海量数据驱动下的高分辨率海洋数值模式发展与展望

宋振亚[1,2]，刘卫国[2,3]，刘 鑫[4]，苏天赟[1,2]，刘海行[1,2*]，尹训强[1,2]

（1. 自然资源部 第一海洋研究所，山东 青岛 266061；
2. 青岛海洋科学与技术试点国家实验室 区域海洋动力学与数值模拟功能实验室，山东 青岛 266061；
3. 山东大学 软件学院，山东 济南 250101；4. 无锡江南计算技术研究所，江苏 无锡 214083）

摘 要：海洋数值模式是定量描述海洋物理现象及其变化的数值模型，也是海洋与气候研究、预测的核心工具。随着海洋观测的不断投入与积累、对海洋认识的不断深入，特别是在高性能计算技术的支撑下，海洋数值模式已有了长足进步，正朝着高分辨率和多物理过程的方向发展。随着分辨率的提高、物理过程的细化，海洋数值模式的发展面临着多个方面的挑战。当前，海洋数据数量和种类不断增多，同时超级计算机、高性能计算和深度学习等技术的快速发展，为海洋数值模式的突破提供了机遇与挑战。本研究回顾了海洋数值模式的发展现状，梳理和分析了其发展中遇到的大规模高效并行计算和参数优化这两个关键问题，探讨和展望了当前海量数据驱动下海洋数值模式的发展趋势。提出计算负载均衡、计算与 I/O 重叠的并行流水线设计以及降低全局交换的算法改进是当前突破高分辨率海洋模式大规模高效并行效率的关键。从海洋科学、高性能计算以及深度学习深度交叉融合的角度，提出了实现海洋科学与深度学习相结合的 6 个途径，在此基础上，探讨了基于深度学习的参数化优化可能实现的途径。

关键词：海洋数值模式；海量数据；物理参数优化；高分辨率；深度学习；高性能计算

中图分类号：P731.3 **文献标识码：**A **文章编号：**1671-6647(2019)02-0161-10

doi：10.3969/j.issn.1671-6647.2019.02.001

引用格式：宋振亚，刘卫国，刘鑫，等. 海量数据驱动下的高分辨率海洋数值模式发展与展望[J]. 海洋科学进展，2019，37(2)：161-170. SONG Z Y，LIU W G，LIU X，et al. Research progress and perspective of the key technologies for ocean numerical model driven by the mass data[J]. Advances in Marine Science，2019，37(2)：161-170.

海洋数值模式(以下简称“海洋模式”)是定量描述海洋现象及其变化的数值模型，通过建立数学物理方程组反映海洋动力、物理以及两者相互作用等复杂过程。它将连续的海洋流体运动离散化，地球空间被划分为经向、纬向和垂向的三维网格结构，进而采用数值积分方式对偏微分方程组进行求解[1-2]。随着人们对海洋认识和理解的深入，以及计算机技术的快速发展，全球海洋模式的研制及其应用已经成为地球科学研究的重要方向之一。海洋模式不仅可以应用于海洋科学本身的研究，同时还广泛地应用于气候与环境变化的模拟和预测研究，并为海洋资源开发和国防安全等重大国家需求提供技术支撑，其发展水平是评估一个国家在海洋与气候整体研究能力的重要指标[3-4]。

收稿日期：2019-02-06

资助项目：国家自然科学基金委员会-山东省人民政府联合基金项目——海量数据驱动下的高分辨率海洋数值模式关键算法研究(U1806205)；中央级公益性科研院所基本科研业务费专项资金资助项目-束星北青年基金项目——地球系统模式 FIO-ESM V2.0 的建立及应用(2016S03)；青岛海洋科学与技术国家实验室鳌山人才计划——优秀青年学者专项(2017ASTCP-ES04)；自然资源部第一海洋研究所英特尔并行计算中心(Intel® Parallel Computing Center at FIO)项目——Large-Scale and Highly-Effective Numerical Simulation of Marine Environment with Global Surface Waves Model

作者简介：宋振亚(1982-)，男，山东聊城人，研究员，博士，主要从事海洋与气候数值模拟及高性能计算方面研究. E-mail：songroy@fio.org.cn

* **通讯作者：**刘海行(1961-)，男，山东青岛人，研究员，主要从事海洋信息技术和高性能计算方面研究. E-mail：liuhx@fio.org.cn

（李 燕 编辑）

随着海洋和气候变化研究的不断深入,海洋模式正逐步朝着更高分辨率(分辨率越高,网格越精细)、更多物理过程(方程组方程数目增加)和更快计算速度的方向发展[5]。当前,随着我国科技综合实力的增强,海洋观测资料的数量和种类不断增多,同时高性能计算机的计算速度正在由“P”级(PetaFlops)迈向“E”级(ExaFlops)[6]。人工智能特别是深度学习发展迅速[7-8],也为高分辨率海洋模式的快速发展提供了机遇,同时也提出了新的挑战。

本研究首先回顾了海洋模式的发展历史和现状,在此基础上分析了当前在海量数据驱动下的高分辨率海洋模式面临的挑战,并结合高性能计算和深度学习技术提出了可能的解决途径,从而最终提高海洋模式的模拟性能。

1 海洋模式发展历史和现状

海洋模式是建立在描述动力和物理过程的偏微分方程基础上,通过求解复杂的纳维斯托克斯(Navier-Stokes)方程对海洋系统进行建模和模拟。海洋模式从其刻画的运动形态来分,可以分为海洋环流模式、海浪模式、潮汐模式、风暴潮模式等,这些刻画不同现象的海洋模式可认为是纳维斯托克斯方程不同的简化。以下以海洋环流模式为例进行说明。

随着世界各国对海洋观测的不断投入和积累,人们对海洋内部过程的认识更加清晰,从而促进了海洋环流模式对实际物理问题的准确刻画。按照其发展历史,海洋环流模式的发展可分为早期海洋环流模式、三维海洋环流模式和海洋动力系统数值模式三个阶段。

1)第一阶段:早期海洋环流模式

早期的海洋环流数值模式是为把解析结果推广到非线性范围这个有限目的而开展起来的。该时期的海洋环流模式描述的物理过程存在较多的简化,其分辨率也非常粗。总体可分为2大类:第一类仅仅是为了阐明物理原理。海区的几何形状取得尽可能简单,并且在模式中不包括对解决问题起重要作用的一些特征。这类模式可以称为“机制性的”(Mechanistic),这一类主要是以Sakisyan的系列工作为代表的大洋环流的诊断计算方法[9-10],随后多个海洋学家发展了多种诊断计算模式并将其应用到大西洋[11-12]、太平洋[13]以及南极环流[14]等计算中。第二类是为了获得能与海洋学观测进行比较的解。一般来说,它需要大洋区域的实际几何形状,也需要精确的边界条件。这类模式可以称为“模拟性的”(Simulation),这一类主要是以Bryan的系列工作为代表的海洋环流模式[15-16],并成为以后Modular Ocean Model(MOM)[17]系列的模式基础。

2)第二阶段:三维海洋环流模式

20世纪末,随着计算机的飞速发展,海洋学家发展了一系列的海洋环流模式,如Modular Ocean Model (MOM)[17]、Parallel Ocean Program (POP)[18]、Hybrid Coordinate Ocean Mode (HYCOM)[19]、MIT General Circulation Model (MITgcm)[20]、Nucleus for European Modelling of the Ocean (NEMO)[21]、LASG/IAP Climate system Ocean Model(LICOM)[22]、Princeton Ocean Model(POM)[23]、Regional Ocean Modeling System(ROMS)[24]、Finite Volume Coastal Ocean Model(FVCOM)[25]等。这些海洋环流模式在水平网格上基本采用Arakawa B(如MOM)或C网格(如POM),通过有限插分(有限体积)格式实现实际海洋过程的数值模拟,有些海洋环流模式还附加了其他功能模块,如ROMS包括了生物学模块等。海洋环流模式的控制方程均采用波斯尼克(Boussinesq)近似,绝大多数采用准静压近似。垂向分层种类较多,如MOM采用了等深坐标、ROMS则采用了伸展坐标、HYCOM采用了混合坐标等。最近结合了拉格朗日和欧拉描述优点的任意拉格朗日-欧拉坐标(Arbitrary Lagrangian-Eulerian,ALE)逐渐兴起,其具有可加大时间步长、减少虚假的混合、直接模拟淡水通量等优势。

IPCC第三次评估报告[26]将海洋环流模式按水平分辨率的不同分为粗分辨率、中等分辨率、涡相容分辨率和涡分辨率海洋环流模式等四种。其中涡分辨率海洋环流模式当前被称为高分辨率模式,其水平分辨率一般不低于10 km。但是需要说明的是,10 km是涡分辨率模式水平分辨率的下限,仅仅能够分辨中低纬度

的中尺度涡旋，还不足以完全分辨高纬度特别是南极绕流区的中尺度涡旋特征。要真正建立一个全球高分辨率海洋环流模式，其理想的水平分辨率应该在 5 km 左右。

近年来高性能计算机的更新换代、并行算法的改进等为涡分辨率海洋数值模式的建立奠定了技术基础。在世界大洋环流实验（WOCE）计划的倡导下，国际上已经开始建立涡分辨率的海洋环流数值模式。与粗分辨率的大洋环流数值模式相比，涡分辨率模式的模拟结果与观测更相符[27-30]。因此，研发涡分辨率海洋环流模式已经成为国际上海洋环流模式和耦合气候系统模式发展的一个重要方向。

同时，受到分辨率较低和对物理过程认识不足的限制，海洋数值模式通过参数化方式将模式无法分辨的物理过程加入到内部混合中。常用的垂向混合方案包括 Mellor-Yamada 二阶半湍流封闭方案[31]、PP 方案[32]和 KPP 方案[33]等。水平混合包括类比分子混合的 Smagorinsky 方案[34]和沿等密度面的 GM 方案[35]等。这些混合方案将海洋内部的中尺度和亚中尺度过程、海浪和内波等小尺度过程对大尺度环流运动的作用进行参数化处理，但关键物理过程的参数化具有很大的不确定性，是海洋数值模式发展的核心问题。

3)第三阶段：海洋动力系统数值模式

进入 21 世纪以来，随着观测技术的提高和观测数据的日益丰富，人们对海洋中物理过程的机理有了更清晰的认识，不同物理过程之间的相互作用不断应用于高分辨率海洋环流模式中。Qiao 等建立了非破碎波浪致混合理论并应用于海洋数值模拟中，实现了海浪-潮流-环流的耦合，能够显著改进不同环流模式对上层海洋的模拟结果[36-37]。很多研究也致力于改善海洋数值模式中海洋亚中尺度过程和内波等过程的混合作用。如 Wang 等利用内潮直接模拟结果进行参数化得到混合系数，将其用于南海高分辨率区域海洋模式，模拟结果得到了较大的改善[38]。

在高分辨率海洋环流模式的基础上，通过耦合的方式充分考虑海洋多种物理过程之间的相互作用已成为当前的一个重要研究方向，海洋环流模式将逐步发展为全物理过程的海洋动力系统数值模式。总的来说，从海洋环流模式的发展来看，随着观测数据数量和种类的增加、物理过程的深入理解以及高性能计算技术的快速发展，海洋模式正朝着更高分辨率、更多物理过程、更优计算性能的方向发展。

2　高分辨率海洋模式面临的挑战及解决途径

高分辨率海洋模式的发展面临着模式框架、物理过程、模式技术、模式应用等多方面的挑战，涉及到数学、物理、高性能计算、可视化等领域，本文仅对模式技术中的高效并行算法和物理过程中的参数化方案优化开展深入讨论。

2.1　面向 E 级计算的高分辨率海洋模式的高效并行算法

建设高分辨率的海洋环流模式，模式的运行效率成为其发展和应用中的一个重要挑战。海洋模式分辨率提高会导致对计算资源的需求和计算量呈几何级数的增长。一般来说，水平分辨率提高一个数量级，计算量会增加 3 个数量级，模式输出数据（I/O）量会增加 2 个数量级。并且，随着物理过程的增多，计算量和输出数据量也会随之提高。显然这也提高了对计算量和计算速度的要求。当前超级计算机的计算速度正在由“P”级迈向“E”级，这为高分辨率模式的进一步发展提供了基础，但更关键的是需要海洋模式能够充分发挥超级计算机的计算能力[39]，因此与超级计算机相适应的海洋模式算法是关键。

Maltrud 和 Mcclean 利用 IBM Power3 计算机 500 个 CPU 核运行水平分辨率为 1/10°、垂向 40 层的全球大洋环流数值模式 1 个模式年需要 8.5 d、700 GB 的存储空间[29]。Masumoto 利用日本地球模拟器 1 500 个 CPU 核运行水平分辨率为 1/10°、垂向 54 层的准全球大洋环流数值模式 100 个模式年需要 17 d[28]。虽然近些年国内超级计算机有了快速发展，计算机系统峰值速度快速提高，如蝉联 4 次全球最快超级计算机桂冠的“神威・太湖之光”超级计算机达到 125 PFlops 计算能力，为高性能应用软件的研发提供了有力的硬件保障[40]，但是，与之相匹配的高可扩展并行应用软件的设计和开发水平还远远落后。以高分辨率全球海洋

模式为例,目前并行效率最好的海洋模式是美国的洛斯阿拉莫斯国家实验室研制的POP2,其针对海洋模式特点采用了基于网格块分区的并行剖分方案,但最高仅能扩展到万核量级,10 km水平分辨率算例的模拟速度仅能达到2~3物理年/d(无I/O)或1物理年/d(有I/O)[3]。若达到5 km水平分辨率,计算速度将变为原来的1/10,也就是说在无I/O的情况下,1 d仅能模拟0.2~0.3 a。这无论是海洋环境预报(40 min内需要完成7 d的预报)还是海洋与气候研究(通常海洋与气候科学试验需要完成10 a甚至更长的模拟),这都是无法接受的。最近,Zhao等提出了非规则并行剖分方案,可使得并行剖分达到准负载均衡[41]。应用到MASNUM海浪模式中,并行规模可达千万核,入围了2016年"戈登贝尔"奖[42]。但与海浪模式相比,海洋环流模式更加复杂,仍需要进一步研究。同时,采用异构架构是国际超级计算机系统的发展趋势,主从核计算负载均衡和合理的并行流水线计算方法是实现高效并行的一个关键[40,42]。

海洋模式中的算法也会对计算效率产生极大影响。比如海洋环流模式中的正压过程,采用预处理共轭梯度法求解线性方程组,在迭代中需要全局交换,使得通讯代价随着并行规模增大而增加,最终造成大规模并行效率低下、并行规模仅可扩展到万核量级[3],这需要对相关算法进行发展和改进[43],减少计算中的全局交换。此外,值得注意的是海洋模式的I/O问题。随着分辨率的提高,I/O量会呈现几何级增长,这会对模式运行效率产生极大影响。前期测试表明(图1),全球水平分辨率为100 km时,I/O开销不到10%;而水平分辨率提高到10 km后,海洋环流模式POP的I/O所占开销至少达到31%,而MASNUM海浪模式的I/O开销甚至超过了60%。这个问题需要从2个方面来改进,一方面采用并行I/O方式提高I/O效率;同时另一方面需要将当前海洋模式的"计算-I/O-计算-I/O"这种串行流水线方式改为计算与I/O重叠的并行流水线方式。

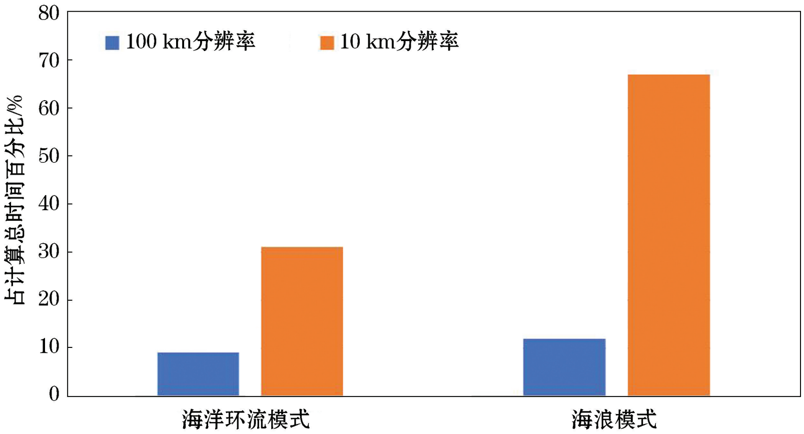

图1 不同水平分辨率海洋模式I/O所占计算总时间的比例

Fig.1 Time overhead of ocean numerical models with different horizontal resolutions

因此,有效利用新一代计算硬件,研发高效并行的海洋模式是海洋学家和计算机专家面临的共同挑战,其中高可扩展性和高效I/O是其中的关键。

2.2 基于深度学习的海洋模式物理参数优化

深度学习兴起于20世纪80年代[44-47],由于其具有灵活、多维、非线性的建模能力,越来越受到重视,已被广泛应用于多个研究领域。目前在地球系统科学的各个领域也开始了探索性的研究[48-61],但深度学习与海洋科学的融合仍是当前面临的一个挑战。机器学习算法是一类从数据中自动分析获得规律,并利用规律对未知数据进行预测的算法。深度学习则是机器学习的一个分支,其通过构建具有很多隐藏层的机器学习模型和海量的训练数据,来让机器自动从大数据中学习更有用的特征,本质上是一种统计学习方法。从这个

角度来看，机器学习（包括浅层学习和深度学习）与海洋科学的深度融合包括了新型统计预报、复杂机理探测、参数化方案优化、模式结果订正、模式过程替代和数值模式仿真六个实现途径（图 2）。当然，机器学习与海洋科学的融合途径可以是 6 个途径之一，也可以是多个途径的组合，以下仅对这 6 个实现途径简要介绍。

1）新型统计预报：与统计预报类似，只是通过机器学习替代科学家对观测数据进行统计分析，寻找模拟或预报量时空上的关联，最终建立一个统计模型，进而开展模拟或预报。基于机器学习的新型统计模型对于提高预报预测水平有所帮助，但由于海洋时空资料的不足，当前仅能针对特定的海洋现象[48]或者小范围区域[49-50]进行预测，存在较大的时空局限性，无法像海洋模式一样刻画出连续的全覆盖的整体海洋状态。

2）复杂机理探测：随着海洋数据量的不断增加，如何从海量的数据发现或分析海洋现象与规律的难度不断加大。机器学习可以通过计算机程序自动地从海量数据中寻找物理量之间的关联，特别是深度学习更擅长发现高维数据中的复杂关系，如台风路径变化与背景场之间的关系[51-52]等，这将有助于发现和认识海洋物理现象和规律，提高对海洋过程的认知，进而推动海洋模式的发展与完善。但机器学习主要关注的是两个或者多个指标之间的关联关系，不关心这些之间的内在因果关系和物理机制，通过机器学习发现的海洋物理现象和规律仍需海洋学家结合理论分析或海洋模式进一步甄别和深入分析。

3）参数化方案优化：海洋模式中存在大量的参数化方案来近似物理过程，这使得海洋模式模拟和预报结果存在一定的偏差和不确定性。参数化方案的参数调优是海洋模式发展和应用中一项必不可少的工作，但目前主要依赖于人工调试和启发式经验。机器学习可以从大量参数实验中寻找规律并预测出最优参数的取值[53-54]，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模型[55]，进而提高海洋模式的模拟和预报能力。需要指出的是，当前研究多是围绕大气环流模式[53]或台风预报模式[54]开展的参数优化，结合机器学习开展海洋模式参数化研究刚刚起步。

4）模式结果订正：当前，由于对海洋认识的不足、参数化方案不完善等原因，使得海洋模式模拟和预报结果存在一定偏差，限制了海洋模式的应用。类似于已开始探索的大气模式结果订正方法[56]，机器学习可以从海洋模式已有模拟和预报结果与观测之间的偏差中寻找偏差的规律，建立偏差预测模型，进而修正海洋模式新的模拟和预报，减少模式结果与观测之间的偏差。需要注意的是，模式结果订正能够提高海洋模式的模拟和预报能力，但只是对模式结果进行了后处理，海洋模式的发展与完善仍需海洋学家对偏差模型进行深入分析，寻找偏差来源进而改进模式过程。此外 Zhu 等[57]基于深度学习对模式结果订正的方法开展了模式同化的探索，但其工作是基于简单的洛伦兹模式开展的，对于复杂的海洋模式来说还有很长的路要走。

5）模式过程替代：海洋模式中存在某些造成计算瓶颈的物理过程，如海浪模式中的非线性波波相互作用过程、前文提到的海洋环流模式的正压过程等。尽管这些过程在模式中已被尽可能的简化，但仍然会消耗大量计算资源，而减少这些过程的计算次数又会对模拟结果的准确性产生负面的影响。机器学习，特别是深度学习能够逼近任意连续的非线性函数。已在大气模式和化学模式开展的研究工作[58-59]表明，机器学习可以通过海洋模式中这些过程的相关数据进行函数拟合，将复杂的非线性函数求解过程替换为快速的计算过程，有效降低计算开销，加速海洋模式的模拟计算。但通常海洋模式中对这些过程的结果精度要求较高，这对机器学习是一个挑战。

6）数值模式仿真：与新型统计模型类似。但相对海洋观测资料，海洋模式的模拟数据无论是在时间还是空间上都相对完备。同时从数学上理解，海洋模式可以看成由输入和输出变量构成的映射关系。因此机器学习可以通过训练来实现从海洋模式输入到海洋模式输出的一个非线性映射关系，进而实现海洋模式的快速计算和模拟，如海浪模式预报[60]。仿真模式模拟对应急预报具有重要的意义，但其由于是对海洋模式的模拟结果学习得到的模型，其模拟和预报精度很难超过海洋模式。

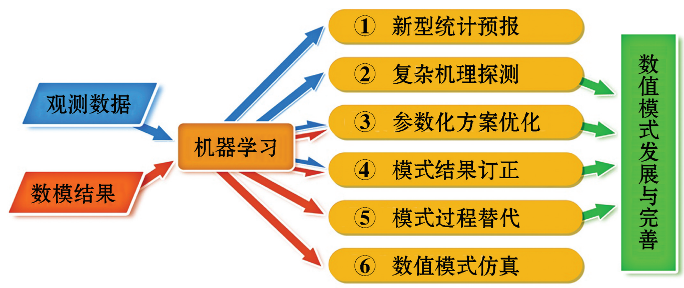

图 2　机器学习与海洋科学融合框架示意图

Fig.2　Schematic diagram of machine learning and ocean science integration

深度学习与数值模拟两者结合应是未来机器学习与海洋科学融合的发展趋势[61]，利用深度学习方法对海洋模式物理参数进行优化也许会是当前两者结合的一个突破点。参数化方案是将海洋模式中低于网格分辨率尺度的重要物理现象通过统计方法来近似得到的一种物理简化模型，其参数的选取直接影响到海洋模式的模拟与预报能力，因此参数调优是一项必不可少的工作。但目前模式物理参数调优主要依赖于人工调试和启发式经验，对于模式发展人员而言，缺乏理论支撑和有效的工程方法进行参数调优，往往使得模式调优过程成为一项复杂、繁琐、重复和低效的工作。从数学上理解，海洋模式都可看成由输入、参数和输出变量构成的映射关系，对于参数优化问题，可以借助于当前超级计算机的计算能力，选取大量不同的参数集进行模拟，通过与观测资料的对比，形成模式输入样本集和最优物理参数样本集，进而通过深度学习的方法建立模式输入与最优参数之间的关联模型，即模式物理参数估计函数(图 3)。在采用海洋模式开展模拟与预测时，可以通过模式物理参数估计函数得到可能最优的参数值，以尽可能的达到最优模拟与预测效果。

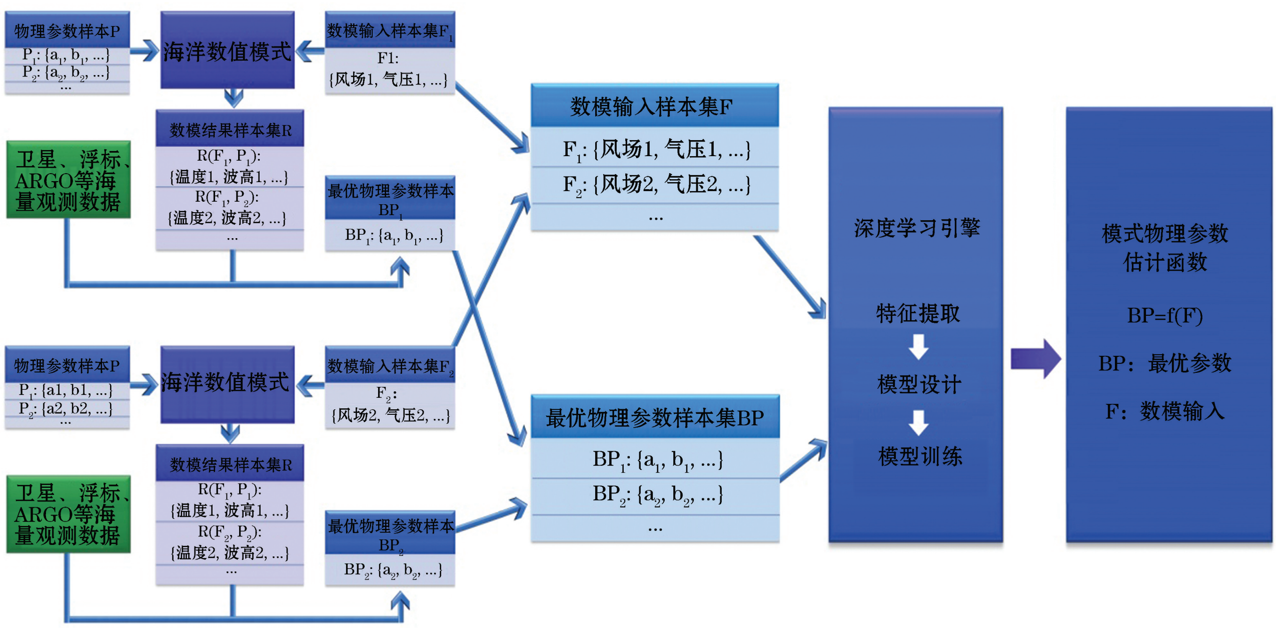

图 3　基于深度学习的参数优化方案示意图

Fig.3　Schematic diagram of parameterization optimization method based on deep learning

此外，还可以通过深度学习的方法对海洋模式中的参数化方案进行改进。一种途径是基于观测资料，采用深度学习的方法寻找一个更优的参数化方案表达，如海浪模式和部分海洋环流模式中拖曳系数的参数化方案；另外一种途径是基于更高分辨率海洋模式的模拟结果，通过深度学习得到相对低分辨率模式中的更优

的参数化方案表达，如粗分辨率海洋环流模式中的中尺度过程或涡分辨率模式中的次中尺度过程。

因此，基于大规模超算平台和海量的海洋数据，采用深度学习的方法对海洋模式中的物理参数进行合理估计，有望在提高模式模拟能力和预报精度、促进先进计算技术与海洋科学研究的交叉学科发展等方面发挥重要的作用。需要指出的是，针对海洋具有时空连续性特点的特征提取方法和深度学习神经网络模型方面，当前尚在探索，仍需要进一步完善和发展。

3　结　语

海洋模式作为海洋科学研究与预报的核心工具，正在朝着更高分辨率、更多物理过程发展。海洋模式作为重要的高性能计算程序，大规模并行计算已经成为必不可少的手段之一，新型计算方法的研究和高效并行实现技术的发展将成为推动模拟水平进步的重要因素。无论是海洋环境预报，还是气候变化的千年模拟，算得快、模拟准是其核心目标，也是衡量模式发展水平的重要指标。基于 E 级超算和海量的海洋数据，开展高效并行算法研究，提高模式计算速度；基于深度学习的新方法开展模式参数的合理估计，提高模式模拟能力，这对海洋模拟、气候变化研究以及海洋环境预报具有重要的科学意义和应用价值，同时也将推动海洋科学、高性能计算以及人工智能的深度交叉融合。

参考文献(References)：

[1] ZHANG X H, LIANG X Z. A numerical world ocean general circulation model[J]. Advances in Atmospheric Sciences, 1989, 6(1): 43-61.

[2] ZHAO Y L, ZHANG M, SI G Y. Review and perspectives of study on the thermohaline circulation[J]. Journal of PLA University of Science and Technology, 2006, 7(3): 281-290. 赵艳玲，张铭，司广宇. 海洋环流模式研究回顾与展望. 解放军理工大学学报(自然科学版)，2006，7(3)：281-290.

[3] ZHAO W, LEI X Y, CHEN D X, et al. Porting and application of global eddy-resolving parallel ocean model POP to SW supercomputer[J]. Computer Applications and Software, 2014, 31(5): 42-45. 赵伟，雷晓燕，陈德训，等. 全球涡分辨率并行海洋模式 POP 在神威蓝光上的移植和应用[J]. 计算机应用与软件，2014，31(5)：42-45.

[4] QIAO F L. Ocean models system and the surface wave-circulation coupled theory[J]. Frontier Science, 2007, 3: 81-86. 乔方利. 海洋动力系统数值模式体系及海浪-环流耦合理论[J]. 前沿科学，2007，3：81-86.

[5] SHU Q, QIAO F L, SONG Z Y, et al. A comparison of two global ocean-ice coupled models with different horizontal resolutions[J]. Acta Oceanoglogica Sinica, 2013, 32(8): 1-11.

[6] DONGARRA J, BECKMAN P, MOORE T. The International exascale software project roadmap[J]. International Journal of High Performance Computing Applications, 2011, 25(1): 3-60.

[7] YU Z L. Review of progress on artificial intelligence[J]. Journal of Nanjing University of Information Science and Technology (Natural Science Edition), 2017, 9(3): 297-304. 俞祝良. 人工智能技术发展概述[J]. 南京信息工程大学学报(自然科学版)，2017，9(3)：297-304.

[8] ZHANG R, LI W P, MO T. Review of deep learning[J]. Information and Control, 2018, 47(4): 385-397, 410. 张荣，李伟平，莫同. 深度学习研究综述[J]. 信息与控制，2018，47(4)：385-397，410.

[9] SARKISYAN A S. The role of the pure drift advection of density in the dynamics of the wind-driven currents of a baroclinic ocean[J]. Izvestiya Akademii Nauk SSSR Seriya Geologicheskaya, 1961, 9: 1396-1407.

[10] SARKISYAN A S. On the dynamics of occurrence of wind-driven currents in a baroclinic ocean[J]. Okeanologiya, 1962, 2(3): 393-409.

[11] GORMATYUK Y K, SARKISYAN A S. The results of calculations of currents in the north Atlantic according to the four-level model[J]. Izvestiva Akademii Nauk SSSR, Fizika Atmosfery i Okeana, 1965, 1(3): 313-326.

[12] HOLLAND W R, HIRSCHMAN A D. A numerical calculation of the circulation in the north Atlantic ocean[J]. Journal of Physical Oceanography, 1972, 2: 336-354.

[13] DEMIN Y L, SARKISYAN A S. Calculation of equatorial currents[J]. Journal of Marine Research, 1977, 35: 339-356.

[14] GILL A E, BRYAN K. Effects of geometry on the circulation of a three-dimensional southern-hemisphere ocean model[J]. Deep Sea Re-

search，1971，18：685-721.

［15］ BRYAN K，COX M. A numerical investigation of the oceanic general circulation[J]. Tellus，1967，19(1)：163-174.

［16］ BRYAN K. A numerical method for the study of the circulation of the world ocean[J]. Journal of Computational Physics，1969，4(3)：347-376.

［17］ GRIFFIES S M，STOUFFER R J，ADCROFT A J，et al. A historical introduction to MOM[EB/OL]. [2018-03-12]https：//www.gfdl.noaa.gov/wp-content/uploads/files/model_development/ocean/mom_history_v15.09.05.pdf.

［18］ SMITH R D，MALONE R C，DUKOWICZ J K. Parallel ocean general circulation modeling[J]. Physica D：Nonlinear Phenomena，1992，60(1-4)：38-61.

［19］ BLECK R. An oceanic general circulation model framed in hybrid isopycnic-cartesian coordinates[J]. Ocean Modelling，2002，4(1)：55-88

［20］ ADCROFT A J，HILL C N，MARSHALL J C. A new treatment of the Coriolis terms in C-Grid Models at both high and low Resolutions[J]. Monthly Weather Review，1999，127(127)：1928.

［21］ MADEC G. THE NEMO TEAM. Nemo Ocean Engine v3.4[EB/OL]. [2018-03-12] http：//www.nemo-ocean.eu/27.

［22］ LIU H L，LIN P F，YU Y Q，et al. The baseline evaluation of LASG/IAP climate system ocean model (LICOM) version 2[J]. Acta Meteorologica Sinica，2012，26(3)：318-329.

［23］ BLUMBERG A F，MELLOR G L. A description of a three-dimensional coastal ocean circulation model[C]//HEAPS N S. Three-dimensional coastal ocean models，Coastal and Estuarine Series volume 4. Washington D C，USA：American Geophysical Union (AGU)，2013：1-16.

［24］ SHCHEPETKIN A F，MCWILLIAMS J C. The regional oceanic modeling system (ROMS)：a split-explicit，free-surface，topography-following-coordinate oceanic model[J]. Ocean Modelling，2005，9(4)：347-404.

［25］ CHEN C S，LIU H，BEARDSLEY R C. An unstructured grid，finite-volume，three-dimensional，primitive equations ocean model：application to coastal ocean and estuaries[J]. Journal of Atmospheric & Oceanic Technology，2003，20(20)：159-186.

［26］ IPCC. Climate Change 2001：The scientific basis[M]. Cambridge，USA：Cambridge University Press，2001：476-479.

［27］ SMEDSTAD O，HURLBURT H，METZGER E，et al. An operational real-time eddy-resolving 1/16°global ocean nowcast/forecast system[J]. Journal of Marine Systems，2003，40-41(3)：341-361.

［28］ MASUMOTO Y. A fifty-year eddy-resolving simulation of the world ocean：preliminary outcomes of OFES (OGCM for the Earth Simulator)[J]. Journal of Earth Simulator，2004(1)：35-56.

［29］ MALTRUD M，MCCLEAN J. An eddy resolving global 1/10° ocean simulation[J]. Ocean Modelling，2005，8(1-2)：31-54.

［30］ YU Y Q，LIU H L，LIN P F. A quasi-global (1/10)° eddy-resolving ocean general circulation model and its preliminary results[J]. Chinese Science Bulletin，2012，57(3)：3908-3916.

［31］ MELLOR G L，YAMADA T. Development of a turbulence closure model for geophysical fluid problems[J]. Reviews of Geophysics，1982，20(4)：851-875.

［32］ PACANOWSKI R C，PHILANDER S G H. Parameterization of vertical mixing in numerical models of tropical oceans[J]. Journal of Physical Oceanography，1981，11(11)：1443-1451.

［33］ LARGE W G，MCWILLIAMS J C，DONEY S C. Oceanic vertical mixing：a review and a model with a nonlocal boundary layer parameterization[J]. Reviews of Geophysics，1994，32(4)：363-403.

［34］ SMAGORINSKY J. General circulation experiments with the primitive equations：I. the basic experiment[J]. Monthly Weather Review，1963，91：99-164.

［35］ GENT P R，MCWILLIAMS J C. Isopycnal mixing in ocean circulation models[J]. Journal of Physical Oceanography，1990，20(20)：150-160.

［36］ QIAO F L，YUAN Y L，YANG Y Z，et al. Wave-induced mixing in the upper ocean：distribution and application to a global ocean circulation model[J]. Geophysical Research Letters，2004，31(11)：293-317.

［37］ QIAO F L，YUAN Y L，EZER T，et al. A three-dimensional surface wave-ocean circulation coupled model and its initial testing[J]. Ocean Dynamics，2010，60(5)：1339-1355.

［38］ WANG X W，PENG S Q，LIU Z Y，et al. Tidal mixing in the south China sea：an estimate based on the internal tide energetics[J]. Journal of Physical Oceanography，2016，46(1)：107-124.

［39］ YANG X D，SONG Z Y，ZHOU S. et al. Code modernization optimization of MASNUM wave model[J]. Advances in Marine Science，2017，35(4)：473-482. 杨晓丹，宋振亚，周姗，等. MASNUM海浪模式的代码现代化优化[J]. 海洋科学进展，2017，35(4)：473-482.

［40］ FU H H，LIAO J J，YANG J Z，et al. The Sunway TaihuLight supercomputer：system and applications[J]. Science China：Information Sciences，2006，59(7)：072001.1-072001.16.

[41] ZHAO W, SONG Z Y, QIAO F L, et al. High efficient parallel numerical surface wave model based on an irregular quasi-rectangular domain decomposition scheme[J]. Science China: Earth Science, 2014, 57(8): 1869-1878.

[42] QIAO F L, ZHAO W, YIN X Q, et al. A highly effective global surface wave numerical simulation with ultra-high resolution[C]//SC16: International Conference for High Performance Computing, Networking, Storage and Analysis. Piscataway N J, USA: IEEE Press, 2016.

[43] HU Y, HUANG X M, BAKER A H, et al. Improving the scalability of the ocean barotropic solver in the community earth system model[C]//SC15: International Conference for High Performance Computing, Networking, Storage and Analysis. New York, USA: IEEE Press, 2015.

[44] KOHONEN T. A simple paradigm for the self-organized formation of structured feature maps[M]//Competition and Cooperation in Neural Nets. Springer Berlin Heidelberg, 1982.

[45] HOPFIELD J J. Neural networks and physical systems with emergent collective computational abilities[J]. Proceedings of the National Academy of Sciences, 1982, 79(8): 2554-2558.

[46] LIPPMANN R, BECKMAN P. Adaptive neural net preprocessing for signal detection in non-Gaussian noise [C]//TOURETZKY D S. Advances in Neural Information Processing Systems 1. San Francisco C A, USA: Morgan Kaufmann Publishers Inc, 1989: 124-132.

[47] LECUN Y, BENGIO Y, Hinton G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.

[48] GAO S, ZHAO P, PAN B, et al. A nowcasting model for the prediction of typhoon tracks based on a long short term memory neural network[J]. Acta Oceanologica Sinica, 2018, 37(5): 8-12.

[49] DEO M C, NAIDU C S. Real time wave forecasting using neural networks[J]. Ocean Engineering, 1998, 26(3): 191-203

[50] MANDAL S, PRABAHARAN N. Ocean wave forecasting using recurrent neural networks[J]. Ocean Engineering, 2006, 33(10): 1401-1410.

[51] ZHANG W, LEUNG Y, CHAN J C. The analysis of tropical cyclone tracks in the western north Pacific through data mining. Part I: tropical cyclone recurvature[J]. Journal of Applied Meteorology and Climatology, 2013, 52: 1394-1416.

[52] PARK M, KIM M, LEE M, et al. Detection of tropical cyclone genesis via quantitative satellite ocean surface wind pattern and intensity analyses using decision trees[J]. Remote Sensing of Environment, 2016, 183: 205-214.

[53] ZHANG T, XIE F, XUE W, et al. Quantification and optimization of parameter uncertainty in the grid-point atmospheric model GAMIL2[J]. Chinese Journal of Geophysics, 2016, 59(2): 465-475. 张涛，谢丰，薛巍，等. 格点大气环流模式 GAMIL2 参数不确定性的量化分析与优化[J]. 地球物理学报，2016，59(2)：465-475.

[54] JIANG G-Q, XU J, WEI J. A deep learning algorithm of neural network for the parameterization of tyhpoon-ocean feedback in typhoon forecast models. Geophysical Research Letters, 2018, 45(8): 3706-3716.

[55] BOLTON T, ZANNA L. Applications of deep learning to ocean data inference and subgrid parameterization[J]. Journal of Advances in Modeling Earth Systems, 2018, 11. DOI:10.1029/2018MS001472.

[56] MEN X L, JIAO R L, WANG D, et al. A temperature correction method for multi-model ensemble forecast in North China based on machine learning[J]. Climatic and Environmental Research, 2019, 24(1): 116-124. 门晓磊，焦瑞莉，王鼎，等. 基于机器学习的华北气温多模式集合预报的订正方法[J]. 气候与环境研究，2019，24(1)：116-124.

[57] ZHU J C, HU S, ARCUCCI R, et al. Model error correction in data assimilation by integrating neural networks[J]. Big Data Ming and Analytics, 2019, 2(5): 83-91.

[58] BRENOWITZ N D, BRETHERTON C S. Prognostic validation of a neural network unified physics parameterization[J]. Geophysical Research Letter, 45(12): 6289-6298.

[59] WILLIS M J, VON STOSCH M. Simultaneous parameter identification and discrimination of the nonparametric structure of hybrid semi-parametric models[J]. Computers & Chemical Engineering, 104: 366-376.

[60] JAMES S C, ZHANG Y, O'DONNCHA F. A machine learning framework to forecast wave conditions[J]. Coastal Engineering, 2018, 137: 1-10.

[61] REICHSTEIN M, CAMPS-VALLS G, STEVENS B, et al. Deep learning and process understanding for data-driven earth system science[J]. Nature, 2019, 566: 195-204.

Research Progress and Perspective of the Key Technologies for Ocean Numerical Model Driven by the Mass Data

SONG Zhen-ya[1,2], LIU Wei-guo[2,3], LIU Xin[4], SU Tian-yun[1,2], LIU Hai-xing[1,2], YIN Xun-qiang[1,2]

(1. *First Institute of Oceanography, MNR*, Qingdao 266061, China;
2. *Laboratory for Regional Oceanography and Numerical Modeling, Pilot National Laboratory for Marine Science and Technology (Qingdao)*, Qingdao 266061, China;
3. *School of Software, Shandong University*, Jinan 250101, China;
4. *Jiangnan Institute of Computing Technology*, Wuxi 214083, China)

Abstract: The ocean model is the key tool for ocean and climate research and prediction, which is a numerical model for the quantitative description of marine phenomena based on the physical laws. With the continuous marine investment, accumulation of ocean observation data, deeply understanding of the ocean process, and high-performance computing technology development, ocean numerical models have made great progress. Now, the main stream of ocean model development focuses on higher resolution and more accurate parameterization of unresolved physical processes. With the finer resolution and the more physical process, the development of ocean numerical models faces several challenges. The increasing ocean data, rapid development of high-performance computing and neural network depth learning technologies provide an opportunity for the breakthrough of ocean numerical models. This paper mainly reviewed the history and research status of the ocean numerical models, clarified and analyzed the two bottlenecks, performance of large-scale parallel computing and physical parameterization scheme, of ocean numerical models' development and applications. Then, from the perspective of marine science, high-performance computing and deep learning integration, we proposed the six approaches of deep learning and ocean science convegence, the future development and trend of ocean numerical models driven by the mass data.

Key words: ocean numerical model; mass data; physical parameter optimization; high resolution; deep learning; high-performance computing

Received: February 6, 2019