. and Tukey, J. W.: An algorithm for the machine calculation of complex Fourier series, Math. Comput., 19, 297–301, https://doi.org/10.1090/S0025-5718-1965-0178586-1, 1965.

Cowen, R. K. and Sponaugle, S.: Larval Dispersal and Marine Population Connectivity, Ann. Rev. Mar. Sci., 1, 443–466, https://doi.org/10.1146/annurev.marine.010908.163757, 2009.

Craver, C. F.: When mechanistic models explain, Synthese, 153, 355–376, https://doi.org/10.1007/s11229-006-9097-x, 2006.

Craver, C. F. .: Explaining the brain : mechanisms and the mosaic unity of neuroscience, Oxford University Press, 308 pp., 2007.

https://doi.org/10.5194/egusphere-2026-4208
Preprint. Discussion started: 21 August 2026

Cushman-Roisin, Benoit. and Beckers, J.-Marie.: Introduction to geophysical fluid dynamics: physical and numerical aspects, Academic Press, 828 pp., 2011.
Daubechies, I.: The wavelet transform, time-frequency localization and signal analysis, IEEE Trans. Inf. Theory, 36, 961–1005, https://doi.org/10.1109/18.57199, 1990.
Dean, R. G. . and Dalrymple, R. A. .: Coastal processes : with engineering applications, Cambridge University Press, 475 pp., 2002.
Deike, L.: Mass Transfer at the Ocean–Atmosphere Interface: The Role of Wave Breaking, Droplets, and Bubbles, Annu. Rev. Fluid Mech., 54, 191–224, https://doi.org/10.1146/annurev-fluid-030121-014132, 2022.
Dethier, C.: When is an ensemble like a sample? “Model-based” inferences in climate modeling, Synthese, 200, 52, https://doi.org/10.1007/s11229-022-03477-5, 2022.
Douglas, H. E.: Science, Policy, and the Value-Free Ideal, University of Pittsburgh Press, https://doi.org/10.2307/j.ctt6wrc78, 2009.
Edwards, P. N.: A Vast Machine: Computer Models, Climate Data, and the Politics of Global Warming, MIT Press Ltd, 2022.
Emerson, Steven. and Hedges, John.: Chemical oceanography and the marine carbon cycle, Cambridge University Press, 453 pp., 2010.
Emery, W. J. . and Thomson, R. E. .: Data analysis methods in physical oceanography, Elsevier, 638 pp., 2004.
Evensen, Geir.: Data assimilation : the ensemble Kalman filter, Springer, 307 pp., 2009.
Feest, Uljana. and Steinle, Friedrich.: Scientific concepts and investigative practice, De Gruyter, 301 pp., 2012.
Ferrari, R. and Wunsch, C.: Ocean Circulation Kinetic Energy: Reservoirs, Sources, and Sinks, Annu. Rev. Fluid Mech., 41, 253–282, https://doi.org/10.1146/annurev.fluid.40.111406.102139, 2009.
Frigg, R. and Hartmann, S.: Models in Science, in: The Stanford Encyclopedia of Philosophy , edited by: Edward N. Zalta & Uri Nodelman (eds.), https://plato.stanford.edu/archives/win2025/entries/models-science/, 2025.
Gelci, R., Cazale, H., and Vassal, J.: Prevision de la houle. La methode des desites spectroangulaires, Bull. Inf . Comite Central Oceanogr. Etudes Cotes, 9, 416–435, 1957.
Gill, A. E.: Atmosphere-ocean dynamics, Academic Press, 662 pp., 1982.
Griffies, S.: Fundamentals of ocean climate models, Princeton University Press, 2004.
Gumbel, E. Julius.: Statistics of extremes, edited by: Columbia Univesrity Press, Dover Publications, New York, 1958.
Hacking, I.: Representing and Intervening, Cambridge University Press, https://doi.org/10.1017/CBO9780511814563, 1983.
Hasselmann, K., Barnett, T. P., Bouws, E., Carlson, H., Cartwright, D. E., Enke, K., Ewing, J. A., Gienapp, H., Hasselmann, D. E., Kruseman, P., Meerburg, A., Müller, P., Olbers, D. J., Richter, K., Sell, W., and Walden, H.: Measurements of wind-wave growth and swell decay during the Joint North Sea Wave Project (JONSWAP), Deutches Hydrographisches Institut, A8, 1–95, 1973.
Holthuijsen, L. H.: Waves in Oceanic and Coastal Waters, Cambridge University Press, 387 pp., 387 pp., 2007.
Holton, J. R. and Hakim, G. J.: An Introduction to Dynamic Meteorology, Elsevier, https://doi.org/10.1016/C2009-0-63394-8, 2013.
Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N.-C., Tung, C. C., and Liu, H. H.: The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, 454, 903–995, https://doi.org/10.1098/rspa.1998.0193, 1998.
Hume, D.: An Enquiry concerning Human Understanding, edited by: A Millan, Oxford University Press, London, 134–198, https://doi.org/10.1093/oseo/instance.00032980, 1748.
Intergovernmental Panel on Climate Change (IPCC): Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change[Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K.

https://doi.org/10.5194/egusphere-2026-4208
Preprint. Discussion started: 21 August 2026

Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)], edited by: Cambridge University Press, C. U. K. and N. Y. N. U., Cambridge University Press, 2391 pp., https://doi.org/10.1017/9781009157896, 2021.

Janssen, Peter.: The interaction of ocean waves and wind, Cambridge University Press, 300 pp., 2009.

Karnauskas, Kris.: Physical oceanography and climate, Cambridge University Press, 247 pp., 2020.

Katzav, J.: The epistemology of climate models and some of its implications for climate science and the philosophy of science, Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics, 46, 228–238, https://doi.org/10.1016/j.shpsb.2014.03.001, 2014.

Kawamleh, S.: Confirming (climate) change: a dynamical account of model evaluation, Synthese, 200, 122, https://doi.org/10.1007/s11229-022-03659-1, 2022.

Kim, Y. C. and Powers, E. J.: Digital Bispectral Analysis and Its Applications to Nonlinear Wave Interactions, IEEE Transactions on Plasma Science, 7, 120–131, https://doi.org/10.1109/TPS.1979.4317207, 1979.

Komen, G. J., Cavaleri, L., M. Donelan, K. H., Hasselmann, S., Janssen, P. A. E. M., Donelan, M., Hasselmann, K., Hasselmann, S., and Janssen, P. A. E. M.: Dynamics and Modelling of Ocean Waves., edited by: Cambridge University Press, Cambridge University Press, Cambridge, 532 pp., https://doi.org/10.1017/CBO9780511628955, 1994.

Kuhn, T. S.: The Structure of Scientific Revolutions., edited by: University of Chicago Press, Chicago, 1962.

Lakatos, I.: The Methodology of Scientific Research Programmes, edited by: Worrall, J. and Currie, G., Cambridge University Press, https://doi.org/10.1017/CBO9780511621123, 1978.

Leckler, F., Ardhuin, F., Peureux, C., Benetazzo, A., Bergamasco, F., and Dulov, V.: Analysis and interpretation of frequency-wavenumber spectra of young wind waves, J. Phys. Oceanogr., 45, 2484–2496, https://doi.org/10.1175/JPO-D-14-0237.1, 2015.

Lenhard, J. and Winsberg, E.: Holism, entrenchment, and the future of climate model pluralism, Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics, 41, 253–262, https://doi.org/10.1016/j.shpsb.2010.07.001, 2010.

Lloyd, E. A.: Confirmation and Robustness of Climate Models, Philos. Sci., 77, 971–984, https://doi.org/10.1086/657427, 2010.

Lloyd, E. A.: Model robustness as a confirmatory virtue: The case of climate science, Studies in History and Philosophy of Science Part A, 49, 58–68, https://doi.org/10.1016/j.shpsa.2014.12.002, 2015.

Longuet-Higgins, M. S.: On the statistical distribution of the heights of sea waves, J. Mar. Res., 11, 245–265, 1952.

Longuet-Higgins, M. S.: Statistical properties of wave groups in a random sea state, Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences, 312, 219–250, https://doi.org/10.1098/rsta.1984.0061, 1984.

Maestrini, D., Dematteis, G., Benetazzo, A., and Onorato, M.: Revealing wave-wave resonant interactions in ocean wind waves, Commun. Phys., 9, 193, https://doi.org/10.1038/s42005-026-02653-0, 2026.

Malila, M. P., Thomson, J., Breivik, Ø., Benetazzo, A., Scanlon, B., and Ward, B.: On the Groupiness and Intermittency of Oceanic Whitecaps, J. Geophys. Res. Oceans, 127, https://doi.org/10.1029/2021JC017938, 2022.

Mann, K. H. and Lazier, J. R. N.: Dynamics of marine ecosystems: biological-physical interactions in the oceans, Blackwell Pub., 496 pp., 2006.

McPhee, M.: Air-Ice-Ocean Interaction, Springer New York, New York, NY, https://doi.org/10.1007/978-0-387-78335-2, 2008.

McWilliams, J. C.: Submesoscale currents in the ocean, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 472, 20160117, https://doi.org/10.1098/rspa.2016.0117, 2016.

Moldovan, G., Pinnington, E., Prieto Nemesio, A., Lang, S., Ben Bouallègue, Z., Dramsch, J., Alexe, M., Santa Cruz, M., Hahner, S., Cook, H., Theissen, H., Clare, M., O'Brien, C., Polster, J., Magnusson, L., Mertes, G., Pinault, F., Raoult, B., de Rosnay, P., Forbes, R., and Chantry, M.: AIFS Single 1.1.0: an

https://doi.org/10.5194/egusphere-2026-4208
Preprint. Discussion started: 21 August 2026

update to ECMWF's machine-learned weather forecast model AIFS, Geosci. Model Dev., 19, 4703–4724, https://doi.org/10.5194/gmd-19-4703-2026, 2026.
Molin, Bernard.: Offshore structure hydrodynamics, Cambridge University Press, 426 pp., 2023.
Munk, W.: One Man's Noise: Stories of An Adventuresome Oceanographer, https://www.youtube.com/watch?v=je3QvqNdHl0, 2008.
Ochi, M. K.: Ocean Waves, Cambridge University Press, 319 pp., https://doi.org/10.1017/CBO9780511529559, 1998.
Oreskes, N.: The Scientific Consensus on Climate Change, Science (1979)., 306, 1686–1686, https://doi.org/10.1126/science.1103618, 2004.
Oreskes, N., Shrader-Frechette, K., and Belitz, K.: Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences, Science (1979)., 263, 641–646, https://doi.org/10.1126/science.263.5147.641, 1994.
Oreskes, Naomi.: Science on a mission : how military funding shaped what we do and don't know about the ocean, The University of Chicago Press, 738 pp., 2022.
Parker, W. S.: Understanding Pluralism in Climate Modeling, Found. Sci., 11, 349–368, https://doi.org/10.1007/s10699-005-3196-x, 2006.
Parker, W. S.: II - Confirmation and adequacy-for-Purpose in Climate Modelling, Aristotelian Society Supplementary Volume, 83, 233–249, https://doi.org/10.1111/j.1467-8349.2009.00180.x, 2009.
Parker, W. S.: Predicting weather and climate: Uncertainty, ensembles and probability, Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics, 41, 263–272, https://doi.org/10.1016/j.shpsb.2010.07.006, 2010.
Parker, W. S.: Computer Simulation, Measurement, and Data Assimilation, Br. J. Philos. Sci., 68, 273–304, https://doi.org/10.1093/bjps/axv037, 2017.
Parker, W. S.: Climate Science, Cambridge University Press, https://doi.org/10.1017/9781009619301, 2024.
Pearl, J.: Causality: models, reasoning, and inference, Cambridge University Press, 465 pp., 2009.
Pedlosky, Joseph.: Geophysical Fluid Dynamics, Springer, 723 pp., 2013.
Percival, D. B. and Walden, A. T.: Spectral Analysis for Physical Applications, Cambridge University Press, https://doi.org/10.1017/CBO9780511622762, 1993.
Pereira, Leonel. and Pardal, Miguel.: Oceanography : relationships of the oceans with the continents, their biodiversity and the atmosphere, IntechOpen, 140 pp., 2024.
Polanyi, M.: The Logic of Tacit Inference, Philosophy, 41, 1–18, 1966.
Popper, Karl.: The Logic of Scientific Discovery, edited by: Basic Books., New York, NY, 1959.
Portilla, J., Ocampo-Torres, F. J., and Monbaliu, J.: Spectral Partitioning and Identification of Wind Sea and Swell, J. Atmos. Ocean. Technol., 26, 107–122, https://doi.org/10.1175/2008JTECHO609.1, 2009.
Potochnik, Angela.: Idealization and the aims of science, University of Chicago Press, 2017.
Pugh, David. and Woodworth, Philip.: Sea-level science : understanding tides, surges, tsunamis and mean sea-level changes, Cambridge University Press, 395 pp., 2014.
Roemmich, D., Johnson, G., Riser, S., Davis, R., Gilson, J., Owens, W. B., Garzoli, S., Schmid, C., and Ignaszewski, M.: The Argo Program: Observing the Global Oceans with Profiling Floats, Oceanography, 22, 34–43, https://doi.org/10.5670/oceanog.2009.36, 2009.
Rothman, K. J., Greenland, Sander., and Lash, T. L.: Modern epidemiology, Wolters Kluwer Health/Lippincott Williams & Wilkins, 758 pp., 2008.
Rothschild, B. J. and Osborn, T. R.: Small-scale turbulence and plankton contact rates, J. Plankton Res., 10, 465–474, https://doi.org/10.1093/plankt/10.3.465, 1988.
Rudin, C.: Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead, Nat. Mach. Intell., 1, 206–215, https://doi.org/10.1038/s42256-019-0048-x, 2019.
Salmon, W. C. .: Scientific Explanation and the Causal Structure of the World, Princeton University Press, 2005.
Sarmiento, J. Louis. and Gruber, Nicolas.: Ocean biogeochemical dynamics, Princeton University Press, 503 pp., 2006.

https://doi.org/10.5194/egusphere-2026-4208
Preprint. Discussion started: 21 August 2026

Schupbach, J. N.: Robustness Analysis as Explanatory Reasoning, Br. J. Philos. Sci., 69, 275–300, https://doi.org/10.1093/bjps/axw008, 2018.

Shackley, S., Young, P., Parkinson, S., and Wynne, B.: Uncertainty, Complexity and Concepts of Good Science in Climate Change Modelling: Are GCMs the Best Tools?, Clim. Change, 38, 159–205, https://doi.org/10.1023/A:1005310109968, 1998.

Stanev, E.: Trajectories in Oceanography, Springer Nature Switzerland, Cham, https://doi.org/10.1007/978-3-031-33720-8, 2023.

Steinle, F.: Entering New Fields: Exploratory Uses of Experimentation, Philos. Sci., 64, S65–S74, https://doi.org/10.1086/392587, 1997.

Stewart, R. H.: Introduction to Physical Oceanography., edited by: Texas A & M University, 121–275 pp., 2008.

Stokes, G. G.: On the Theory of Oscillatory Waves., Transactions of the Cambridge Philosophical Society, 8, 441–455, 1847.

Storto, A., Alvera-Azcárate, A., Balmaseda, M. A., Barth, A., Chevallier, M., Counillon, F., Domingues, C. M., Drevillon, M., Drillet, Y., Forget, G., Garric, G., Haines, K., Hernandez, F., Iovino, D., Jackson, L. C., Lellouche, J.-M., Masina, S., Mayer, M., Oke, P. R., Penny, S. G., Peterson, K. A., Yang, C., and Zuo, H.: Ocean Reanalyses: Recent Advances and Unsolved Challenges, Front. Mar. Sci., 6, https://doi.org/10.3389/fmars.2019.00418, 2019.

Sverdrup, H. U., Johnson, M. W., and Fleming, R. H.: Sverdrup, H.U., Johnson, M.W. and Fleming, R.H. (1942) The Oceans, Their Physics, Chemistry, and General Biology., Prentice-Hall, New York, 1942.

Talagrand, O.: Assimilation of Observations, an Introduction, Journal of the Meteorological Society of Japan. Ser. II, 75, 191–209, https://doi.org/10.2151/jmsj1965.75.1B_191, 1997.

Talley, L. D., Pickard, G. L., Emery, W. J., and Swift, J. H.: Descriptive physical oceanography : an introduction, Academic Press, 555 pp., 2011.

Thomson, R. E. . and Emery, W. J. .: Data analysis methods in physical oceanography, Elsevier Science, 2014.

Thorpe, S. A.: An introduction to ocean turbulence, Cambridge University Press, 240 pp., 2007.

Vallis, G. K.: Atmospheric and Oceanic Fluid Dynamics, Cambridge University Press, https://doi.org/10.1017/9781107588417, 2017.

Wanninkhof, R.: Relationship between wind speed and gas exchange over the ocean, J. Geophys. Res., 97, 7373, https://doi.org/10.1029/92JC00188, 1992.

Weisberg, Michael.: Simulation and similarity: using models to understand the world, Oxford University Press, 190 pp., 2013.

Welch, P. D.: The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on Time Averaging Over Short, Modified Periodograms, IEEE Transactions on Audio and Electroacoustics, AI-15, https://doi.org/10.1109/TAU.1967.1161901, 1967.

Winsberg, E. B. .: Science in the age of computer simulation, University of Chicago Press, 152 pp., 2010.

Winsberg, E. B. .: Philosophy and climate science, Cambridge University Press, 270 pp., 2018.

Wunsch, C.: The Ocean Circulation Inverse Problem, Cambridge University Press, https://doi.org/10.1017/CBO9780511629570, 1996.

Young, I. R. .: Wind generated ocean waves, Elsevier, 288 pp., 1999.