

▶ 前沿交叉研判



人工智能海洋学

前沿交叉领域研究报告

人工智能海洋学前沿交叉领域研究 项目组

国家自然科学基金委与中国科学院联合资助

2024

报告撰写组

主 编：陈大可 王 凡

副主编：董昌明 李晓峰

参加编写者（按姓氏笔画排序）：

王久珂	王丽娜	刘 健	朱学明	宇婧婧	张 斌	张守文
张红春	张瑞子	李小燕	李晓龙	沈向宇	周书逸	郑 罡
高 乐	徐广珺	韩国庆	韩 磊	韩 莹	谢华荣	

摘要

随着海洋大数据时代的来临和人工智能技术的蓬勃发展，人工智能海洋学作为新兴的交叉学科逐渐崭露头角。该学科深度融合海洋科学和人工智能技术，为海洋研究和应用提供了新的方向和强有力的科技手段。然而，在人工智能海洋学发展初期，我们也面临一系列挑战。

为了应对当前的诸多挑战，本研究系统梳理人工智能海洋学在建立海洋大数据系统、开发适用于海洋大数据的人工智能算法、海洋智能大模型和海洋数字孪生系统等方面的最新进展。研究发现，人工智能海洋学在数据整合、存储与分析领域正迅速发展，不仅拥有庞大的数据资源，还积累了丰富的数据处理经验。借助深度学习、机器学习等前沿技术，人工智能算法能够更加智能和高效地解读和应用海洋数据，为海洋科学研究开辟了新的可能性，展现了广阔的应用前景。

人工智能海洋学已涉足多个关键领域，包括海洋特征智能识别、海洋要素与现象智能预报、模式参数智能估算等。智能识别海洋特征使我们更准确地了解海洋环境的动态变化，为海洋环境安全提供更为精准的保障。基于人工智能算法的预报模型提高了海洋气象、海洋生态和动力过程等方面的预报精度，为海洋生态环境保护 and 资源可持续利用提供了可靠的支撑。同时，智能估算海洋模式中的各类参数，可节省计算资源并提升模式的模拟和预报能力。

展望未来，我们需要研发定制化的自然语言与信息处理技术(如 Ocean ChatGPT)，需要发展基于人工智能的地球系统耦合大模型，需要进一步夯实海洋数字孪生系统的技术框架和数据底座，需要在算力调度、硬件资源、计算效率和稳定性等方面下功夫，为人工智能海洋学这一海洋科学的革命性分支打下坚实的基础。随着人工智能技术的迅猛发展和相关领域人才队伍的不断壮大，人工智能海洋学的发展前景将远超我们的想象，将为人类探索海洋的奥秘和应对全球气候变化等重大问题提供更为强大的、智慧高效的工具。

Abstract

With the advent of the era of ocean big data and the flourishing development of artificial intelligence (AI) technology, AI oceanography has emerged as a new interdisciplinary science. By deeply integrating marine sciences with AI technology, it can provide strong support to the field of oceanography in a totally unprecedented manner. However, AI oceanography is still in its infancy and is facing many challenges, such as the lack of clarity in the direction it goes, in the scientific problems it addresses, and in the planning of its development pathways.

This study reviews the current state of AI oceanography, including new advancements in the establishment of ocean big data systems, the development of relevant AI algorithms, and the construction of the digital twins of the ocean. It is found that AI oceanography is rapidly advancing in the areas of data integration, storage and analysis, accumulating not only vast data resources but also rich experience in data processing. Via cutting-edge technologies such as deep learning and machine learning, AI oceanography enables more intelligent analysis and application of ocean data, proving new research opportunities and possibilities.

In terms of applications, AI oceanography has made significant strides in research areas such as intelligent recognition of ocean features, intelligent prediction of ocean phenomena, and intelligent estimation of model parameters. Intelligent recognition of ocean features allows for a more accurate understanding of dynamic processes, providing precise safeguards for marine environment. AI-based intelligent models can enhance the prediction accuracy for ocean meteorology and marine ecosystems, offering reliable support for marine environment protection and sustainable resource utilization. Intelligent estimation of model parameters can largely increase models' computational efficiency as well as their simulation and forecast ability.

Looking ahead, we need to develop customized natural language processing technologies (i.e. Ocean ChatGPT), to build AI-based Earth system big models, to further solidate the framework and database for the digital twins of the ocean, and to work

collectively on computing power scheduling, hardware construction, computing efficiency and stability, providing a robust foundation for AI oceanography, a truly revolutionary branch of ocean sciences. With the rapid development of AI technology and the continued growth of research teams, AI oceanography will go far beyond our imagination, providing powerful intelligent support for us to explore the mysteries of the ocean and address critical issues such as global climate change.

I 目录

摘要	I
Abstract	II
一、人工智能海洋学的发展现状	1
1. 海洋大数据系统的建立	1
2. 面向海洋大数据的人工智能算法开发	2
3. 人工智能在海洋学中的应用	3
3.1 海洋特征智能识别	3
3.2 海洋要素与现象的智能预报	10
3.3 模式参数的智能估算	15
3.4 模式结果的智能订正	16
3.5 偏微分方程的智能求解	18
4. 人工智能大模型的发展	19
5. 海洋数字孪生系统构建	22
二、人工智能海洋学的重点方向	24
1. 海洋灾害的智能预警预报	24
2. Ocean ChatGPT	24
3. 海洋人工智能大模型	25
4. 海洋数字孪生系统	25
5. 算力优化与统筹	26
三、结语	27
参考文献	28

一、人工智能海洋学的发展现状

人工智能的发展始于 20 世纪 50 年代，随着 21 世纪初大数据、深度学习和计算能力的提升，人工智能进入高速发展时期。人工智能已深入医疗、金融、制造等领域，成为推动科学研究和社会经济发展的核心技术，将塑造未来，对人类社会发
展进程产生革命性影响。人工智能海洋学是海洋科学与人工智能相融合的前沿学科，近年来发展迅速，在海洋领域中发挥的作用主要体现在以下五个方面。

1. 海洋大数据系统的建立

海洋大数据的特征被概括为典型的 5V，即数据体量大（Volume）、增长速度快（Velocity）、种类和来源多样（Variety）、价值密度高（Value）、质量上乘（Veracity）（Mayer-Schönberger 等，2013）。这五个特征反映了海洋领域数据处理的复杂性与挑战性。

随着近几十年来海洋观测和数值模拟技术的飞速发展，海洋数据的储存量已经达到了 PB 量级，标志着海洋科学已经迈入了大数据时代。现有用于智能算法开发的海洋数据主要包括卫星遥感观测数据、数值模式数据和再分析数据（Qian 等，2022）。

传统的海洋遥感数据依赖于具备专业知识的人工处理和判读，存在准确性受人为因素影响、耗时、消耗人力资源等问题。在人工智能海洋学中，智能化处理海洋遥感数据已成为一个重要研究方向。借助人工智能技术，特别是机器学习和深度学习算法，能够实现对海洋遥感数据的自动化处理和分析，提高数据处理的速度和准确性。

海洋数值模式是海洋动力学、连续性方程、热力学等方程组的数值表达，用于模拟和预测海水的运动、温度、盐度等动态特征的变化。这些模式持续产生大规模的海洋数据，模拟海洋系统的复杂性，预测海洋环境的变化，为科学研究和决策提供重要依据。

海洋再分析是综合利用多种观测数据与数值模型进行耦合，采用数据同化技术，生成高质量的数据集，用于对过去和现有海洋状态的分析 and 重构。对于海洋科学研究以及预测模式的初始化，网格化的海洋再分析产品是最主要的数据支撑之

一。

除了大数据的 5V 特征外，海洋数据还具有高维、异构、时空高度稀疏的特征。这些特征给海洋数据的处理和分析带来挑战，需要采用特定的方法和技术来处理 and 提取其中的有用信息。

2. 面向海洋大数据的人工智能算法开发

海洋大数据涉及物理海洋、海洋遥感、海洋化学、海洋生物、海洋地质和海洋生态等多个研究领域的相关变量，可为海洋环境预报、灾害防范、作业生产和经济政策制定等提供关键的信息服务和决策支持。海洋大数据不仅涵盖了广泛的研究领域，而且各海洋要素之间具有复杂的时空关联性。

为了消除海洋数据的孤立性，提高数据的可用性和价值，需要进行海洋多源异构数据的集成。首先，从不同来源获取数据，经过预处理和清洗，确保数据的准确性和一致性。然后，对数据进行有效的存储和管理，以便快速访问和检索。在数据集成和匹配阶段，解决异构数据的结构和格式差异，确保数据能够被有效集成。同时，为了控制数据质量，还需要建立相应的评估体系，确保数据的可信度。

在利用人工智能算法进行海洋多源异构数据集成的研究中，Zhang 等(2023a)使用了一种时空生态集合模型的集合学习方法，构建了 1997-2020 年全球尺度上的每月浮游植物群产品（图 1）。该研究表明，人工智能算法对于集成海洋多源异构数据有无可比拟的优势，值得进一步探索相关的技术方法和应用场景。

面对海洋数据的稀疏性，研究者们正在积极探索数据融合和同化技术，将来自不同观测源的数据整合在一起，通过数值模式将其融合，从而填补数据的空白，提高数据的空间和时序连续性。其中，数据同化是一种重要的技术方法，它能结合观测数据与数值模式，输出比观测更精细、比模式更准确的数据产品。人工智能技术目前已开始应用于海洋大数据的融合和同化，有望为海洋研究和预警预报提供更完整、及时的数据，也为政策制定者提供更可靠的科学依据。

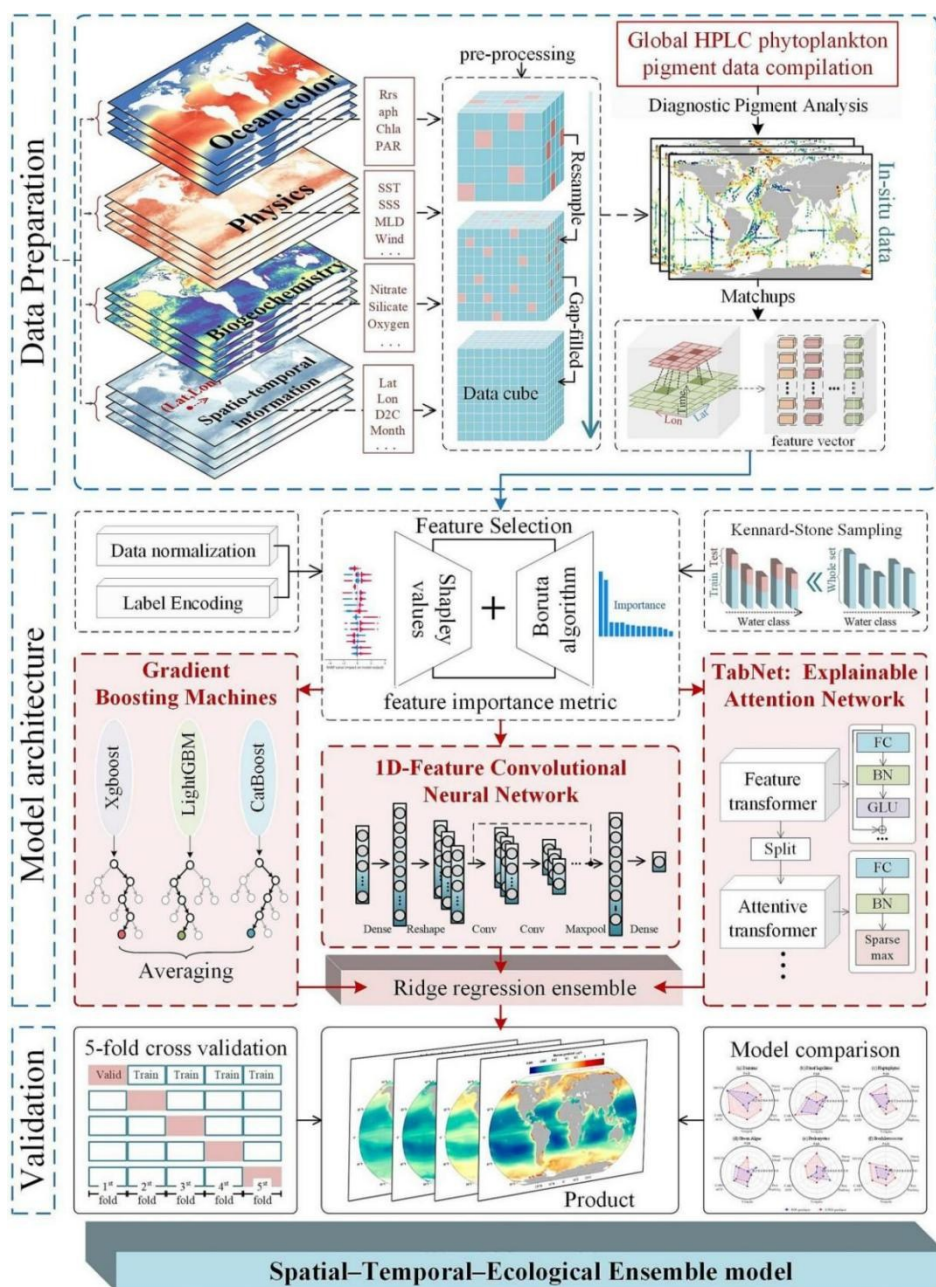


图 1. 时空生态集合（Spatial-Temporal-Ecological Ensemble, STEE）模型示意图。采用 STEE 的集合学习方法，引入全球主要海洋区域的水色、物理海洋学、生物地球化学以及时空信息数据，生成了 1997-2020 年全球尺度上的每月浮游植物群产品。（Zhang 等，2023a）

3. 人工智能在海洋学中的应用

3.1 海洋特征智能识别

海洋特征智能识别是利用人工智能技术实现海洋中物理、生物和地质等特征的快速、准确判别，对深入理解海洋系统具有重要意义。下面将系统回顾海洋涡旋、

海洋内波、海上溢油、海上船只、海洋藻类、海洋珊瑚和深海沉积物智能识别的研究进展。

（1）海洋涡旋

海洋涡旋是一种广泛存在于海洋中的封闭环状旋转水体，它在全球海洋中物质、能量和热量的输运中扮演着重要的角色。早在 2016 年，国外学者就采用机器学习方法进行海洋涡旋探测，其识别准确率高达 97%，远超传统识别算法（Ashkezari 等，2016）。这标志着人工智能在海洋涡旋自动识别领域的首次成功应用。随后，深度学习模型（如基于编-解码框架的 U-net、REDN 等模型）也被引入到海洋涡旋的智能识别研究，进一步提升了识别的精度和效率（Franz 等，2018；Santana 等，2022）。与此同时，国内学者也积极开展相关研究工作，不仅改进模型结构以提高涡旋识别准确性（Duo 等，2019；Xu 等，2019，2021；Sun 等，2021），还利用融合多源遥感数据的人工智能模型，获得新发现（Liu 等，2021a，图 2）。然而，海洋涡旋具有复杂的三维结构（Zhang 等，2016），现有智能识别方法难以揭示涡旋的完整形态。未来通过高度融合海洋知识和人工智能算法，有望通过不断增加的训练数据，进一步实现涡旋三维结构的智能识别。

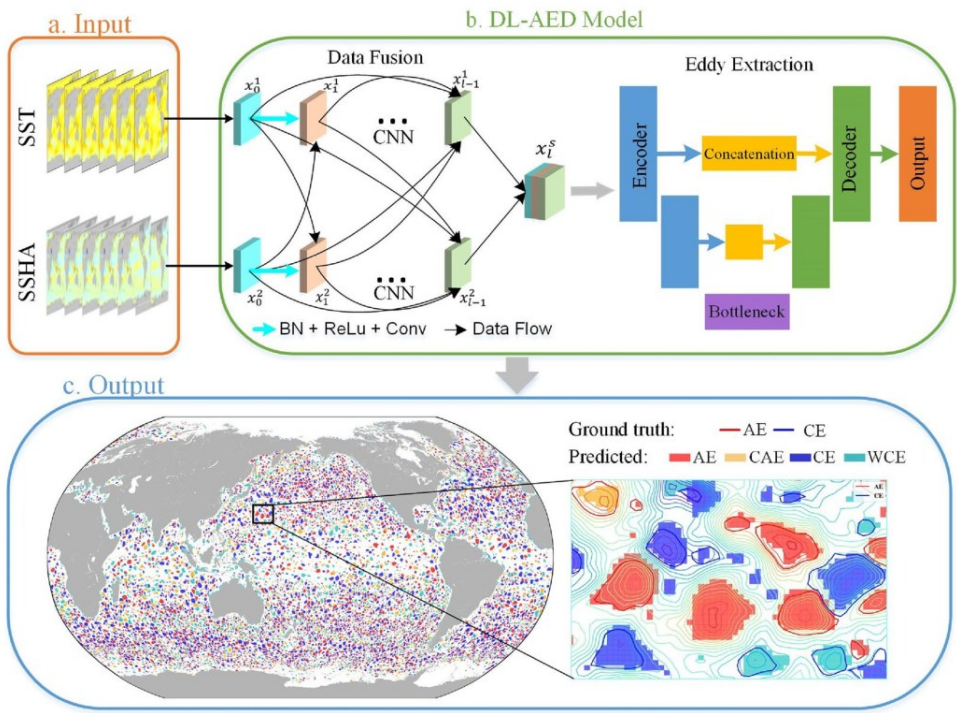


图 2. 基于海面高度和温度异常的海洋涡旋探测深度学习模型。（Liu 等，2021a）

（2）海洋内波

内波（Internal Wave, IW）是海洋中一种常见的自然现象，对海洋声学、海水混合、近海工程和海底导航等都有重要影响。卫星遥感是唯一可以对内波开展大范围长时序观测的手段，目前已成为研究内波时空分布、传播、演变等规律的主要手段（Li 等，2008）。21 世纪以来，国际学者率先利用传统的图像处理方法从卫星图像中提取内波信息（Simonin 等，2009），但这些传统方法容易受到图像噪声的干扰。相比之下，机器学习方法在从卫星图像中提取内波信息方面表现更为可靠（Li 等，2020）。Zhang 等（2022）基于深度学习网络设计了 IWE-Net，针对内波特征的模型结构和损失函数，实现了对不同载荷、不同分辨率遥感图像内波信息的自动提取，达到较高精度（图 3）。基于深度学习模型可以显著提高在有挑战性的成像条件下内波信号提取的准确性，以弥补现场观测的不足，促进内波的生成、传播、演化和消散研究。然而，深度学习模型在少数海域的准确性仍有提升空间，未来可以通过获取更多含有丰富内波特征的影像进行训练来解决。

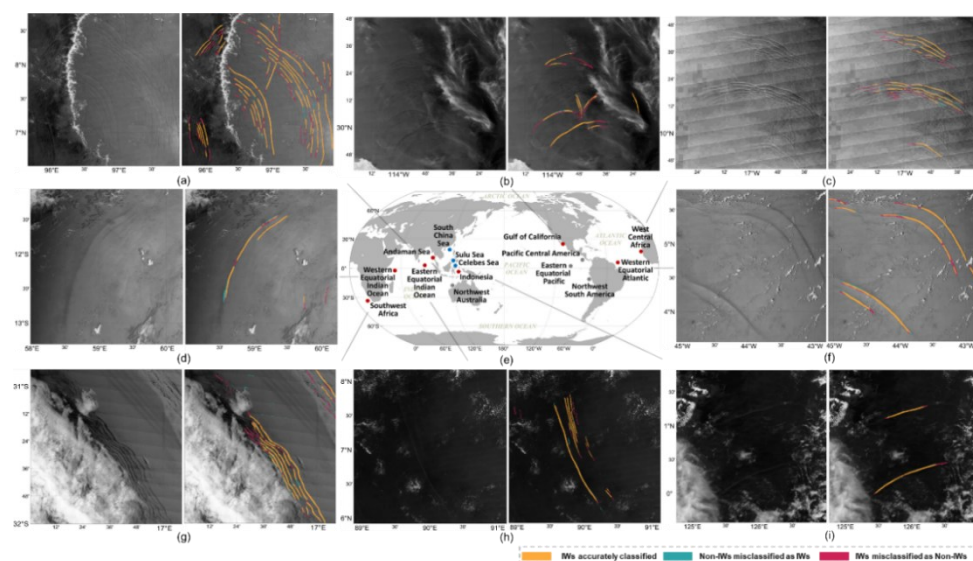


图 3. IWE-Net 在全球范围内应用的八个 MODIS 影像示例。（Zhang 等，2022）

（3）海上溢油

海上溢油是一种典型的海洋污染事件，通常由船舶事故、海上油井爆炸、输油管道破裂等原因引发，对海洋渔业、水产养殖、生态系统、海上旅游和运输等都会带来严重影响（Crone 等，2010；Li 等，2021）。卫星遥感影像被广泛应用于海上溢油事件的探测。通过不同深度学习模型，可以深度挖掘影像中的油膜特征，不仅

能够快速识别海面溢油特征（Krestenitis 等，2019；Yekeen 等，2020），还可以基于影像的纹理和形状对溢油区域进行分类（Bianchi 等，2020，图 4）。相较于语义分割模型，实例分割模型在溢油智能识别方面表现更佳。国内学者也积极从事海上溢油智能识别的研究工作。通过改进模型结构或引入更多特征信息，不仅能有效提取影像中海面溢油的细节特征（Wang 等，2022），还成功实现了对不规则形状溢油事件的准确识别（Dong 等，2023）。卫星遥感影像的质量对海上溢油事件的智能识别至关重要。未来通过改进光学和合成孔径雷达的成像和预处理技术，有望进一步提升海面溢油智能识别模型的性能。

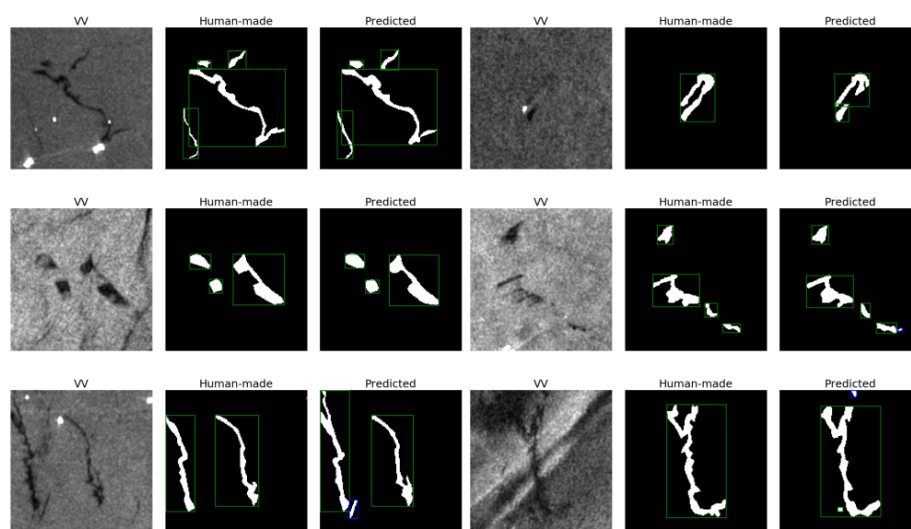


图 4. 溢油智能识别模型在验证集上的结果。从左到右依次为：VV 极化输入、人工标定、模型输出。绿色方框标记人工和模型共同识别部分，蓝色方框标记仅模型识别部分，红色方框标记仅人工识别部分。（Bianchi 等，2020）

（4）海上船只

船只检测是海洋监测中的一项重要任务，对提高海防预警能力和加强海运监测与管理具有重要意义。合成孔径雷达（SAR）因其全天候、全天时的监测能力，在海洋船只检测中得到广泛应用（Zhu 等，2010；Ouchi 等，2004）。学者们针对 SAR 影像中的船只检测提出了一系列算法，可以分为传统方法和人工智能方法两大类。基于阈值分割的方法是一种传统的船只检测方法，通过查找明暗差异明显的像素值，利用滤波统计建模。传统方法的主要缺点在于需要先验知识来手动设计参数和特征，这也是大数据时代许多领域所共有的挑战（LeCun 等，2015）。与传统方法相比，深度学习可以自动提取船只特征，更迅速、更准确地检测船只。例如，

基于深度学习和旋转边界框的 DRBox (Liu 等, 2017)、基于快速区域卷积的 Faster-RCNN (Deng 等, 2019, 图 5) 等方法都已经应用于船只检测。基于深度神经网络的船只检测模型大大简化了特征提取过程, 具有更高的精度和稳定性。未来, 利用智能模型从 SAR 影像或光学影像中进一步检测船只的类型、几何尺寸等参数, 将有助于提高海上监测和管理的效率。

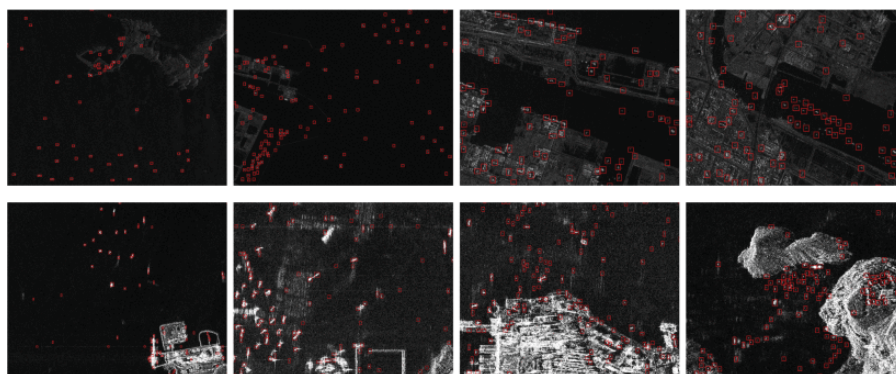


图 5. 使用 Faster-RCNN 于多张 SAR 影像的船舶检测结果。(Deng 等, 2019)

(5) 海洋藻类

海洋藻类在生态系统中发挥着积极作用, 但近年来不断爆发的有害藻华事件对海洋生态系统和人类社会造成重大负面影响。智能检测海洋藻类, 对防控藻华灾害的发生至关重要。卫星遥感能够大幅宽、周期性地监测海洋, 是藻华监测的主要技术手段, 同时深度学习在光学和 SAR 遥感影像的藻华识别中展现出巨大潜能。Arellano-Verdejo 等 (2011) 提出了一种基于一维 CNN 的 ERISNet 模型, 应用于墨西哥海岸线的大型马尾藻提取, 准确率高达 90.08%。Guo 等 (2022) 提出了一种用于 SAR 图像绿藻检测的 GA-Net 模型 (图 6), 相较于经典的深度学习模型, 该模型综合考虑了归一化雷达散射截面和纹理信息, 具有更高的准确性和泛化能力。光学和 SAR 影像由于不同的成像机制和时空分辨率, 其海藻识别的结果虽然总体分布特征一致, 但细节往往有很大差异, 比如光学的 MODIS 影像往往常漏检海藻细微条带, 而 SAR 影像却可以同步获取更多细节 (Gao 等, 2022)。未来通过融合不同卫星的海藻特征数据, 或者综合考虑不同类型的损失函数, 有望进一步提高智能模型识别藻类的准确性和鲁棒性。

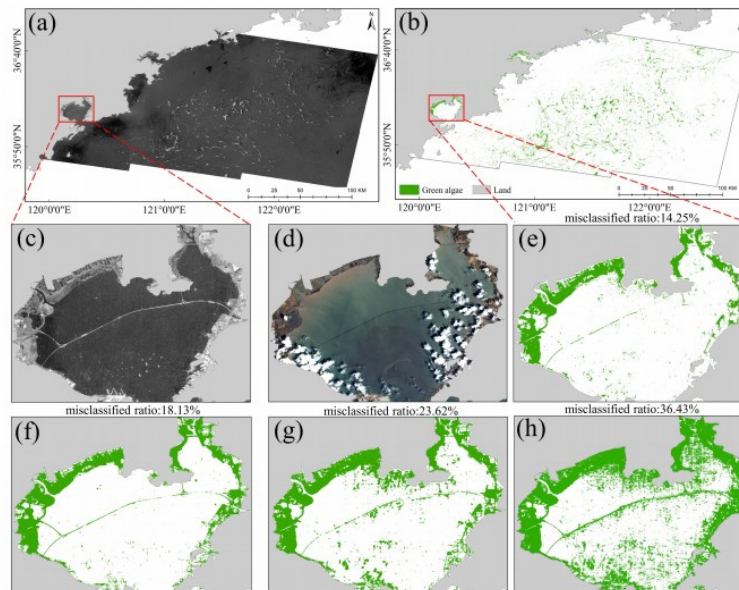


图 6. 2019 年 6 月 23 日基于 SAR 图像识别的绿藻。(a) UTM 投影下 VV 极化的 SAR 影像；
(b) GA-Net 模型检测结果；(c) - (h) 为 GA-Net/VGG-16/AlgaeNet/U-net 模式在该区域的
识别结果。(Guo 等，2022)

(6) 海洋珊瑚

珊瑚是海洋生态系统的重要组成部分，也是气候和环境变化的指示物，识别其分布与变化对于保护海洋生态、应对气候变化至关重要。然而，传统的基于形态学和分子遗传学的生物识别方法在实际海洋环境中的应用受到限制。随着计算机视觉和深度学习技术的发展，学者们开展了海洋珊瑚的智能识别研究。Jaisakthi 等 (2019) 提出了一种基于卷积神经网络的方法，用于定位和检测不同类型底栖生物，但这类传统的检测构架很难实现高精度的实时检测。针对这一问题，Li 等 (2020a) 利用深度学习网络实现了对珊瑚的智能识别 (图 7)。他们通过分帧处理和手工标注，构建了不同类型珊瑚的视频数据集，对图像进行标准化、裁剪和增强处理，以提高模型的鲁棒性和准确性。Liu (2024) 等利用深海图像数据，进一步提升了细粒度识别的准确性。珊瑚的形态多样性和物种丰富性导致高细粒度的分类检测很难实现，同时海洋的成像环境异常复杂且特殊，导致高精度实时检测非常困难。未来针对海洋特殊复杂成像环境，优化数据增强算法的，扩大珊瑚样本数量，有望进一步提高海洋珊瑚的识别精度和颗粒度。

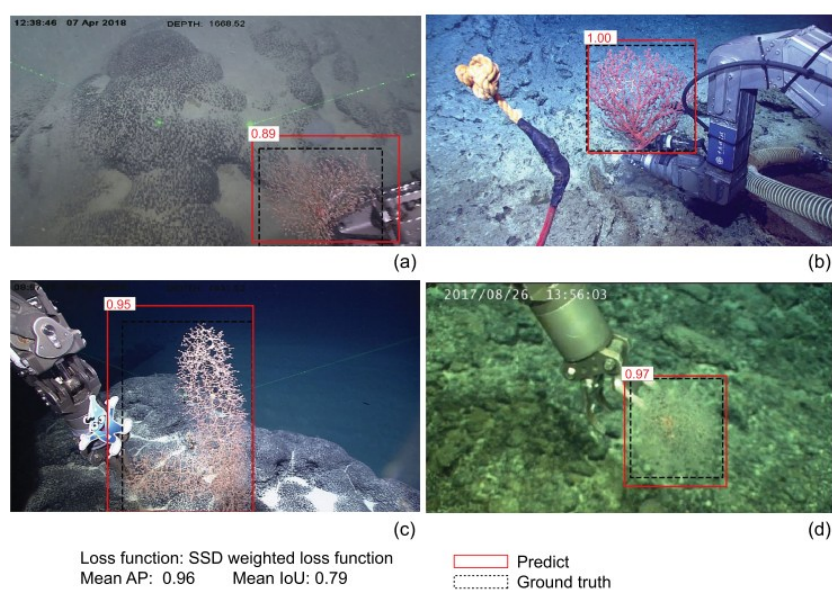


图 7. 基于 VGG 模型识别的不同类型珊瑚。白框中的数字代表每种珊瑚的识别置信度。（Li 等，2020a）

（7）深海沉积物

深海沉积物蕴藏着大量的资源，记录了地球环境变迁的历史。沉积物的广泛分布和多样性使其识别成为一项具有挑战性的任务（图 8）。近年来深海沉积物识别开始转向深度学习技术。例如，江峦等（2020）提出了基于深度学习的沉积物测试系统，通过神经网络对数据进行误差修正并实现智能识别。Ruan 等（2022）则设计了基于图像识别技术的采样系统，利用智能识别模型框架提高识别准确性。未来通过优化特征提取和数据增强算法，将样本数据转化为高维特征空间，并应用先进的人工智能算法，有望进一步提升深海沉积物识别的准确性，促进对深海环境和资源的深入了解，以支持相关科学研究和应用。

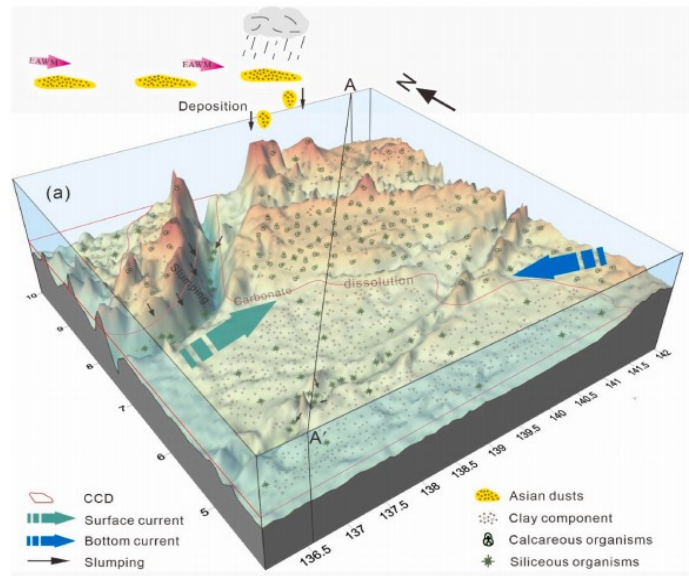


图 8. 热带西太平洋不同组成成分的沉积物的空间分布。(Guo 等, 2023)

3.2 海洋要素与现象的智能预报

数据驱动的海洋智能模型利用人工智能技术，学习数据中隐含的海洋变化内在特征和规律，实现对各种海洋要素与现象的预报，目前已应用于从短期气候变化到中小尺度海洋过程的诸多领域。

(1) ENSO 智能预测

近年来，AI 技术在 ENSO（厄尔尼诺-南方涛动）预测方面的应用取得了显著进展。Ham 等（2019）建立了一个基于 CNN 的 ENSO 预测模型，宣称能提前 17 个月预测 NINO3.4 指数，预测能力显著高于传统的动力模型。Gupta 等（2020）使用卷积长短时记忆网络（ConvLSTM），可提前一年预测月平均 NINO 3.4 指数，甚至能够预测极端厄尔尼诺现象。为了更好地解释神经网络的预测结果，Cachay 等（2021）将时空图神经网络应用于 ENSO 预测。与传统的 CNN 相比，这一模型能更充分地利用输入数据的时空信息，提高 ENSO 预测的准确性。Liang 等（2021）采用梁氏-克莱曼信息流的因果人工智能模型，声称提前 12 年再现了 ENSO 事件。Zhou 等（2023）创建了一种分时空三维卷积神经网络模型 TS-3DCNN，用于预测 2020 年至 2021 年的 La Niña（图 9），结果表明该模型可以在一定程度上捕获这次特殊的 La Niña 事件。Wang 等（2023）开发了一个可解释的深度学习 ENSO 预测

模型，发现全球海洋与 ENSO 相关的区域随着 ENSO 的变化，有效 ENSO 预测提前量达到 22 个月，并且几乎不受春季可预测性障碍影响。

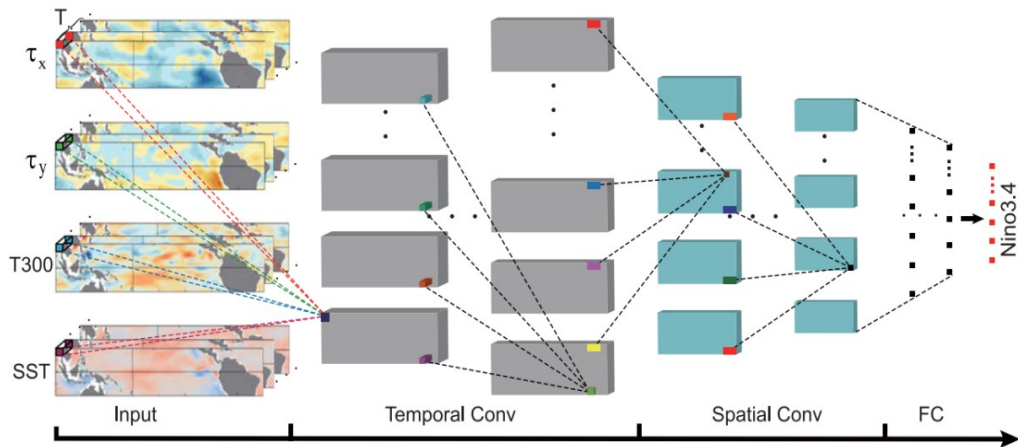


图 9. 用于 NINO3.4 海温异常预测的 TS-3DCNN 模型。(Zhou 等，2023)

（2）风暴潮智能预报

风暴潮是一种由剧烈的大气扰动引发的水位急剧上升现象，其是否成灾主要取决于风暴潮增水是否与天文大潮的潮汐周期相叠加。近年来，人工智能技术的崛起为风暴潮预报研究的发展添加了新的动力。Lee（2006）、Rajasekaran 等（2008）和 Hashemi 等（2016）将风速、风向、气压等台风要素作为输入，使用人工神经网络（ANN）或支持向量机对风暴潮进行预报。基于长短时记忆网络（LSTM），刘媛媛（2020）等利用气象要素和潮汐前期时间序列，建立了单个站点的风暴潮近期预报模型。该模型同时考虑了四种不同输入参数的组合，对未来 1 至 3 小时内的潮汐高度进行预报。谢文鸿等（2022）利用 CNN 模型直接提取二维风场和气压场的特征，并将其与潮汐时间序列融合，用于预报风暴潮站点的水位序列（图 10）。与站点观测相比，模型预报的 24 小时内风暴潮水位的相关系数超过 0.95，表明模型的良好性能。此外，Xie 等（2023）将 ConvLSTM 模型应用于风暴潮漫滩过程的智能预测。与传统的动力数值预报相比，该模型的计算效率更高，且在短期预报中能够取得与数值预报准确性相当的结果。

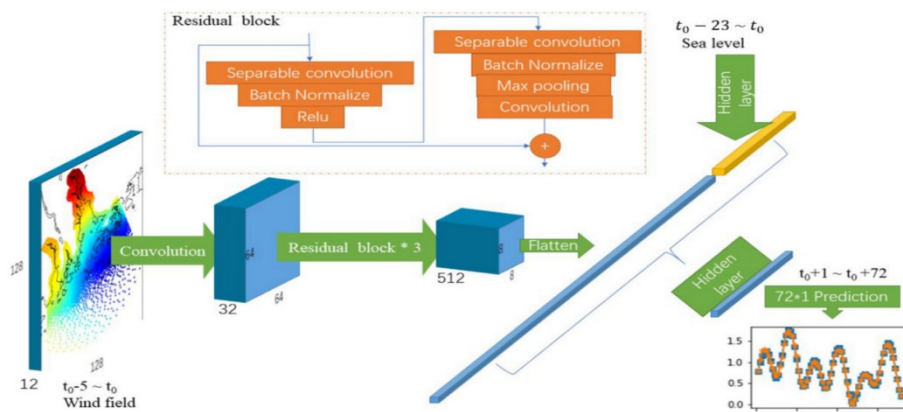


图 10. 基于 CNN 的风暴潮站点水位智能预测模型。(Xie 等, 2023)

(3) 海洋波浪智能预报

海浪具有极强的非线性特性,对海洋工程、海上作业活动和交通运输等有重要影响。已有大量研究利用人工智能技术对单点海浪进行预报。Londonhe 和 Panchang (2006) 采用 ANN 对海浪的有效波高开展智能预报研究,结果表明该模型可以较准确地预测未来 6 小时的有效波高,对未来 12 小时的波高预测仍具有 67% 的相关性。Fan 等 (2020) 则将 LSTM 和 SWAN 波浪模式相结合,对波高进行预报。与仅使用 SWAN 模式的预测结果相比,SWAN-LSTM 模型的预报精度提高了约 65%。Zhou 等 (2021a) 提出了一种基于经验模态分解 (EMD) 的 LSTM 算法,将其应用于单点海浪的预报研究。与加勒比海的浮标观测相比,该模型取得了优于 LSTM 算法的预测精度。二维波浪场智能预报通常使用再分析数据作为标签数据训练模型。Zhou 等 (2021b) 使用 ConvLSTM 算法对我国东部和南部海域的有效波高开展预报试验 (图 11),最长预报时长可达 24 小时。随后, Bai 等 (2022) 使用 CNN 算法,通过 56 组敏感性试验进行了南海有效波高的预报试验。

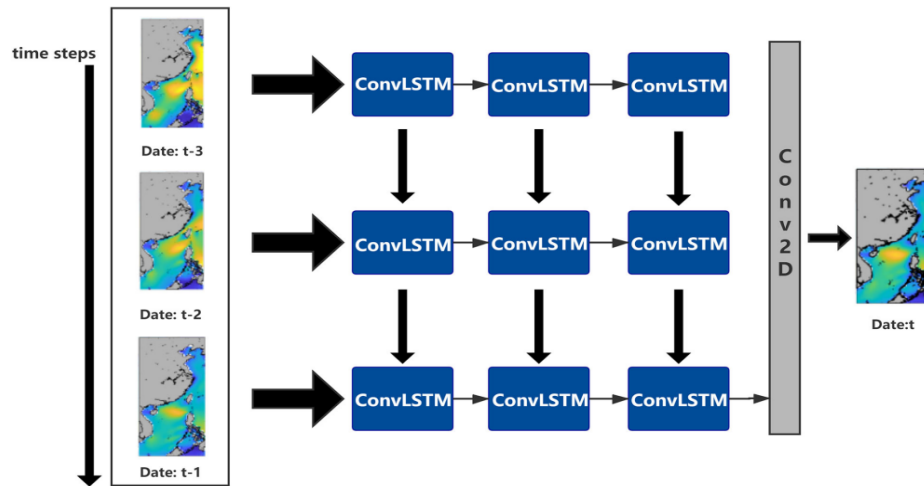


图 11. 基于 ConvLSTM 的有效波高预报模型。(Zhou 等, 2021b)

(4) 海表温度智能预报

对海表温度 (Sea Surface Temperature, SST) 时空变化规律的深入理解及精确预报, 对海洋学、气象学、航海等多个领域有重要意义。目前 SST 的预测方法主要分为两类: 数值模式和数据驱动的智能方法。近年来, 由于深度学习的自适应特征, 尤其是循环神经网络及其扩展, 如 LSTM 和门控循环单元, 利用深度学习实现 SST 的准确预报已经成为当前的研究热点。然而, 单一的循环神经网络预报模型存在学习能力较弱和容易产生过拟合的问题。为解决这一问题, 研究人员尝试将循环神经网络与其他预报模型相结合, 以提高预报准确性。例如, Xiao 等 (2019a) 构建了 LSTM 与 AdaBoost 的组合模型, 对我国东海的 SST 进行预报。针对循环神经网络仅考虑时间特性而忽略空间特性的不足, Xiao 等 (2019b) 将 CNN 与 LSTM 结合, 进一步提升了 SST 的预测水平 (图 12)。另一方面, 考虑到海洋数据的非线性和多噪音特性, 学者们还尝试在海洋预报模型中引入数据预处理模块以消除噪音, 从而提高预报精度。

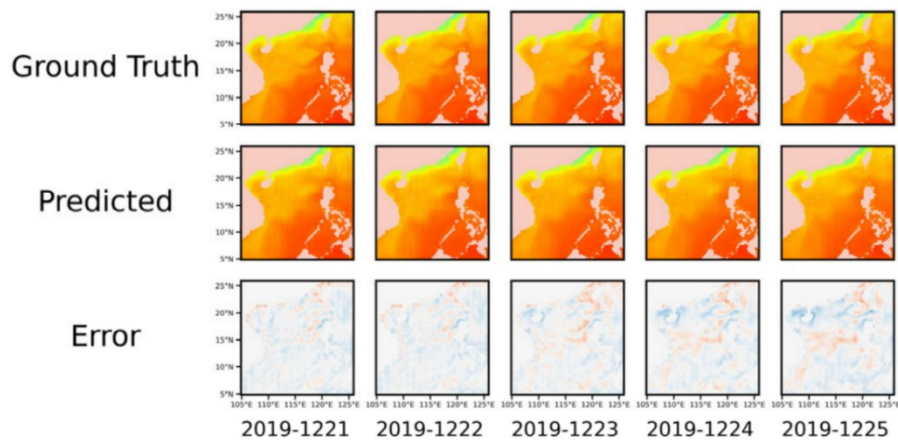


图 12. 基于 ConvLSTM 提前 5 天预报的海表温度。(Xiao 等, 2019b)

(5) 海面风速的智能预报

利用不同的神经网络算法对海面风速的预报, 取得了优于传统方法的结果。一种常见的方法是采用分解集成框架, 通过提取具有周期性和趋势性的子序列来提高风速的预报精度 (Jiang 等, 2021)。另一种方法是采用组合加权框架, 首先针对相同的数据利用不同的预报模型分别进行预报, 然后利用加权算法或优化算法将所有结果组合, 得到最终的风速预报结果 (Niu 等, 2019)。Nie 等 (2021) 利用 ICEEMDAN 分解方法处理风速数据, 然后分别使用 Bi-LSTM、ELM 和 BP 神经网络对其子序列进行预报, 通过多目标优化算法融合子序列的结果, 得到最终的风速预报。同样地, Wang 等 (2021) 利用 ELM、BPNN 和小波神经网络等分别对风速进行预报, 然后使用灰狼优化算法获得最终结果。此外, 一些新的技术, 包括物理信息方法 (Geng 等, 2021)、图神经网络以及图注意力机制等, 也被应用于风速的预报研究。未来通过不断丰富和改进预报的技术工具箱, 有望进一步提高预报能力。

(6) 海冰的智能预报

北极海冰预报的主流方法包括数值模式和统计模型。统计模型从数据本身出发挖掘和拟合海冰的变化规律, 比数值模式灵活易用, 特别是在面对复杂的因素和尚未明确的影响规律时具有优势 (Horvath 等, 2020)。深度学习方法从本质上来说是一种新兴的统计方法。通过构建多种深度神经网络模型, 学者们从海冰历史时空序列中挖掘规律, 实现对多尺度海冰时空变化的预报, 显著提高了预报精度 (Chi 和 Kim, 2017; Choi 等, 2019; Liu 等, 2021b)。深度学习能够较好地捕捉海冰的变化规律, 提高对海冰密集度等关键参数的预报能力。但由于海冰变化涉及多尺度

的物理过程，尤其在夏秋季节融池现象较为严重时，影响其变化的热力和动力因素更加复杂，纯数据驱动的智能模型通常难以准确预报异常海冰变化。如何将深度学习与海冰变化的物理机制相结合，是未来需要解决的问题。

（7）海洋热浪的智能预报

海洋热浪通常被定义为持续至少五天的极端海表高温事件。在全球气候变化背景下，海洋热浪的强度和频率逐渐增加，对海洋环境和生态系统产生了显著影响。近来人工智能也被应用于海洋热浪的预报研究。Sun 等（2023）提出了一种结合 U-Net 和 ConvLSTM 网络的方法，用于预报南海海域的海洋热浪。他们使用 U-Net 模型对海洋热浪的强度进行预报，并利用 ConvLSTM 模型预报其发生概率。尽管随着预报提前时间的增加，热浪的预报精度会逐渐降低，但对于 7 天的预报提前时间，模型的准确性仍能保持在 80%以上（图 13）。未来随着人工智能技术的不断进步和更多海洋数据的积累，有望进一步提升对热浪的预报能力，为应对气候变化、保护和管理海洋提供支撑。

3.3 模式参数的智能估算

对数值模式中的次网格过程进行参数化是一个难题，而参数的计算通常需要耗费大量资源。数据驱动的人工智能方法对参数化不需要做任何的物理假设，且模型一旦建立则计算非常简便。这类参数化在一些理想模式中已经得到了应用（Berloff, 2005; Zanna 等, 2017;

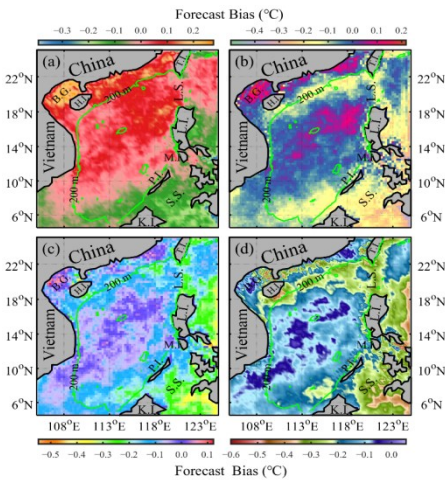


图 13. 提前 1 天（a）、3 天（b）、5 天（c）和 7 天（d）热浪强度预报误差的空间分布。（Sun 等，2023）

Duraisamy 等, 2019)。通过从大量试验中寻找规律并估算最优参数取值 (Zhang 等, 2016; Han 等, 2022), 或从海量观测数据或高分辨率模式结果中寻找规律并建立新的参数化方案 (Bolton 和 Zanna, 2019), 智能参数化模型能提升模式的模拟和预报能力。当前研究多是围绕大气环流模式 (Zhang 等, 2016) 或台风预报模式 (Jiang 等, 2018) 展开, 结合机器学习的海洋模式参数化研究刚刚起步。

基于深度学习的湍流参数化方案最早被应用于数值模式 (Tracey 等, 2015; Ling 等, 2016)。Ling 等 (2016) 使用深度全连接神经网络学习各向异性应力张量的涡动量参数化, 发现只有将物理约束引入网络, 才能取得优于线性回归模型的结果。此外, 部分研究使用神经网络对自由衰减二维湍流模型 (Maulik 和 San, 2017; Cruz 等, 2019) 和大涡模拟的涡动量通量进行了参数化。在海洋学研究领域, 学者们利用 CNN 模型分别对海洋中尺度涡、垂向混合速率等进行参数化, 并指出智能参数化模型可以推广到不同动力状态的海洋数值模式中。Zhu 等 (2022) 利用热带太平洋的湍流观测记录, 设计了首个基于深度学习的海洋垂向混合参数化方案 (图 14), 并将其应用于海洋环流和海气耦合数值模式。他们发现, 数据驱动的湍流混合参数化方案可以很好地模拟上层海洋的垂向热通量, 从而提升热带太平洋海水温度的模拟精度。

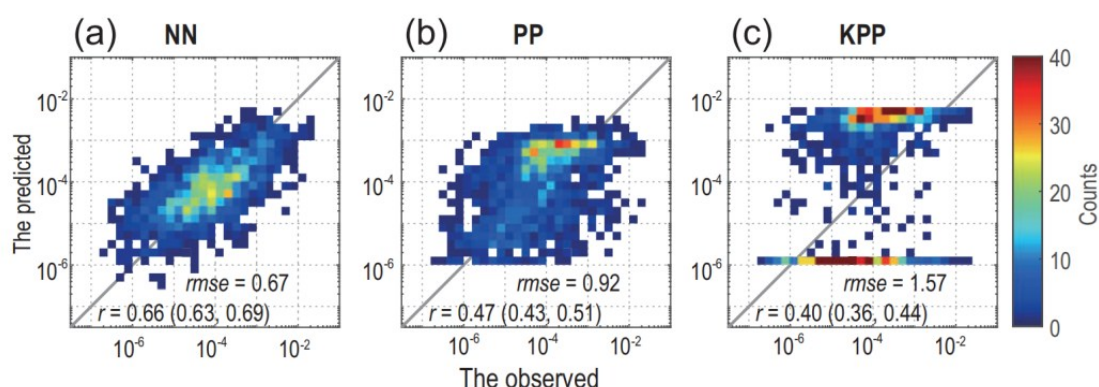


图 14. 观测的和不同方法估算的垂直涡扩散系数之间的二维直方图。NN 表示神经网络, PP 表示 PP 参数化方案, KPP 表示 KPP 参数化方案。(Zhu 等, 2022)

3.4 模式结果的智能订正

数值模式是当今气候预测的主要工具, 但由于模式中物理过程的参数化处理和微分方程的离散化表达, 数值模式存在系统性偏差。相较于基于数学统计的传统

订正方法，基于深度学习的订正模型可以更有效地捕捉数值模式预测结果与实际观测之间的非线性关系，从而实现更精准的模式订正。Chapman 等（2019）利用 CNN 订正了 GFS 模式输出的东太平洋和美国西部的水汽传输，预测结果的均方根误差减小了 9%-17%，相关性增加 0.5%-12%。Sayeed 等（2021）使用深度 CNN 对 WRF 模式的预报结果进行订正，结果表明，90%以上的站点风速、相对湿度和每小时降水的预报精度得到了显著提升。Zhang 等（2022）使用了线性回归、LSTM-FCN 和 LightGBM 这三种机器学习方法对 GRAPES-3km 模式预报的 2 米温度进行订正，结果表明三种机器学习方法均能很好地修正温度预报结果，均方根误差分别下降了 33%、32%和 40%，LightGBM 方法订正后的模型准确性更是高达 84%。Son 等（2022）将 CFS-v2 与图像超分辨率深度学习技术相结合，构建了混合火灾天气预报系统，显著提高了预报火灾天气的准确性，提前期长达 7 天，空间分辨率提升至 4 公里。

受海上观测资料限制，海洋数值模式的误差订正主要集中在海洋波浪、盐度和温度等领域。Makarynskyy（2004）使用 ANN 对未来 24 小时波参数进行预报，并对初始预报值进行修正，进一步提高预报精度。Londhe 等（2016）使用 INCOIS 模式波高历史预报误差作为输入，对第 24 小时的波浪预报数据进行订正。Gracia 等（2021）结合多层感知机和 LightGBM，将数值模式输出的有效波高、波周期等参数的预报误差减小了 36%。Jang 等（2021）分别利用 ANN、随机森林和支持向量回归对 SMAP 海表面盐度（SSS）产品进行改进，结果表明智能模型能够有效降低 SSS 产品的误差，且随机森林的性能优于另外两种模型。Choi 等（2022）提出了一套基于深度生成图像网络、卫星数据与海洋数值模式初始场的海洋温度实时订正系统，使 SST 的预报误差降低了 0.5°C，有效缩小了资料同化与预报的差距，并提高了远海预报系统的精度。Yuan 等（2023）提出了一种基于生成对抗网络的 SST 偏差订正及降尺度整合模型，并针对 ENSO、IOD 以及海洋热浪事件开展了试验（图 15）。该模型在提高分辨率到 0.0625°×0.0625°的同时将预测误差减少约 90.3%，突破了观测数据分辨率的限制，且与观测结果的结构相似性高达 96.46%。

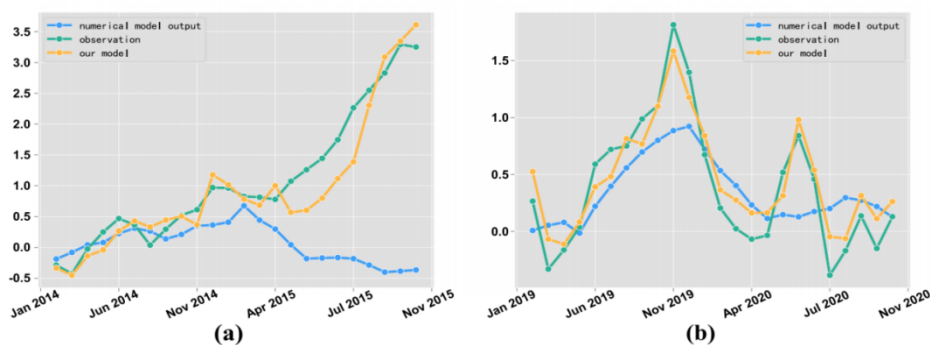


图 15. 智能订正的 (a) NINO3.4 和 (b) DMI 与数值模式和观测值的对比。(Yuan 等, 2023)

3.5 偏微分方程的智能求解

偏微分方程 (Partial Differential Equation, PDE) 在物理规律的建模中具有重要地位, 常用于表征复杂的物理系统。物理约束神经网络 (Physics-Informed Neural Networks, PINNs) 将物理系统的 PDE 作为正则化项融入损失函数中 (Raissi 等, 2019), 在解决关键动力学问题的机器学习中具有广泛的应用潜力。然而, 在海洋科学等领域, 作为传统数值方法 (如有限差分或有限体积) 的替代方案, PINN 仍然面临一些未解决的问题, 从而限制了结果的准确性 (Zhu 等, 2022)。

求解 PDE 是由数值求解器在高维坐标空间上对多个耦合方程进行离散化的结果, 输入与输出之间的非线性映射关系极为复杂。为了求解适用于真实场景的非线性 PDE, 学者们尝试使用泛化性更强的神经算子来获得近似解。例如, DeepONet (Lu 等, 2021)、Fourier Neural Operator (Li 等, 2020b) 和 Koopman Neural Operator (Xiong 等, 2023a, 图 16) 等均是目前广泛使用的神经算子。然而, 通过学习单一神经算子仍然难以拟合复杂的非线性映射关系, 而且这些深度神经网络模型通常局限于高维坐标空间, 难以实现模型泛化能力的提高。为了避免冗余且高维的坐标空间, Wu 等 (2023) 提出了一种隐谱模型, 将原始高维数据投影到一个隐谱特征空间, 将复杂的非线性映射近似为多个基算子线性加权的和。此外, Rao 等 (2023) 提出了一种泛化性更强且具有更高抗噪能力的深度学习框架, 基于物理编码递归卷积神经网络, 用于非线性 PDE 的求解。这一方法将给定的 PDE、初始条件和边界条件的先验物理结构知识强制编码到递归卷积神经网络中。

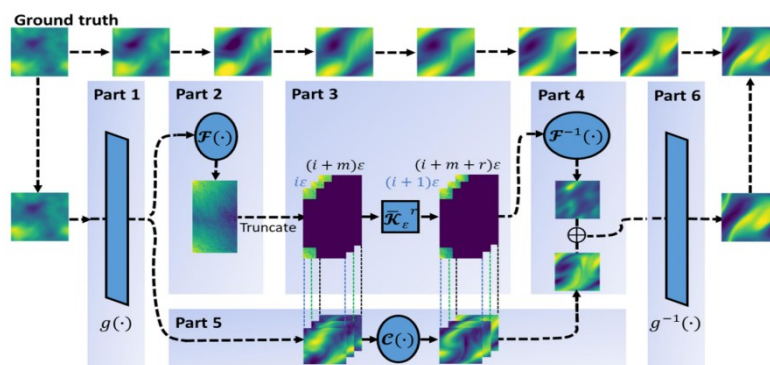


图 16. Koopman 神经算子的网络结构。（Xiong 等，2023a）

偏微分方程的智能求解在海洋和气象预报等领域具有广泛的应用前景。然而，如何在观测数据相对稀疏的情况下，利用人工智能模型实现偏微分方程的快速、准确求解，从而获得具有高精度、鲁棒性、可解释性和泛化性的模拟和预报结果，仍需要进一步探索。

4. 人工智能大模型的发展

人工智能大模型是在大规模数据集上进行预训练形成的模型。近年来，气象大模型的发展取得了令人瞩目的成果，在降水临近预报方面表现尤为突出。谷歌公司 2020 年 3 月推出了 MetNet 模型（Sønderby 等，2020，图 17），是大模型在气象预报领域的首次尝试，耗时仅需几秒钟，远快于数值模式。英伟达等机构合作开发了 FourCastNet 模型（Pathak 等，2022），以 6 小时、 $0.25^{\circ} \times 0.25^{\circ}$ 的时空分辨率完成 10 天全球关键气象要素预测（图 18）。同年 11 月，华为推出了盘古气象大模型（Bi 等，2023，图 19），具备秒级的全球气象预测能力。12 月，DeepMind 和谷歌合作发布了 GraphCast 大模型（Lam 等，2022），其 99.2% 的结果具有优于盘古大模型的预测精度。

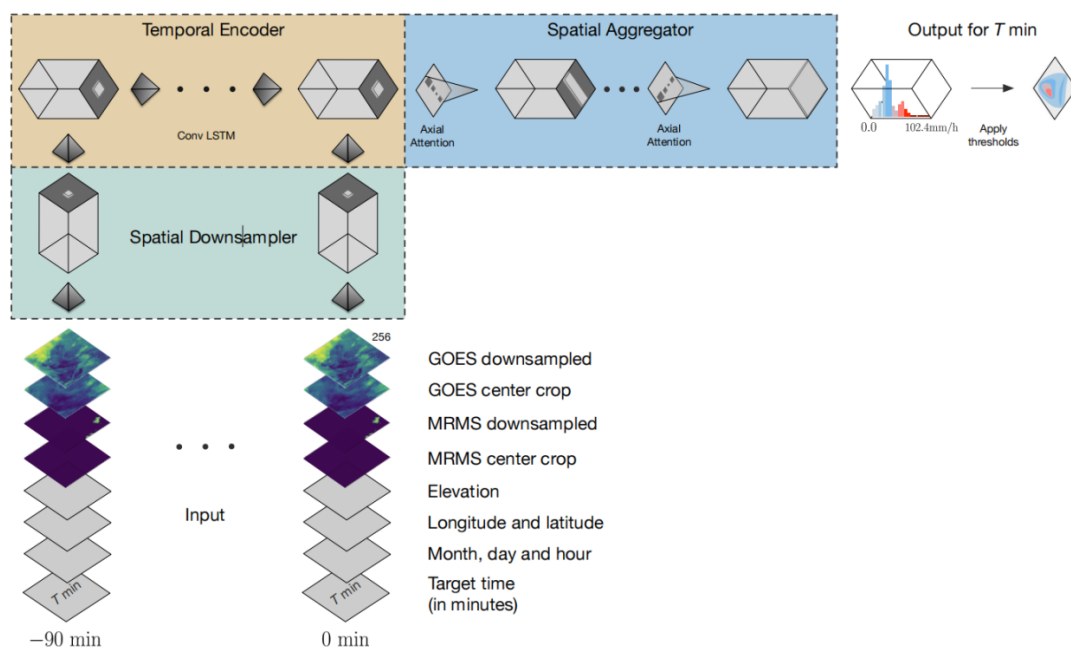


图 17. MetNet 模型的基本框架。(Sønderby , 2020)

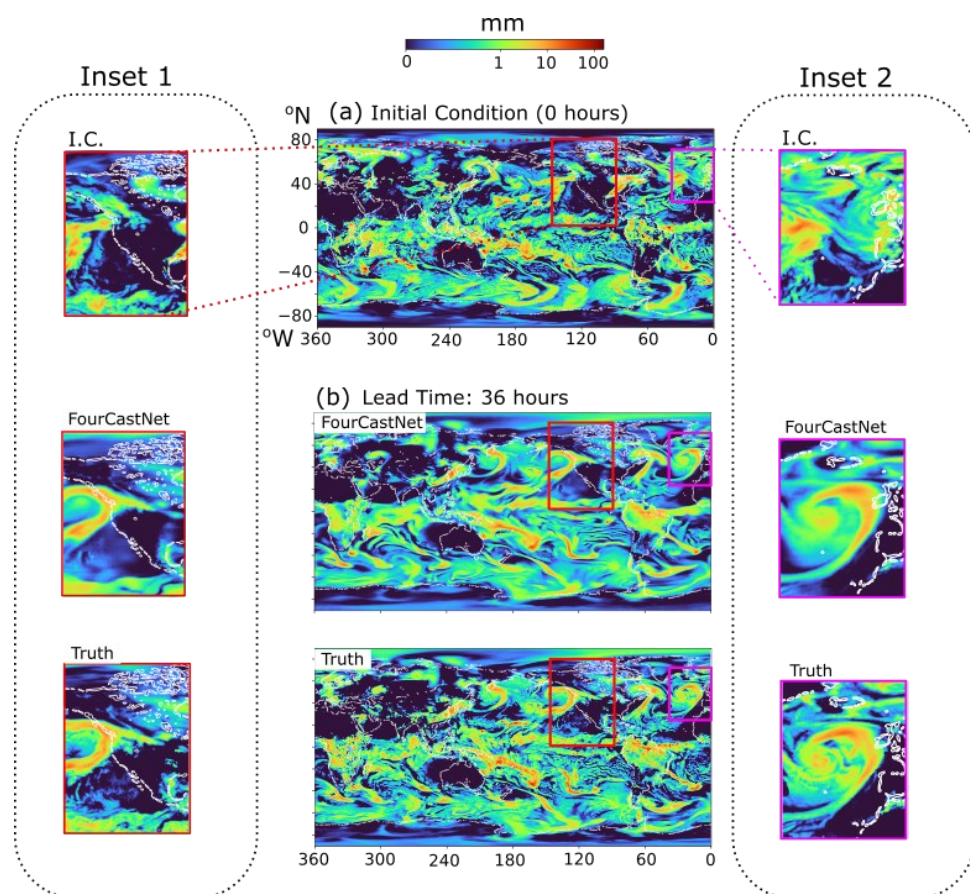


图 18. 利用 FourCastNet 模型预测的全球总降水量。(Pathak 等, 2022)

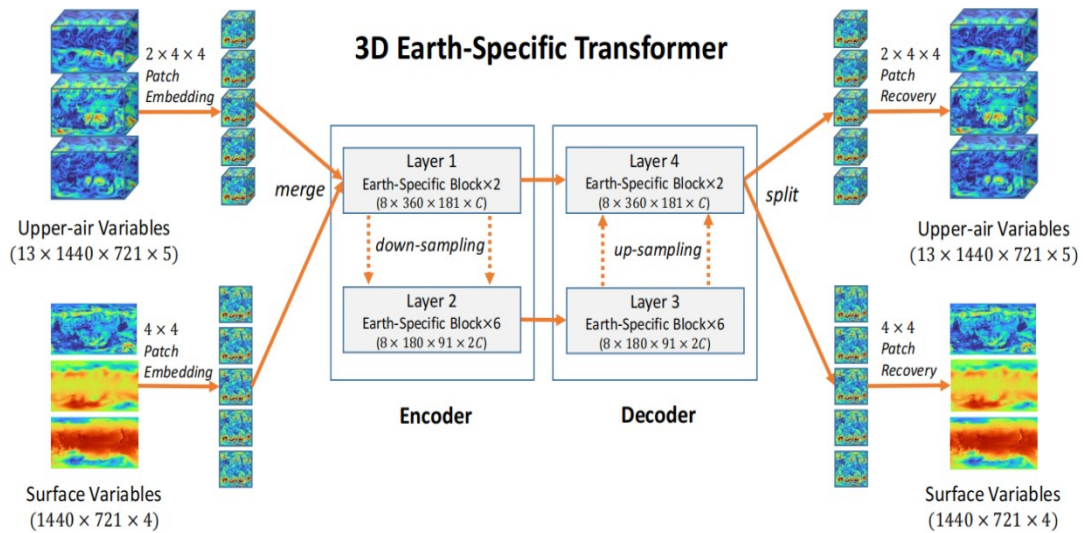


图 19. 盘古气象大模型中 3DEST 模型的网络框架。（Bi 等，2023）

2023 年 1 月，ClimaX 大模型——全球首个能够跨越天气和气候尺度的智能大模型出现（Nguyen 等，2023）。同年 4 月，上海人工智能实验室等机构发布了全球中期天气预报大模型“风乌”（Chen 等，2023a），其预报时效首次超过 10 天。随后，“伏羲”气象大模型也被发布（Chen 等，2023b），首次将大模型的预报时效扩展到 15 天。7 月，NowcastNet 大模型被提出（Zhang 等，2023），首次将降水临近预报的时效延长至 3 小时，解决了极端降水预报的短板。

与气象领域不同，海洋数据尤其是水下数据高度稀疏，大模型研究仍处于起步阶段。清华大学于 2023 年 8 月推出了 AI-GOMs 海洋大模型（Xiong 等，2023b，图 20），用于预测基本海洋变量，同时还包括区域降尺度、海浪和生物化学耦合模块。AI-GOMS 具有 $0.25^\circ \times 0.25^\circ$ 的空间分辨率和 15 个垂向层，在全球海洋基础变量的 30 天预测方面表现出卓越的性能。此外，AI-GOMS 还成功地模拟了黑潮区域 $9 \text{ km} \times 9 \text{ km}$ 空间分辨率的中尺度涡旋和热带太平洋的海洋层结。

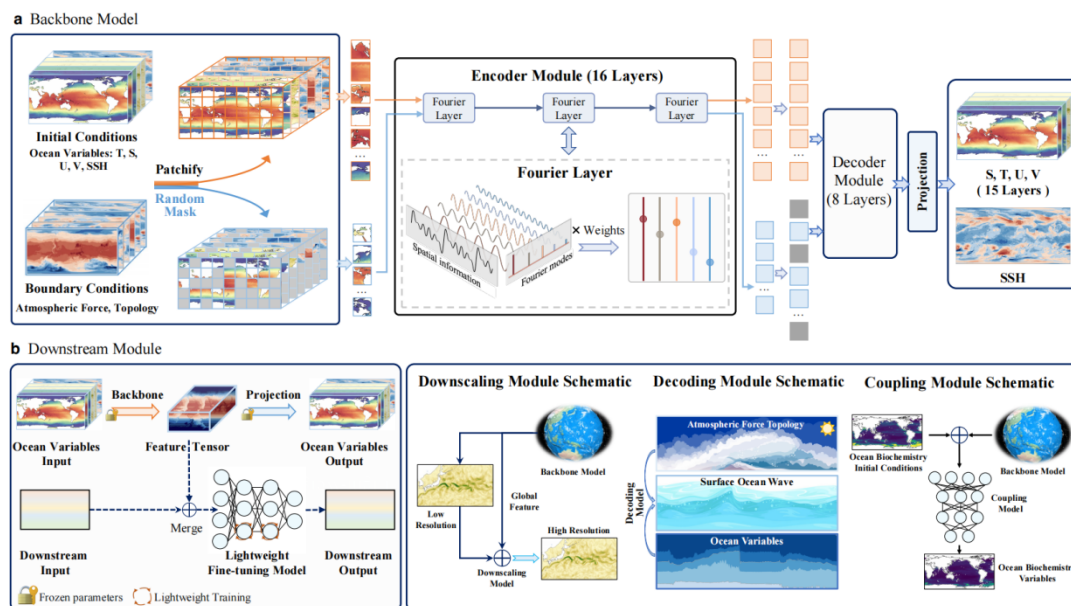


图 20. AI-GOMS 海洋大模型的框架。(Xiong 等, 2023b)

5. 海洋数字孪生系统构建

数字孪生技术早期主要应用于工业领域，通过数字化手段创建物理实体的虚拟模型，以模拟物理实体在真实环境中的行为。在国际上，数字孪生领域的研究与应用发展较早。美国宇航局（NASA）在 2011 年构建了航天器的数字孪生模型，标志着数字孪生在工业界的首次实际应用（Tuegel 等, 2011）。美国海洋大气局（NOAA）于 2022 年宣布与英伟达、洛克希德马丁公司合作，启动了“地球观测数字孪生”（Earth Observations Digital Twin）项目，旨在监测全球环境，同时针对极地、陆地、天气和海洋进行高分辨率实时模拟。

我国的海洋数字孪生技术虽然起步较晚，但目前基本与国际并跑。一些科研机构正在积极开展海洋数字孪生研究，在海洋大数据底座打造、智能海洋模式发展、海洋和气候智能预测等方面进行了一系列前期工作。清华大学深圳国际研究生院在三维地理信息与海洋信息的集成技术方面取得了突破，构建了基于 BIM 和 GIS 的数字海洋模型。崂山实验室资助了多个重要项目，包括黄海冷水团海底沉积灾害的数字孪生研究项目、基于人工智能的孪生海洋关键技术与应用示范。南方海洋科学与工程广东省实验室（珠海）与腾讯公司合作于 2023 年 9 月 8 日发布了全球首个海洋数字孪生引擎，并应用于粤港澳大湾区。

随着云计算技术的快速发展，孪生海洋数据池采用计算与存储相分离的设计，既能够满足突发性资源高峰需求，又有助于提升整体资源利用率，从而实现了海洋大数据资源的集约化管理。尽管孪生海洋系统已经在信息挖掘与预报技术方面取得了初步进展，但在基础数据的积累和及时获取、AI 算法与数值模式的结合、系统运行的效率和精度等方面仍有较大的改进空间。孪生海洋实体的基础建立在多源观测和模拟数据的集成之上，不仅需要对全球大尺度海洋动力过程的呈现，还需要实时映射与人类相关的局地中小尺度海洋与气象过程。因此，建立全球海洋观测网和精细化的区域海洋观测系统，融合各类传统数值模式和 AI 预报技术，提升服务平台和工具包的智能化、开放性和互动性，是发展海洋数字孪生系统的必须。

二、人工智能海洋学的重点方向

1. 海洋灾害的智能预警预报

海洋灾害包括动力灾害（如风暴潮、巨浪等）、生态灾害（如赤潮、绿潮及珊瑚白化等）和地质灾害（如深海泥石流、海底地震及火山喷发等），对海洋环境和人民生命财产安全有巨大威胁。近年来，基于深度学习的海洋灾害智能预警预报取得了显著进展，在一定程度上弥补了传统业务化数值预报系统在时效性和便捷性方面的不足，对于提升海洋灾害防御能力、保障沿海地区安全有重要意义。

目前的海洋灾害智能预警预报主要依靠纯数据驱动的模式，不仅缺乏科学依据，而且需要海量的训练数据。为提高预报的可靠性和可解释性，并降低对训练数据量的依赖，将物理约束引入智能模型正逐渐发展为新的研究方向。这种物理约束不仅有助于模型更好地表征海洋动力过程的复杂性，还有助于提高其预报结果的可信度。此外，因为海洋灾害预警预报通常在硬件条件相对简易的沿海海洋台站进行，必须发展轻量化的智能预报模型，从而有效应对不同地区的海洋灾害。

2. Ocean ChatGPT

GPT 是 Generative Pre-Trained Transformer 的缩写，是一项基于深度学习和 Transformer 架构的重要技术，已成为处理自然语言任务的强大工具。尽管通用的 GPT 模型具有不断增强的语言理解和生成能力，但在特定的专业领域通常表现不佳。随着简化版的开源大语言模型如 Meta LLaMa 及微调方法的问世，大语言模型正逐渐从通用向专用转变。通过特定训练，不同领域的大语言模型或基于专业知识与自然语言处理模型层出不穷，正在为各专业领域提供更精准的服务。

因此有必要发展面向海洋的大语言模型 Ocean ChatGPT。该模型应能深入理解海洋科学、海洋人文、海洋环境、海洋地理等方面的专业知识，并能提供海洋领域的知识问答、数据处理、技术支持、政策解读、旅游推荐，以及环境保护与可持续发展解决方案等服务。为提升模型的适应性和泛化能力，应广泛而深入地收集来自国际海洋科学数据中心、公开出版物、政府报告等方面的数据资料，并进行数据清洗、整理和标注预处理工作。基于 GPT 系列模型设计 Ocean ChatGPT 模型架构，

利用分布式计算资源进行模型训练，并将其部署到大型服务器或云平台提供在线服务，定期更新海洋领域的新知识和新应用，同时收集用户反馈数据，不断优化 Ocean ChatGPT 模型。

3. 海洋人工智能大模型

地球系统由大气、海洋、陆地等多个相互耦合的圈层组成，极其复杂和庞大。对这一系统的模拟和预测目前主要依靠针对各个圈层的数值模式，以及耦合所有圈层的地球系统模式。近年来，人工智能大模型在地球系统科学、尤其是大气科学领域展现出卓越的应用价值和发展潜力，成为重要的研究方向。然而，这些大模型大都是纯数据驱动，高度依赖于训练数据的数量和来源，而且在预测极端天气时因样本不均衡普遍表现欠佳。与此同时，人工智能大模型对计算资源的消耗与日俱增，如何降低成本也是未来发展所面临的严峻挑战。

对海洋领域而言，构建人工智能大模型的最大困难是缺乏可靠的训练数据，特别是长时序、高分辨率、覆盖各种海洋现象的四维观测数据。因此，必须引入物理约束，在提升可解释性的同时减少对数据量的依赖。在可预见的未来，人工智能模型将与传统数值模式以相辅相成的方式共存。人工智能可在维持数值模式动力框架不变的情况下，改进模式的参数化、资料同化和偏差订正方案，提升其计算效率。同时，可发展融合人工智能模型和动力模式的集合预报系统，提高对机理不明确过程的预报技巧。最后，建立具备物理约束（经典力学和信息理论）的海洋大模型和地球系统大模型，通过生成式人工智能，根据需求持续修正模型并输出各种形态的产品，提供多样性服务。

4. 海洋数字孪生系统

海洋数字孪生系统是认知和管理海洋的万能工具。数字孪生可高效处理海量且快速涌入的非结构海洋数据，挖掘海洋大数据的潜在信息资源，建立多维度、标准化的海洋信息数据池，并结合新型虚拟现实与人工智能技术，完成从数字表达到可视孪生的跨越，实现沉浸式、高感知度、可互动的海洋全景可视化。海洋数字孪生是海洋科学与信息技术融合并提质增效的前沿与热点方向，可为海洋信息感知、知识发现和决策支持提供重要科技基石，具有广阔的发展前景。

海洋数字孪生系统的核心是数据底座和智能技术。需要通过智能、高效和精细化的新型感知技术获取更多的海洋观测数据，需要通过物联网、区块链、数据中台等先进技术集成、共享和处理各类数据，从而夯实海洋数字孪生系统的数据底座。此外，海洋数值模式正朝着高分辨率、多运动形态耦合等方向蓬勃发展。通过先进的数据同化方法，结合高分辨率模式输出和各种观测资料，建立可靠的海洋再分析数据集，不仅可用于智能模式训练，而且能为数字孪生海洋系统提供支撑。

5. 算力优化与统筹

人工智能进步与革新的核心驱动力在于算力的不断演进与提升。根据 ChatGPT 公开的数据，该模型训练过程中消耗的算力资源高达 3640PF-days（即假设每秒计算一千万亿次，需要 3640 天）。算力资源的稀缺和高昂的计算成本已成为制约人工智能发展的关键瓶颈。海洋科学研究和应用领域已步入大数据时代，预计到 2030 年，全球海洋数据总量将达到 275PB，日增量也达到 TB 级。

为适应海洋智能大模型的发展需求，从多个维度提升算力成为当务之急。首先，需要加大海洋智能算力硬件资源的投入，建立面向海洋应用的大规模海洋智能算力中心和公共基础设施，采用先进的人工智能芯片，针对海洋特色人工智能场景，打造技术领先、生态成熟且可持续迭代升级的智能计算平台，全面支撑海洋人工智能技术的应用和创新。其次，需要优化算力调度技术，从系统架构、算法优化等方面提升算力调度能力，解决效率低下、稳定性差等问题，积极响应国家的东数西算战略，实现全国网络的算力高效调度。

三、结语

人工智能近年来取得了巨大进步，不仅在计算机科学领域欣欣向荣，更开始在各行各业绽放异彩。人工智能与海洋学结合，将开启全新的探索之旅，拓展人类对海洋和地球的认知边界。人工智能海洋学的未来发展将远超越我们的想象，将产生深刻的科学和社会影响。

人工智能将推动海洋科学和技术的发展进程。相较于传统研究方法，人工智能具备自动收集、高效处理和精准解读海洋数据的能力，能极大提升科研效率。人工智能与机器人技术的结合，正在为海洋智能感知与观测平台及其协同组网系统，带来革命性的技术变革。

人工智能将成为管理和保护海洋的重要工具。智能化的海洋监测、探测和预报系统，将为解决日益严峻的海洋污染、过度捕捞和气候变化等问题提供科技支撑，并为优化海洋新能源的开发、海洋资源的可持续利用提供智慧高效的手段，服务国家战略和社会经济发展。

人工智能海洋学的发展方兴未艾，前景光明。我们需要进一步完善这一前沿交叉学科的理论框架和内涵，需要针对海洋研究特点开发原创算法和算力资源。同时，重视过度使用人工智能技术可能带来的隐患，妥善处理隐私和数据安全问题，确保海洋数据的合法使用。

参考文献

- Arellano-Verdejo, J., Lazcano-Hernandez, H.E., & Cabanillas-Terán, N. (2018). *ERISNet: Deep learning network for Sargassum detection along the coastline of the Mexican Caribbean*, *PeerJ*, 7, 1. doi: 10.7287/peerj.preprints.27445.
- Ashkezari, M., Hill, C., Follett, C., & et al. (2016). *Oceanic eddy detection and lifetime forecast using machine learning methods*. *Geophysical Research Letters*, 43(23), 12234-212241.
- Bai, G., Wang, Z., Zhu, X., & et al. (2022). *Development of a 2-D deep learning regional wave field forecast model based on convolutional neural network and the application in South China Sea*. *Applied Ocean Research*, 118, 103012.
- Berloff, P. S. (2005). *Random-forcing model of the mesoscale oceanic eddies*. *Journal of Fluid Mechanics*, 529, 71-95.
- Bi, K., Xie, L., Zhang, H., & et al. (2023). *Accurate medium-range global weather forecasting with 3D neural networks*. *Nature*, 619, 533-538.
- Bianchi, F., Espeseth, M., & Borch, N. (2020). *Large-scale detection and categorization of oil spills from SAR images with deep learning*. *Remote Sensing*, 12, 2260.
- Bolton, T., & Zanna, L. (2019). *Applications of deep learning to ocean data inference and subgrid parameterization*. *Journal of Advances in Modeling Earth Systems*, 11(1), 376-399.
- Cachay, S.R., Erickson, E., Bucker, A.F.C., & et al. (2020). *Graph Neural Networks for Improved El Niño Forecasting*. *arXiv preprint arXiv:2012.01598*.
- Chapman, W. E., Subramanian, A. C., Delle Monache, L., & et al. (2019). *Improving atmospheric river forecasts with machine learning*. *Geophysical Research Letters*, 46(17-18), 10627-10635.
- Chen, K., Han, T., Gong, J., & et al. (2023a). *FengWu: Pushing the Skillful Global Medium-range Weather Forecast beyond 10 Days Lead*. *ArXiv Preprint ArXiv:2304.02948*.

- Chen, L., Zhong, X., Zhang, F., & et al. (2023b). *FuXi: A cascade machine learning forecasting system for 15-day global weather forecast*. *ArXiv Preprint ArXiv:2306.12873*.
- Chi, J., & Kim, H. (2017). *Prediction of Arctic sea ice concentration using a fully data driven deep neural network*. *Remote Sensing*, 9(12),1305.
- Choi, Y., Park, Y., Hwang, J., & et al. (2022). *Improving ocean forecasting using deep learning and numerical model integration*. *Journal of Marine Science and Engineering*, 10(4), 450.
- Choi, M., de Silval, W., & Yamaguchi, H. (2019). *Artificial neural network for the short-term prediction of Arctic sea ice concentration*. *Remote Sensing*, 11(9), 1071.
- Crone, T., & Tolstoy, M. (2010). *Magnitude of the 2010 Gulf of Mexico oil leak*. *Science*, 330(6004), 634-634.
- Cruz, M. A., Thompson, R. L., Sampaio, L. E., & et al. (2019). *The use of the Reynolds force vector in a physics informed machine learning approach for predictive turbulence modeling*. *Computers & Fluids*, 192, 104258.
- Deng, Z., Sun, H., Zhou, S., & et al. (2019). *Learning deep ship detector in SAR images from scratch*. *IEEE Transactions on Geoscience and Remote Sensing*, 57(6), 4021-4039.
- Dong, X., Li, Ji., Li, B., & et al. (2023). *Marine oil spill detection from low-quality SAR remote sensing images*. *Journal of Marine Science and Engineering*, 11(8), 1552.
- Duo, Z., Wang, W., & Wang, H. (2019). *Oceanic mesoscale eddy detection method based on deep learning*. *Remote Sensing*, 11(16), 1-14.
- Duraisamy, K., Iaccarino, G., & Xiao, H. (2019). *Turbulence modeling in the age of data*. *Annual review of fluid mechanics*, 51, 357-377.
- Fan, S., Xiao, N., & Dong, S. (2020). *A novel model to predict significant wave height based on long short-term memory network*. *Ocean Engineering*, 205, 107298.
- Franz, K., Roscher, R., Milioto, A., & et al. (2018). *Ocean eddy identification and tracking using neural networks*. *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*, 6887-6890.

- Gao, L., Li, X., Kong, F., & et al. (2022). *AlgaeNet: A Deep-Learning Framework to Detect Floating Green Algae From Optical and SAR Imagery. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 15, 2782-2796.
- Geng, X., Xu, L., He, X., & et al. (2021). *Graph optimization neural network with spatio-temporal correlation learning for multi-node offshore wind speed forecasting. Renewable Energy*, 180, 1014-1025.
- Gracia, S., Olivito, J., Resano, J., & et al. (2021). *Improving accuracy on wave height estimation through machine learning techniques. Ocean Engineering*, 236, 108699.
- Guo, Y., Gao, L., & Li, X. (2022). *A deep learning model for green Algae detection on SAR images. IEEE Transactions on Geoscience and Remote Sensing*, 60, 4210914.
- Gupta, M., Kodamana, H., & Sandeep, S. (2020). *Prediction of ENSO beyond spring predictability barrier using deep convolutional LSTM networks. IEEE Geoscience and Remote Sensing Letters*, 19, 1501205.
- Ham, Y.G., Kim, J.H., & Luo, J.-J. (2020). *Deep learning for multi-year ENSO forecasts. Nature*, 573, 568-572.
- Han, G., Cen, H., Jiang, J., & et al. (2022). *Applying machine learning in devising a parsimonious ocean mixing parameterization scheme. Deep Sea Research Part II: Topical Studies in Oceanography*, 203, 105163.
- Hashemi, M.R., Spaulding, M.L., Shaw, A., & et al. (2016). *An efficient artificial intelligence model for prediction of tropical storm surge. Natural Hazards*, 82(1), 471-491.
- Horvath, S., Stroeve, J., Rajagopalan, B., & et al. (2020). *A Bayesian logistic regression for probabilistic forecasts of the minimum September Arctic sea ice cover. Earth and Space Science*, 7(10), e2020EA001176.
- Jaisakthi, S. M., Mirunalini, P., & Aravindan, C. (2019). *Coral Reef Annotation and Localization using Faster R-CNN. In CLEF (Working Notes)*.
- Jang, E., Kim, Y. J., Im, J., & et al. (2021). *Improvement of SMAP sea surface salinity in river-dominated oceans using machine learning approaches. GIScience & Remote Sensing*, 58(1), 138-160.

- Jiang, G.Q., Xu, J., & Wei, J. (2018). A deep learning algorithm of neural network for the parameterization of typhoon-ocean feedback in typhoon forecast models. *Geophysical Research Letters*, 45(8), 3706-3716.
- Krestenitis, M., Orfanidis, G., Ioannidis, K., & et al. (2019). Oil spill identification from satellite images using deep neural networks. *Remote Sensing*, 11(15), 1762.
- Lam, R., Sanchez-Gonzalez, A., Willson., & et al. (2022). GraphCast: Learning skillful medium-range global weather forecasting. *ArXiv Preprint ArXiv:2212.12794*.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444.
- Lee, T. L. (2006). Neural network prediction of a storm surge. *Ocean Engineering*, 33(3/4), 483-494.
- Li, K., Ouyang, J., Yu, H., & et al. (2021). Overview of research on monitoring of marine oil spill. *IOP Conference Series: Earth and Environmental Science*, 787(1), 012078.
- Li, X., Liu, B., Zheng, G., & et al. (2020a). Deep-learning-based information mining from ocean remote-sensing imagery. *National Science Review*, 7(10), 1584-1605.
- Li, X., Zhao, Z., & Pichel, W. G. (2008). Internal solitary waves in the northwestern South China Sea inferred from satellite images. *Geophysical Research Letters*, 35, L13605.
- Li, Z., Kovachki, N.B., Azizzadenesheli, K., & et al. (2020b). Fourier neural operator for parametric partial differential equations. In: *International Conference on Learning Representations*.
- Liang, X.S., Xu, F., Rong, Y., & et al. (2021). El Niño Modoki can be mostly predicted more than 10 years ahead of time. *Scientific reports*, 11, 17860.
- Ling, J., Kurzawski, A., & Templeton, J. (2016). Reynolds averaged turbulence modelling using deep neural networks with embedded invariance. *Journal of Fluid Mechanics*, 807, 155-166.
- Liu, A., Liu, Y., Xu, K., & et al. (2024). DeepSeaNet: A Bio-Detection Network Enabling Species Identification in the Deep Sea Imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 62, 4202113.
- Liu, L., Pan, Z., & Lei, B. (2017). Learning a rotation invariant detector with rotatable bounding box. *arXiv preprint arXiv:1711.09405*.

- Liu, Q., Zhang, R., Wang, Y., & et al. (2021b). Daily prediction of the Arctic sea ice concentration using reanalysis data based on a convolutional LSTM network. *Journal of Marine Science and Engineering*, 9(3),330.
- Liu, Y., Zheng, Q., & Li, X. (2021a). Characteristics of global ocean abnormal mesoscale eddies derived from the fusion of sea surface height and temperature data by deep learning. *Geophysical Research Letters*, 48, e2021GL094772.
- Londhe, S.N., Shah, S., Dixit, P.R., & et al. (2016). A coupled numerical and artificial neural network model for improving location specific wave forecast. *Applied Ocean Research*, 59, 483-491.
- Londhe, S.N., & Panchang, V. (2006). One-Day Wave Forecasts Based on Artificial Neural Networks. *American Meteorological Society*, 23(11), 1593-1603.
- Lu, L., Jin, P., Pang, G. , & et al. (2021). Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators. *Nature in Machine Intelligence*, 3, 218-229.
- Makarynskyy, O. (2004) Improving wave predictions with artificial neural networks. *Ocean Engineering*, 31(5-6), 709-724.
- Maulik, R., & San, O. (2017). A neural network approach for the blind deconvolution of turbulent flows. *Journal of Fluid Mechanics*, 831, 151-181.
- Mayer-Schönberger, V., & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work, and think*. Houghton Mifflin Harcourt.
- Nguyen, T., Brandstetter, J., Kapoor, & et al. (2023). *ClimaX: A foundation model for weather and climate*. *ArXiv Preprint ArXiv:2301.10343*.
- Nie, Y., Liang, N., & Wang, J. (2021). Ultra-short-term wind-speed bi-forecasting system via artificial intelligence and a double-forecasting scheme. *Applied Energy*, 301, 117452.
- Niu, X., & Wang J. (2019). A combined model based on data preprocessing strategy and multi-objective optimization algorithm for short-term wind speed forecasting. *Applied Energy*, 241, 519-539.

- Ouchi, K., Tamaki, S., Yaguchi, H., & et al. (2004). Ship detection based on coherence images derived from cross correlation of multilook SAR images. *IEEE Geoscience and Remote Sensing Letters*, 1(3), 184-187.
- Pathak, J., Subramanian, S., Harrington, P., & et al. (2022). Fourcastnet: A global data-driven high-resolution weather model using adaptive fourier neural operators. *ArXiv Preprint ArXiv:2202.11214*.
- Qian, C., Huang, B., Yang, X., & et al. (2022). Data science for oceanography: from small data to big data. *Big Earth Data*, 6(2), 236-250.
- Raissi, M., Perdikaris, P., & Karniadakis, G. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378, 686-707.
- Rajasekaran, S., Gayathri, S., & Lee, T.L. (2008). Support vector regression methodology for storm surge predictions. *Ocean Engineering*, 35(16), 1578-1587.
- Rao, C., Ren, P., Wang Q., & et al. (2023). Encoding physics to learn reaction-diffusion processes. *Nature in Machine Intelligence*, 5, 765-779.
- Ruan, D., Chen, J., Jiang, J., & et al. (2022). Research on Sampling System for Abyssal Sediment Based on Image Recognition Technology. *Marine Technology Society Journal*, 56(4), 21-29.
- Santana, O., Hernández-Sosa, D., & Smith, R. (2022). Oceanic mesoscale eddy detection and convolutional neural network complexity. *International Journal of Applied Earth Observation and Geoinformation*, 113, 102973.
- Sayeed, A., Choi, Y., Jung, J., & et al. (2021). A deep convolutional neural network model for improving WRF simulations. *IEEE Transactions on Neural Networks and Learning Systems*, 34(2), 750-760.
- Simonin, D., Tatnall, A. R., & Robinson, I. S. (2009). The automated detection and recognition of internal waves. *International Journal of Remote Sensing*, 30(17), 4581-4598.

- Son, R., Ma, P. L., Wang, H., & et al. (2022). Deep learning provides substantial improvements to county-level fire weather forecasting over the western united states. *Journal of Advances in Modeling Earth Systems*, 14(10), e2022MS002995.
- Sønderby, C. K., & Espeholt. (2020). Metnet: A neural weather model for precipitation forecasting. *ArXiv Preprint ArXiv:2003.12140*.
- Sun, W., Zhou, S., Yang, J., & et al. (2023). Artificial Intelligence Forecasting of Marine Heatwaves in the South China Sea Using a Combined U-Net and ConvLSTM System. *Remote Sensing*, 15(16), 4068.
- Sun, X., Zhang, M., Dong, J., & et al. (2021). A deep framework for eddy detection and tracking from satellite sea surface height data. *IEEE Transactions on Geoscience and Remote Sensing*, 59(9), 7224-7234.
- Tracey, B.D., Duraisamy, K., & Alonso, J.J. (2015). A machine learning strategy to assist turbulence model development. In *53rd AIAA aerospace sciences meeting*.
- Tuegel, E.J., Ingraffea, A.R., Eason, T.G., & et al. (2011). Reengineering aircraft structural life prediction using a digital twin. *International Journal of Aerospace Engineering*, 2011, 154798.
- Wang, B., Wang, B., Wu, W., & et al. (2020). Sea-water-level prediction via combined wavelet decomposition, neuro-fuzzy and neural networks using SLA and wind information. *Acta Oceanologica Sinica*, 39(5), 157-167.
- Wang, H., Hu, S., & Li, X. (2023). An interpretable deep learning ENSO forecasting model. *Ocean-Land-Atmosphere Research*, 2, 0012.
- Wu, H., Hu, T., Luo, H., & et al. (2023). Solving High-Dimensional PDEs with Latent Spectral Models. *ICML 2023*.
- Xiao, C.J., Chen, N.C., Hu, C.L., & et al. (2019a). Short and mid-term sea surface temperature prediction using time-series satellite data and LSTM-AdaBoost combination approach. *Remote Sensing of Environment*, 233, 111358.
- Xiao, C.J., Chen, N.C., Hu, C.L., & et al. (2019b). A spatiotemporal deep learning model for sea surface temperature field prediction using time-series satellite data. *Environmental Modelling and Software*, 2019, 120, 104502.

Xie, W., Xu, G., Zhang, H., & et al. (2023). Developing a deep learning-based storm surge forecasting model. *Ocean Modelling*, 182, 102179.

Xiong, W., Huang X., Zhang Z., & et al. (2023a). Koopman neural operator as a mesh-free solver of non-linear partial differential equations. *ArXiv Preprint ArXiv:2301.10022*.

Xiong, W., & Xiang, Y. (2023b). AI-GOMS: Large AI-Driven Global Ocean Modeling System. *ArXiv Preprint ArXiv:2308.03152*.

Xu, G., Cheng, C., Yang, W., & et al. (2019). Oceanic eddy identification using an AI scheme. *Remote Sensing*, 11, 1349.

Xu, G., Xie, W., Dong, C., & et al. (2021). Application of Three Deep Learning Schemes Into Oceanic Eddy Detection. *Frontiers in Marine Science*, 8, 672334.

Yekeen, S., Balogun, A., & Wan Yusof, K. (2020). A novel deep learning instance segmentation model for automated marine oil spill detection. *ISPRS Journal of Photogrammetry and Remote Sensing*, 167, 190-200.

Yuan, S., Feng, X., Mu, B., & et al. (2023). A Generative Adversarial Network-based unified model integrating bias correction and downscaling for global SST. *Atmospheric and Oceanic Science Letters*, 17(4), 100407.

Zanna, L., Mana, P. P., Anstey, J., & et al. (2017). Scale-aware deterministic and stochastic parametrizations of eddy-mean flow interaction. *Ocean Modelling*, 111, 66-80.

Zhang, H., Wang, Y., Chen, D., & et al. (2022b). Temperature Forecasting Correction Based on Operational GRAPES-3km Model Using Machine Learning Methods. *Atmosphere*, 2022, 13(2), 362.

Zhang, T., Xie, F., Xue, W., & et al. (2016). Quantification and optimization of parameter uncertainty in the grid-point atmospheric model GAMIL2. *Chinese Journal Of Geophysics*, 59(2), 465-475.

Zhang, X., Wang, H., Wang, S., & et al. (2022a). Oceanic internal wave amplitude retrieval from satellite images based on a data-driven transfer learning model. *Remote Sensing of Environment*, 272, 112940.

Zhang, Y., Long, M., Chen, K., & et al. (2023b). Skilful nowcasting of extreme precipitation with NowcastNet. *Nature*, 619, 526-532.

- Zhang, Y., Shen, F., Sun, X., & et al. (2023a). Marine big data-driven ensemble learning for estimating global phytoplankton group composition over two decades (1997–2020). *Remote Sensing of Environment*, 294, 113596.
- Zhang, Z., Tian, J., Qiu, B., & et al. (2016). Observed 3D structure, generation, and dissipation of oceanic mesoscale eddies in the South China Sea. *Scientific Reports*, 6(1), 24349.
- Zhou, L., Gao, C., & Zhang, R.-H. (2023). A spatiotemporal 3D convolutional neural network model for ENSO predictions: A test case for the 2020/21 La Niña conditions. *Atmospheric and Oceanic Science Letters*, 16(4), 100330.
- Zhou, S., Bethel, B.J., Sun, W., & et al. (2021a). Improving significant wave height forecasts using a joint empirical mode decomposition–long short-term memory network. *Journal of Marine Science and Engineering*, 9, 744.
- Zhou, S., Xie, W., Lu, Y., & et al. (2021b). ConvLSTM-based wave forecasts in the South and East China Seas. *Frontiers in Marine Science*, 8, 680079.
- Zhu, C., Zhou, H., Wang, R., & et al. (2010). A novel hierarchical method of ship detection from spaceborne optical image based on shape and texture features. *IEEE transactions on geoscience and remote sensing*, 48(9), 3446–3456.
- Zhu, Y., Zhang, R. H., Moum, J. N., & et al. (2022). Physics-informed deep-learning parameterization of ocean vertical mixing improves climate simulations. *National Science Review*, 9(8), nwac044.
- 江峦, 陈路, 尤蓉蓉, 等. (2020). 一种基于深度学习的海洋沉积物测试系统及方法. 中国, 发明专利, CN110674127A, 2020-01-09.
- 刘媛媛, 张丽, 李磊, 等. (2020). 基于多变量 LSTM 神经网络模型的风暴潮临近预报. *海洋通报*, 39(6), 689–694.
- 谢文鸿, 徐广珺, 董昌明. (2022). 基于 ConvLSTM 机器学习的风暴潮漫滩预报研究. *大气科学学报*, 45(5), 674–687.