中国科学：地球科学 SCIENTIA SINICA Terrae 2026 年 第 56 卷 第 8 期： 3043–3068
论 文
《中国科学》杂志社
SCIENCE CHINA PRESS

基于窗口位移的多源数据融合方法：残差U-Net注意力网络用于海底地形背景场超分辨率优化

柴星宇[1,2]，冷洪泽[2*]，宋君强[1,2*]，肖培坤[2]，鲍森亮[2]，赵军[2]，李金才[2]，邓科峰[2]，杨锦辉[2]

1. 国防科技大学计算机学院，长沙 410073
2. 国防科技大学气象海洋学院，长沙 410073
* 通讯作者，E-mail：hzleng@nudt.edu.cn(冷洪泽)，songjunqiang@nudt.edu.cn(宋君强)

2025-10-11收稿，2026-05-09修回，2026-05-25接受，2026-07-20网络版发表
国家自然科学基金青年科学基金项目(42406195)资助

摘要 海底地形高分辨率重建对海洋科学与工程具有重要意义。针对现有以数字水深模型(Digital Bathymetric Model)作为单一标签的超分辨率研究在真实多波束数据验证上存在局限，且DBM数据分辨率与精度有待提高的问题。本文提出一种基于窗口错位多源数据融合的海底地形背景场超分辨率优化方案，构建了残差U-Net注意力模型。该模型结合卫星测高重力场参量，在西沙群岛与夏威夷-皇帝海山链两个研究区域对1′分辨率的GEBCO背景场进行实验；通过动态滑窗与位置编码进行局部特征融合，实现30″与15″两档分辨率的背景场优化。与未参与训练的高分辨率多波束真值对比表明，残差U-Net注意力模型的输出误差显著低于原始GEBCO背景场及主流基线模型，在15″任务中，西沙与夏威夷-皇帝海山链区域的均方根误差(RMSE)较背景场分别大幅下降约40.3%与60.3%，并在0.1–1 cycles km^{-1}频段内表现出更优的功率谱拟合特征。此外，针对超分辨率优化能力对比了8种基础神经网络结构，结果证实：多尺度上下文模块与通道注意力机制能有效改善全局趋势与局部细节的并行建模，并显著抑制多源数据噪声。本研究在提升海底地形重建精度与多分辨率泛化稳定性方面具有实用价值，为多源物理观测驱动的海底地形重建提供了可行路径。

关键词 海底地形超分辨率构建，注意力机制，多源数据融合，卫星测高重力参量

1 引言

对海底地形进行超分辨率重建是深海探测与测绘领域长期面临的挑战，具有重要的科学意义与显著的社会经济价值。精确的海底地形数据不仅有助于地质学家研究板块构造与海洋盆地演化，还能为海山喷发、海啸等海洋地质灾害的风险评估与预警提供关键支撑(Mayer等，2018)；此外，海底地形作为海洋环流的边界条件，对海洋流体数值模拟与预报结果亦有重要影响(Jayne等，2004)。尽管船载多波束测深是获取高精度海底地形的黄金标准，但其覆盖范围有限、获取成本高昂(Smith，1993；Wölfl等，2019)——迄今船载观测仅完成了全球约25%海域的高分辨率测绘(Yu等，2024)，剩余大面积区域须仰赖间接反演与插值方法。

早期基于重力异常与位势场反演理论的研究为后续海底地形反演奠定了方法基础(Parker，1973)。随着卫星测高技术的发展，学者利用高精度卫星重力数据开展了大尺度反演，为地形刻画提供了重要支撑(Smith和Sandwell，1994；Sandwell和Smith，1997；Andersen和Knudsen，1998；Hu等，2015；Sandwell等，2022)。在此基础上，融合多源数据构建海底地形数字测深模型(DBM)成为了主流，涌现出ETOPO、DTU、S&S、STO_IEU及GEBCO等具有广泛知名度的数据产品(Amante和Eakins，2009；Andersen和Knudsen，2008；Tozer等，2019；Fan等，2022；GEBCO Compilation Group，2023)。目前，利用重力场反演地形的方法主要包括Nettleton法(Nettle-

中文引用格式：柴星宇, 冷洪泽, 宋君强, 肖培坤, 鲍森亮, 赵军, 李金才, 邓科峰, 杨锦辉. 2026. 基于窗口位移的多源数据融合方法：残差U-Net注意力网络用于海底地形背景场超分辨率优化. 中国科学: 地球科学, 56(8): 3043–3068, doi: 10.1360/SSTe-2025-0345
英文引用格式：Chai X, Leng H, Song J, Xiao P, Bao S, Zhao J, Li J, Deng K, Yang J. 2026. Window-shifted multi-source data fusion method: Residual U-Net attention network for seafloor topography background field super-resolution optimization. Science China Earth Sciences, 69(8): 2998–3023, https://doi.org/10.1007/s11430-025-1972-1

© 2026《中国科学》杂志社

ton，1939)、重力地质法(GGM)(Ibrahim和Hinze，1972)以及近些年兴起的利用海洋重力场等参量进行端到端预测的深度学习反演模型(Yang等，2023；Wan等，2023；Zhou等，2023，2024，2025；Harper和Sandwell，2024；Ge等，2025)。然而，无论是传统解析方法还是神经网络反演，均存在固有的物理分辨率瓶颈。受重力向上延拓效应(Upward Continuation Effect)这一低通滤波特性的限制，反演的海底地形分辨率通常被局限在1′左右(Smith和Sandwell，1994；Harper和Sandwell，2024)。现有的高分辨率全球模型(如15″GEBCO)主要依靠叠加稀疏的多波束实测数据才得以实现，鉴于多波束覆盖严重不足，采用超分辨率技术突破反演物理极限、实现高分辨率海底地形重建显得尤为必要。

近年来，研究者广泛尝试以数据驱动的超分技术改进DBM分辨率。典型的图像超分技术(如SRCNN、SRResNet、EDSR)已被用作基线模型，将低分辨率格网转换为更高分辨率的DBM(如ESRGAN在局部域的尝试)(Ledig等，2017；Lim等，2017；Sonogashira等，2020)；也有研究结合可变形卷积与Transformer构建STFET模型(Cai等，2023)，或将陆地高程超分技术迁移至海底(Zhang等，2022)，以及采用图卷积网络对点云数据进行上采样(Irisawa和Iiyama，2024)。在网络优化方面，研究者逐步引入了通道注意力机制(Hu等，2018)、多尺度上下文聚合(ASPP)(Chen等，2017)、膨胀卷积(Dilated U-Net)(Yu和Koltun，2016)等模块以提升高频细节的恢复能力。然而，现有的海底DBM超分研究存在明显不足：一是绝大多数工作侧重以DBM格网本身作为训练与评估标签，未采用独立且未参与训练的船载多波束实测数据作为真实参考以评估实际应用价值；二是这些研究大多停留在纯图像处理的视角，未探讨如何将外部多源物理观测数据融入超分框架以约束重建过程。

基于以上研究现状，为了更充分地结合多源重力参量数据对DBM数据进行优化，本文将海底地形超分任务转化为对DBM数据的背景场超分辨率优化任务，提出并验证了一种基于窗口位移的多源数据融合残差U-Net注意力模型。我们将最新的卫星测高重力参量数据(包括由SWOT提供并经处理得到的高精度重力场参量；Yu等，2024)与作为待优化的海底地形低分辨率背景场(下采样到1′分辨率的GEBCO)共同输入网络，以30″与15″作为目标分辨率进行背景场超分辨率优化，并在未参与训练的高分辨率多波束实测点上进行独立验证并与常见深度学习模型进行对比。在模型设计方面，我们在统一流程下比较分析了SRResNet、EDSR、残差U-Net注意力模型及若干注意力/多尺度变体的性能，同时系统评估了ASPP、膨胀卷积等模块的效用。为证明所提出的多源融合超分模型的背景场优化精度显著优于原始GEBCO，且在物理逻辑上更接近船测精度，本研究的所有实验均在西沙海域与夏威夷-皇帝海山链两个地形差异显著的海域展开。本研究不仅为稀疏测深条件下的高分辨率海底数据库构建提供了技术路径，也为将物理观测与数据驱动超分方法相结合提供了实践依据。

2　数据和方法

2.1　数据

2.1.1　海底地形格网数据产品

本文采用GEBCO2024海底地形网格数据作为模型优化的低分辨率背景场数据。全球地形数据集(GEBCO，General Bathymetric Chart of the Oceans)是一个由国际海道测量组织(IHO)和联合国教科文组织政府间海洋学委员会(IOC)联合支持的项目，旨在提供并持续更新全球海底地形数据。其当前分辨率为15″，是目前该领域内具有极高权威性和广泛应用度的公开全球海底地形数据集。为了充分发挥多源数据融合深度学习模型对现有地形产品的物理规律约束与误差订正能力，将广泛使用的GEBCO2024作为待优化的背景场，高度契合本研究“背景场超分辨率优化”的核心目标。该数据可从https://www.gebco.net/data-products/gridded-bathymetry-data/获取。

2.1.2　海洋重力场参量数据

本文引入了由加利福尼亚大学斯克里普斯海洋研究所(SIO)团队公开的、基于最新SWOT卫星测高数据计算的高精度重力衍生产品。该数据集包括重力异常、海洋重力垂直梯度以及海洋垂线偏差的北向和东向分量。基于SWOT数据计算的重力异常网格(分辨率为1′)已被证实能够更精细地反映重力与海底地形小尺度特征之间的关联(Yu等，2024)。此外，分辨率同为1′的垂直重力梯度数据能有效凸显地形的高频细节；而垂线偏差则能够反映地下物质密度的空间变化，进而揭示深部地壳构造活动(Sandwell等，2014)。将上述多源重力参量作为物理约束条件输入模型，可为网络捕捉细微的海底地质特征提供丰富的物理先验信息。以上重力参量数据可从https://topex.ucsd.edu/pub/archive/grav_SWOT/获取。

2.1.3　船载多波束探测水深数据

本研究采用的作为真实参考的多波束测水深数据来自两类独立渠道。其一为合作单位——广州海洋地质调查局提供的西沙群岛海域高密度多波束测深资料。该区

域为本研究的“西沙”试验区，原始采样间隔约为300 m。为便于与其他栅格数据对齐并用于模型构建，我们首先将这些观测数据插值到15″(约500 m，取决于具体纬度)的分辨率网格上。在模型训练阶段，我们采取了稀疏采样策略，即在15″网格的基础上以1′为间隔进行空间采样，提取出的离散点作为训练的监督标签；而15″及30″分辨率下的密集点云仅在模型评估阶段作为独立的测试参考。其二为NOAA/NCEI发布(2024年7月版本)的皇帝海山链区域多波束Mosaic数据集，官方已将该区域历年航次资料插值到规则格网，并对重叠覆盖区域取历史观测均值以削弱单次测量异常(可从https://www.ncei.noaa.gov/products/seafloor-mapping获取)。同理，我们将该Mosaic数据插值至15″网格，以1′间隔采样的点作为训练标签，高分辨率点留作评估使用。

两片海域在平均深度、方差与地形形态上差异显著，具备明显的空间异质性，为评估模型在不同地形条件下的特征提取与泛化能力提供了理想的实证场景(图1展示了数据分布，表1列出了基本统计参数)。具体而言，西沙群岛(16°–18°N，110.5°–112.5°E)水深范围从近岸浅水延伸至2604.28 m，标准差约486.07 m，反映了珊瑚平台与深海盆地并存的地形特征(Ma等，2011)。考虑到该区域数据标准差相对较低，且1′分辨率的有效训练点数较为有限(仅14400个采样点，作为对比，其15″分辨率下的点数达230400个)，为确保模型训练阶段有充足的地形特征样本，我们未对该区域数据进行离群值剔除。皇帝海山链(30°–50°N，165°–175°E)则整体坐落于广阔深海盆地之上，平均深度达7710.41 m，标准差高达1155.26 m，并散布着若干浅水海山(Watts，1978)。针对该海域标准差极大且极易出现极端观测值的特性，我们在数据入库前专门对其插值后的栅格数据按行进行了基于3σ原则的检验，共剔除10585个异常数据点(占该区总有效数据的0.74%)，从而有效提升了作为Ground Truth的标签质量。

2.1.4 数据集构建

海底地形构建与常规视觉图像重建在本质上存在两点重要差异，因此对数据预处理有特殊要求。其一，海底地形数据的数值方差与值域往往很大，局部低幅值信息在全局归一化过程中可能被平滑掩盖，从而使模型难以学习到微小但实际重要的地形细节；其二，海底地形的分界通常并非清晰断裂，而是以梯度渐变的形式存在，表面上看似不同的色块之间实际具有连续的起伏变化，二维场到二维场的映射若不充分考虑这种渐变特性，难以捕捉这些细微但关键的变化。基于上述考虑，本研究在预处理阶段采用“场到点”的高精度局部采样策略：以每个待预测目标点为中心，提取其周围11×11的格点窗口，将该窗口内所有通道的原始数值作为该目标点的输入特征，从而在样本级别保留丰富的局部空间相关信息以增强模型对细节的敏感性。选取11×11的窗口是在算力需求和空间特征提取能力之间的平衡之选，窗口太小不能充分地提取地形的大尺度变化趋势从而会影响模型精度，窗口太大会导致所需算力需求的显存显著变高，超出了我们现有的算力条件。

我们在低分辨率的GEBCO背景场之外引入了四类重力相关特征作为辅助通道：重力异常、垂直重力梯度、垂线偏差北分量与垂线偏差东分量。最终的数据样本组织形式为：5通道的11×11张量映射到中心点的水深标量(即5通道11×11→1的映射关系)。在从窗口中提取周边格点时不可避免地遇到边界越界或缺失点，此类空值会导致输入维度不一致或信息缺失。对此，我们采用了通用的填充策略：对每个通道用该通道的全局均值对窗口边缘或缺失位置进行填充，保证每个样本输入的张量尺寸恒定并尽量维持该通道的统计特性。

在训练与评估的标签分配及掩码(Mask)设计上，为确保评估结果的客观性与实际应用价值，本研究始终将插值至高分辨率格网的多波束实测水深作为真实值(Ground Truth)参照。具体实现策略如下：在训练阶段，模型仅使用位于1′分辨率格点上且有实际多波束覆盖的数据点作为监督标签；所有留存于高分辨率(30″与15″)格网上的多波束数据则被严密隔离，作为独立的评估基准。在评估环节，我们对预测结果施加掩码处理，严格限制仅在有真实多波束观测覆盖的有效点上计算误差评估指标，以此杜绝因插值平滑或外推估计产生的“虚假标签”对模型性能评估的干扰(各分辨率下的有效观测点数量分布详见表2)。需要特别说明的是，后续用于30″和15″评估的测试点集已严格剔除了所有参与过1′训练的坐标点位。虽然测试集与训练集均源自同一海区、本质上属于具有高度空间自相关性的连续地理场，并非绝对的“空间独立”，但本研究通过严密的点位互斥(Point-wise Mutually Exclusive)划分机制，确保测试集中的高频格点在训练阶段对模型而言是“不可见”的未知区域。这种基于未观测格点的评估策略，真实还原了利用低分辨率稀疏观测推演高分辨率未知地形的实际工程挑战，从而能够科学、有效地检验模型在网格亚像素级(Sub-grid)的超分辨率重建能力与泛化推演水平。

上述预处理与数据集构建流程，既通过局部窗口特征保证了多源物理信息的丰富性与空间连续性，又通过严谨的点位互斥划分与掩码策略建立了客观可信的测试基准，为后续的背景场超分辨率优化与模型性能验证奠

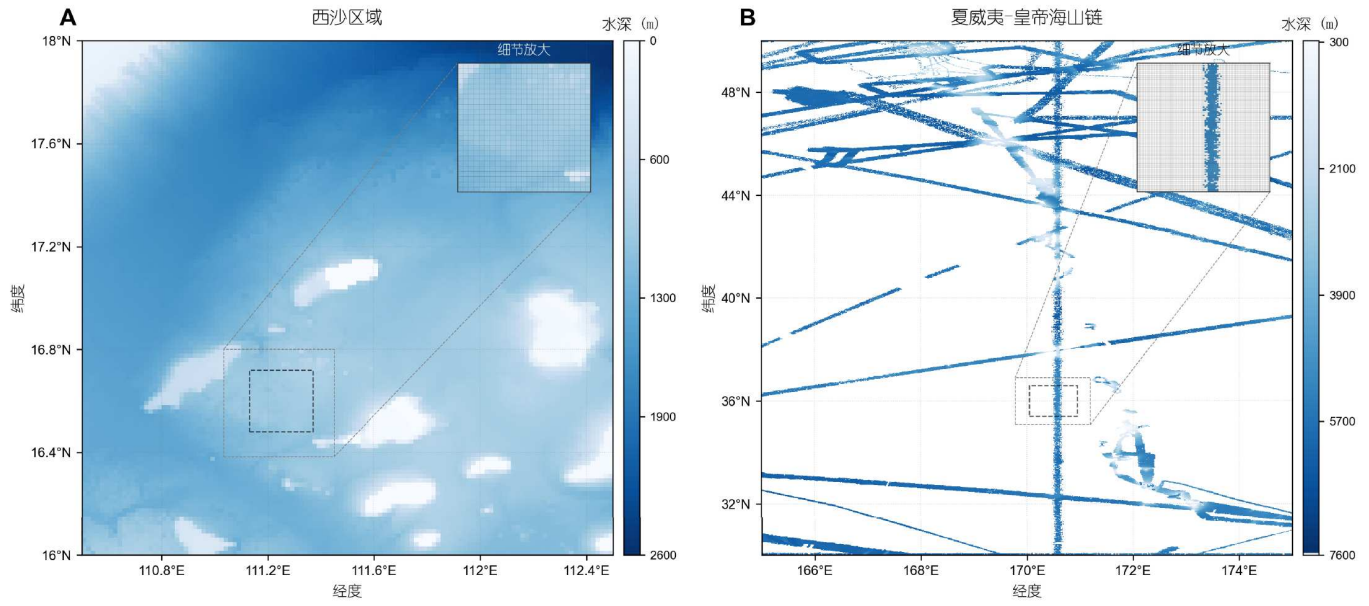


图 1　实验海域多波束测深数据分布及地形渲染图。（A）西沙海域；（B）夏威夷-皇帝海山链区域。图中颜色渐变代表实际水深(m)，虚线框及其连接的实线框展示了对应区域内海底地形的局部细节放大效果

表 1　实验海域数据统计分析表

区域名	经度范围	纬度范围	深度平均值(m)	深度标准差(m)	深度最小值(m)	深度最大值(m)
西沙群岛	110.5°–112.5°E	16°–18°N	1224.13	486.07	0	2604.28
夏威夷-皇帝海山链	165°–175°E	30°–50°N	5218.95	1155.26	279.18	7710.41

表 2　不同海域的不同分辨率有效点数量

区域名	1′分辨率点数(个)	30″分辨率点数(个)	15″分辨率点数(个)
西沙群岛	14400	57600	230400
皇帝海山链	89555	357151	1427963

定了坚实且符合地理学逻辑的数据基础。

在模型训练之前，对所有数据都进行了z-score归一化操作，以确保神经网络能够稳定训练。数据归一化的公式如下所示：

$$\mathbf{X}^{*}=\frac{\mathbf{X}-\mu}{\sigma} \tag{1}$$

其中，$\mathbf{X}$可以是任意数据通道的矩阵；μ表示此类数据的平均值；σ表示此类数据的标准差。

2.2　方法

2.2.1　残差U-Net注意力模型

为了更高效精准地捕捉输入数据的空间特征，本文为海底地形背景场超分辨率优化任务构建了加入位置编码的残差U-Net注意力模型。输入五种多源数据的过程中额外加入了位置编码。模型采用三层残差卷积块(He等，2016)(输出通道为[64,128,256])进行特征提取，每层包含两次卷积、批归一化和ReLU激活(Nair和Hinton，2010)。通过最大池化层进行下采样，模型在瓶颈层应用多头注意力机制(头数为8)(Vaswani等，2017)，以增强特征间的依赖建模。多头注意力的计算公式如下所示：

$$\text{Multihead Attention}(\mathbf{Q},\mathbf{K},\mathbf{V})=\text{Concat}\left(\text{head}_1,\cdots,\text{head}_8\right) \tag{2}$$

其中每个头的计算公式为

$$\begin{aligned}\text{head}_i&=\text{Attention}\left(\mathbf{Q}W_i^{\mathbf{Q}},\mathbf{K}W_i^{\mathbf{K}},\mathbf{V}W_i^{\mathbf{V}}\right)\\&=\text{softmax}\left(\frac{\left(\mathbf{Q}W_i^{\mathbf{Q}}\right)\left(\mathbf{K}W_i^{\mathbf{K}}\right)^{T}}{\sqrt{d_k}}\right)\end{aligned} \tag{3}$$

其中，$\mathbf{Q}$、$\mathbf{K}$和$\mathbf{V}$分别是查询、键和数据矩阵；i对应头数的编号；d_k表示键向量的维度；在模型的运转过程中，

先将特征展平后才能输入到多头注意力计算模块进行注意力计算，之后再重塑回展平之前的数据规格。

解码器通过上采样和跳跃连接逐步恢复特征，最终通过1×1卷积输出水深预测值。此模型利用U-Net(Ronneberger等，2015)进行多尺度特征提取，加入残差连接保证训练稳定性的同时也有着多头注意力机制带来的特征融合能力，专门为捕捉输入数据的复杂空间特征而设计。残差U-Net注意力模型的详细结构如图2所示。

2.2.2 窗口位移的多源数据融合地形背景场超分辨率优化方案

为了将多通道重力参量的信息增益融入背景场进而对低分辨率的背景场进行高分辨率高精度的优化，本文首次提出窗口位移的多源数据融合地形背景场超分辨率优化方案。鉴于当前可用于多源数据融合的SWOT衍生重力数据(重力异常、垂直重力梯度、垂线偏差的北/东分量)均为1′分辨率，通常只能在1′有效观测点上训练出可靠的重建模型。因此，我们以全部1′分辨率的有效观测点作为训练与验证集(80%用于训练，20%用于验证)，并将未参与训练的高分辨率多波束观测点保留用于独立测试评估。模型训练完成后，在超分辨率推理阶段采用滑动窗口错位采样策略：对于30″超分任务，窗口每次沿网格移动30″；对于15″超分任务，窗口移动步长为15″。在每一位置，模型以窗口覆盖范围内可获得的所有观测点为输入，同时记录目标采样点相对于窗口中心的相对偏移作为位置编码(Positional Encoding)。该位置编码连同窗口内多源特征一并输入网络，使模型在窗口平移过程中能够准确感知目标点的相对位置并据此调整预测。

背景场超分优化过程中常见的问题是不同分辨率网格间的错位与数据缺失：由低分辨率(如1′)提取的窗口内观测点往往无法严格对齐到预定义的高分格点，导致某些窗口出现缺项或有效观测点数低于期望。为保证每次输入到网络的特征向量尺寸一致，我们采取两项工程策略：其一，采用边界填充(以该通道的均值填充)保证每个滑窗始终为固定大小(例如11×11)；其二，将上述位置编码显式并入模型输入，从而使模型在面对由错位与缺

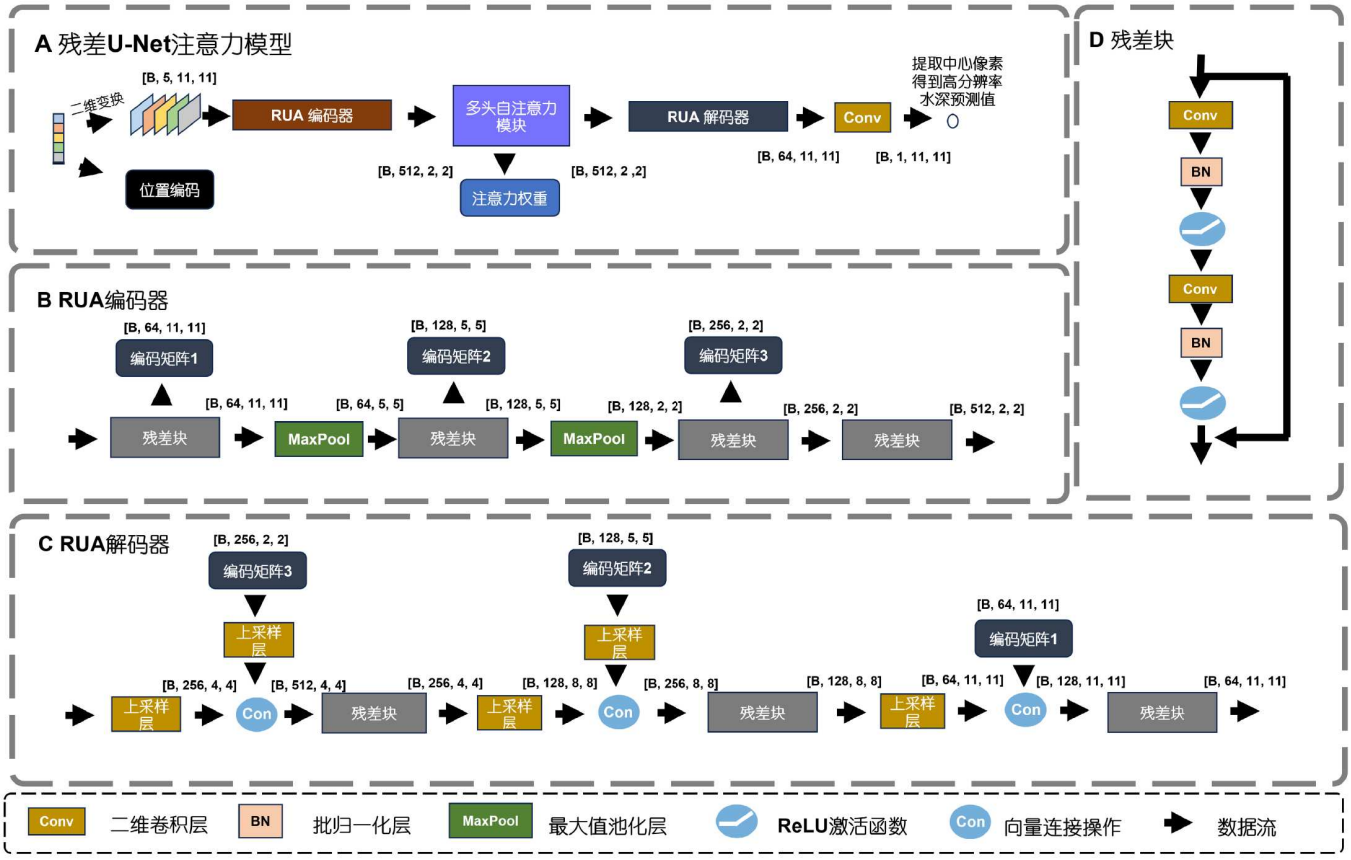


图2 残差U-Net注意力模型(Residual U-Net Attention Model)网络结构图。(A) 模型的整体框架，展示了输入数据经过编码器、多头自注意力(MHSA)模块到解码器的完整特征提取与预测流程；(B) RUA编码器(RUA Encoder)模块的内部结构；(C) RUA解码器(RUA Decoder)模块的内部结构；(D) 残差块(Residual Blocks)的具体卷积层与激活函数组合

测引起的观测不完整性时，仍能基于相对位置与局部多源信息做出准确估计。位置编码的定义如公式(4)所示。

$$\text{pos_encoding}=\left[\frac{x_{\text{pos}}-i_{\text{center}}}{\text{half_window}},\frac{y_{\text{pos}}-j_{\text{center}}}{\text{half_window}}\right] \tag{4}$$

其中，$x_{\text{pos}}=\frac{i}{\text{scale_factor}}$，$y_{\text{pos}}=\frac{j}{\text{scale_factor}}$表示了每个高分辨率预测点$(i,\ j)$在低分辨率网格上的投影位置$(x_{\text{pos}},\ y_{\text{pos}})$。$i_{\text{center}}=\text{round}(x_{\text{pos}})$，$j_{\text{center}}=\text{round}(y_{\text{pos}})$表示了使用round函数取整函数得到的窗口中心坐标$(i_{\text{center}},\ j_{\text{center}})$。

该设计使得模型既能在1′实测点上学习到稳健的观测-地形映射，又能通过滑动窗口与位置编码将训练所得信息推广到更高分辨率网格的预测任务，从而在多源信息受限且采样不规则的海域环境中实现更高精度的背景场超分辨率优化。图3直观展示了窗口位移背景场超分辨率优化方案的流程。

针对深度学习模型在推理过程中可能产生的非物理正值(即错误地将海域预测为陆地)问题，本研究设计了一种基于先验背景场的物理约束策略。具体而言，我们在推断阶段对模型输出实施了严格的“零值约束”：对于任意预测水深大于0 m的异常点，优先采用同位置经双三次插值(Bicubic Interpolation)的GEBCO背景场数值进行替换。该策略在消除模型非线性映射产生的非物理噪点的同时，有效保证了重构地形在物理意义上的合理性与一致性。

在模型效果评估阶段，本研究将模型输出的30″与15″背景场超分优化结果分别与未参与训练的高分辨率多波束观测(插值到目标格网)以及GEBCO数据进行对比。为保证评估的客观性，误差统计仅对多波束观测覆盖的有效点计算，并在后续分析中分别报告非训练点集的RMSE、MAE等指标，以全面反映模型在不同地形类型与空间采样条件下的精度表现。

2.2.3　损失函数

所有模型均采用均方误差(MSE)损失函数，损失函数定义如公式(5)所示。

$$\text{MSE_Loss}=\sqrt{\frac{1}{N}(\text{Pred-Truth})^2} \tag{5}$$

其中，N为有效点个数；Pred为模型输出的预测水深值；Truth为对应区域视为真实值的多波束船测水深数据。

对于基础通道注意力模型和残差U-Net注意力模型，为了正则化模型注意力机制的输出分布，防止模型过度关注某些特征。本文引入了熵损失(Entropy Loss)作为辅助损失函数，如公式(6)所示：

$$\text{Entropy_Loss}=-\sum p\log p(-\lambda)=\lambda\sum p\log p \tag{6}$$

其中，p是注意力权重的概率分布，熵损失为负熵的加权形式。熵损失通过最大化熵来促进平衡，λ在本文的研究中设置为0.01(基于实验经验)。由上可知，对于实验中涉及的两种注意力模型，本研究的损失函数如公式(7)所示：

$$\text{Loss=MSE_Loss+Entropy_Loss} \tag{7}$$

3　实验和结果分析

3.1　实验设置与对比方法

3.1.1　实验基本设置与评价指标

本研究涉及的所有神经网络模型训练的过程中都使用Adam优化器进行参数更新。设置epoch为300轮，初始学习率设为0.0001，并采用线性预热策略：前10个epoch从0.00001逐步增加至目标学习率，避免初始梯度震荡。训练过程中引入早停机制，若验证损失连续50轮未下降则终止训练。学习率根据验证损失通过ReduceLROnPlateau策略动态调整，衰减因子为0.5，最小学习率下限为0.000001。批量大小统一设置为64，最大训练轮数为300，所有实验均在NVIDIA TESLA A100 GPU平台完成，显存为40 G。

本研究对数据的评价指标包括均方根误差(RMSE)和平均绝对误差(MAE)，这两种指标都是用于衡量模型精度的指标。其中均方根误差如公式(8)所示。

$$\text{RMSE}=\sqrt{\frac{1}{N}(\text{Pred-True})^2} \tag{8}$$

平均绝对误差如公式(9)所示：

$$\text{MAE}=\frac{1}{N}|\text{Pred}-\text{True}| \tag{9}$$

本文对模型的精度同时采用以上两种指标衡量，可以更加全面地表现出模型的精度。RMSE在计算过程中可以放大大值误差的影响，而海底地形构建任务本身属于大数值问题，RMSE指标对于模型的精度有着更敏锐的反映。

3.1.2　基础神经网络基线模型

本节介绍用于和残差U-Net注意力模型相对比的基础神经网络基线模型，在实验过程中我们将这些经典的神经网络全部改造成适配滑动窗口超分辨率任务的形式，以突显残差U-Net注意力模型骨干网络相比于其他基础模型骨干网络对于此任务的优越性。

多层感知机：构建基于多层感知机(Multilayer Perceptron，MLP)(Rumelhart等，1986)的基线模型，将展平

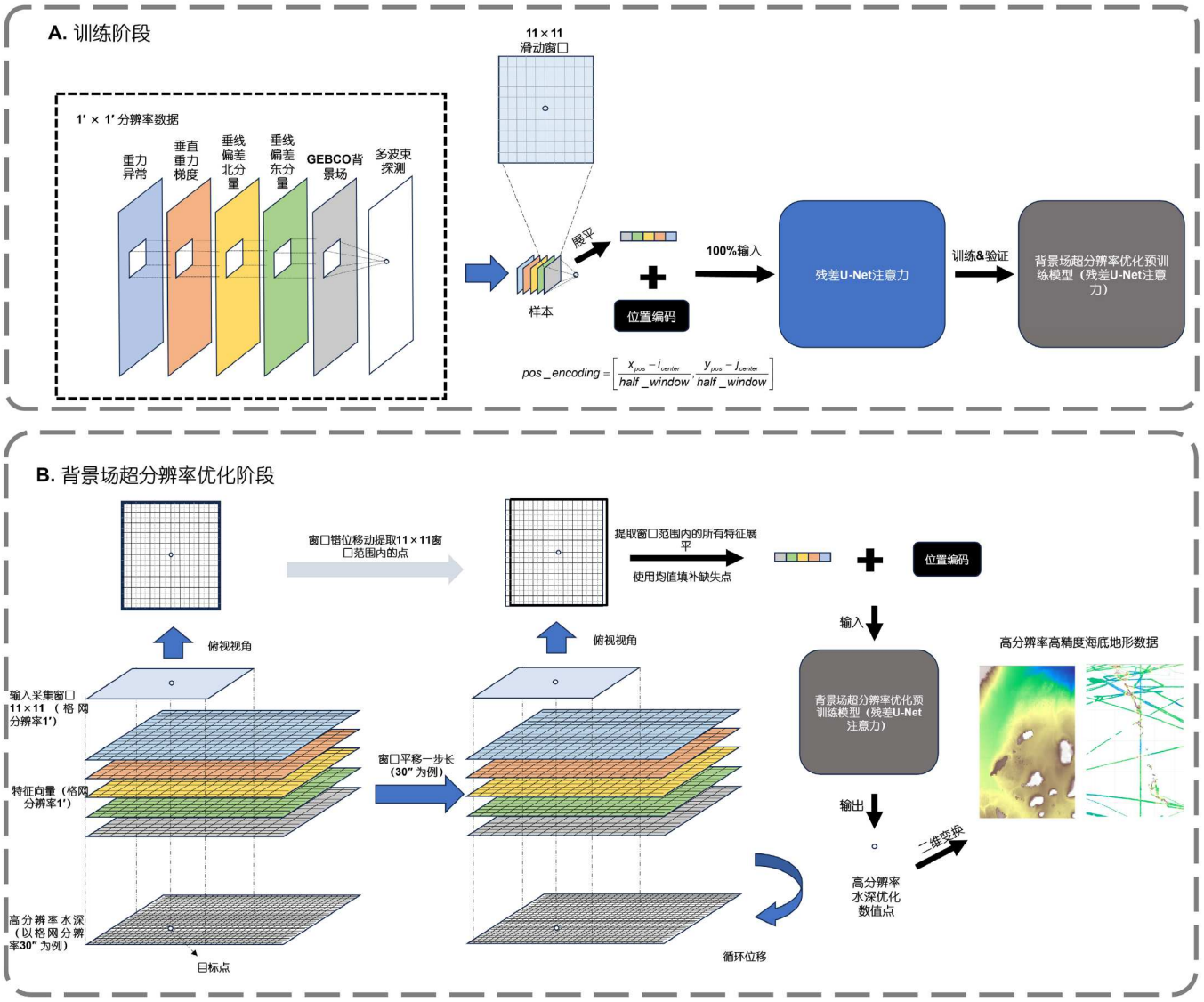


图3 基于残差U-Net注意力模型的海底地形背景场超分辨率优化方案总体流程图。(A) 模型训练阶段：展示了将1′分辨率的多源输入特征(重力异常、重力梯度异常、垂线偏差、GEBCO)与多波束真值通过11×11滑动窗口进行采样提取，并在展平后结合位置编码(Position Encoding)，输入至残差U-Net注意力网络进行训练以获取预训练模型的过程；(B) 背景场超分辨率优化(推理)阶段：展示了基于预训练模型，通过窗口错位移动采样策略提取特征并结合位置编码进行单点推理，最后经二维空间重排(2D Shift)完成高精度、高分辨率海底地形背景场重建的完整流程

后的多源滑窗特征与位置编码作为输入(11×11×5+2)。模型由三层全连接层组成，隐藏单元数依次为256、128和64，各层后接ReLU激活与Dropout(丢弃率0.3)(Srivastava等，2014)，最终输出单一水深值用于回归预测。该模型不显式建模空间结构，作为无空间归纳偏置的基线用于对比。

卷积神经网络：基于卷积神经网络(Convolutional Neural Network，CNN)(Lecun等，1998)构建滑窗模型，输入11×11×5的多通道特征。网络包含两层3×3卷积(通道数分别为16和32)，后接ReLU激活，并通过自适应全局平均池化将特征映射为固定长度向量，最终经全连接层回归水深值。CNN能有效捕捉局部空间模式，较MLP更适合处理具有空间相关性的地球物理数据。

基础通道注意力模型：为刻画多源数据的重要性差异，引入基础通道注意力模型。该模型为每个数据源构建独立的全连接编码器(隐藏单元为128和64)，随后将各通道特征与位置编码拼接，并通过注意力模块(Bahdanau等，2015)计算归一化权重以进行加权融合。模型采用残差连接以增强训练稳定性，并通过全连接层输出水深预测值。该模型可提供通道权重的可解释性，但由于各数据源独立编码，难以刻画其深层交互关系。其结构示意见图4。

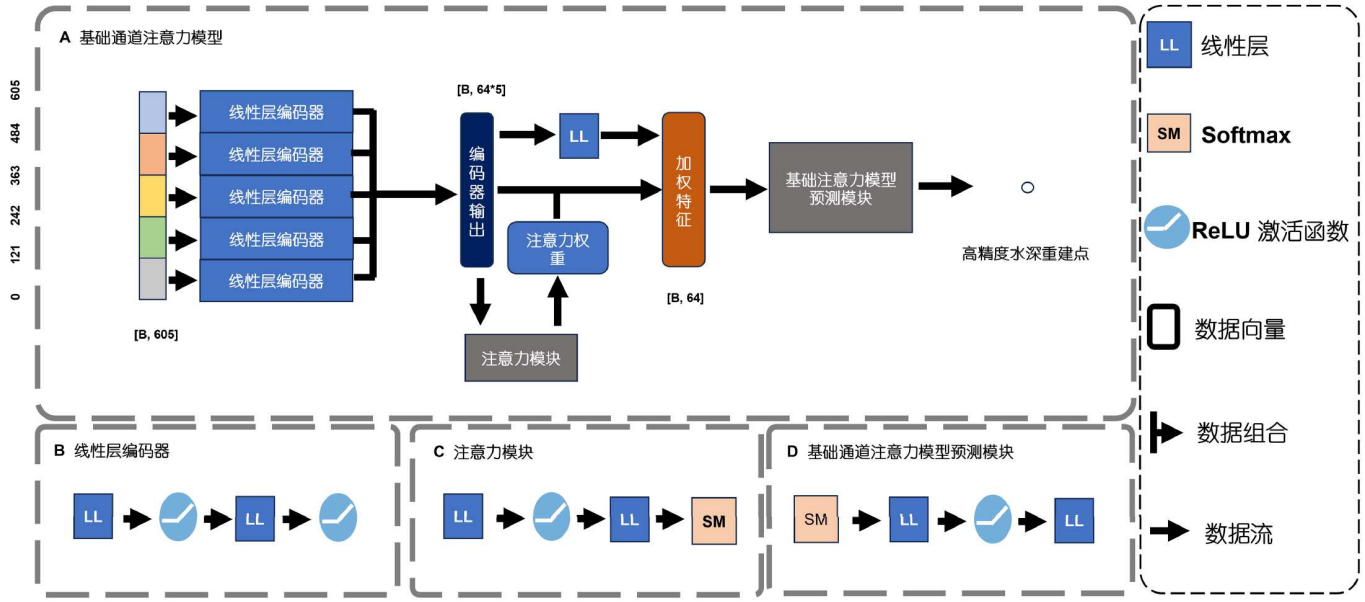


图 4　基础通道注意力模型(Basic Channel Attention Model)网络结构图。(A) 模型的总体架构，展示了多源特征序列的输入、注意力权重的计算与分配以及最终高精度地形点的预测输出过程；(B) 线性编码器(Linear Encoder)模块结构；(C) 注意力块(Attention Block)模块结构；(D) 模型预测块(Prediction Block)的内部层级结构

Transformer(Vaswani等，2017)：构建基于Transformer的基线模型，将各通道展平后的滑窗特征经线性映射并叠加位置投影后输入Transformer编码器，最终通过全连接层输出预测结果。该模型能够建模长程依赖与复杂特征交互，但对训练样本规模和噪声水平较为敏感，计算开销相对较高。

CNN+Transformer混合：采用CNN与Transformer的混合结构，首先利用轻量级CNN从各通道中提取局部特征表示，再将通道特征与位置特征一并送入Transformer编码器进行跨通道建模。该模型兼具局部特征提取与全局关系建模能力，但结构复杂度较高，对训练稳定性和正则化策略要求较高。

超分辨率残差网络(Super Resolution ResNet，SRResNet)：基于超分辨率残差网络SRResNet(Ledig等，2017)构建滑窗回归模型，利用头部卷积和多层残差模块提取局部特征，并通过自适应全局池化与全连接回归头输出水深值。残差结构有利于稳定训练并增强局部模式建模能力，适用于噪声水平较低区域的中高频特征学习。

深度加强超分辨率网络(Enhanced Deep Super-Resolution，EDSR)：采用增强型超分辨率网络EDSR(Lim等，2017)作为基线，其通过去除批归一化并使用宽残差路径以保留幅值信息，增强细节特征累积能力。EDSR在高频成分重建方面具有优势，适合刻画细微地形起伏，但相对更容易受到噪声影响，需要配合正则化或鲁棒损失函数使用。实验中涉及的基础模型信息汇总见表3。

3.1.3　超分神经网络优化变体基线模型

在上述基础模型的基础上，进一步构建多种超分辨率网络的优化变体，以分析不同结构增强策略在海底地形背景场超分辨率优化任务中的有效性。对比重点包括：单纯扩大模型容量、引入多尺度上下文感知以及结合注意力机制三类优化方向。

SRResNet大型模型：通过增加特征通道数与残差块数量对SRResNet进行扩展，用于评估在数据条件允许时，单纯扩大模型参数规模对地形细节重建性能的提升效果。

EDSR大型模型：EDSR的大参数实现，延续EDSR去除批归一化的设计以更好保留幅值信息，扩大宽残差路径与块数以提高高频成分恢复能力，但对训练稳定性与资源要求更高。

残差U-NetASPP：在Residual U-Net骨干网络的瓶颈处加入ASPP模块(多膨胀率并行卷积+全局池化)以并行捕获多尺度上下文，并在跳跃连接处应用通道注意力(Squeeze-and-Excitation，SE)。该方案在不降低空间分辨率的前提下增强感受野，有利于同时恢复大尺度趋势与小尺度细节。

U-Net门控注意力：在编码器各级并入通道门控模块(Channel Gate SE)，通过通道级缩放对跳跃特征做自适应重校准以抑制噪声通道并提升融合时的信噪比；瓶颈处同样保留自注意力以增强全局信息聚合。

膨胀卷积残差U-Net：在编码器中采用不同膨胀率的

表3　基础神经网络模型概况说明

模型名称	架构描述	参数量	主要特点
多层感知机	多层感知机，包含多个全连接层，带有ReLU激活函数和Drop out	~200 K	(1)简单高效；(2)易于训练和调参；(3)捕捉非线性关系
卷积神经网络	卷积神经网络，包含卷积层和全连接层，使用自适应平均池化	~50 K	(1)捕捉局部空间模式；(2)计算成本低；(3)自适应输出尺寸，适合不同窗口大小
基础通道注意力模型	注意力模型，分别编码各输入通道特征，通过注意力机制加权融合，包含残差连接	~100 K	(1)动态融合多源数据；(2)注意力权重提供可解释性；(3)残差连接增强稳定性
残差U-Net注意力模型	加入残差连接和多头注意力机制的U-Net架构，编码器解码器结构，使用卷积块处理空间特征	~2.5 M	(1)结合U-Net的空间特征提取与注意力机制，捕捉局部特征和全局特征；(2)残差连接增强模型训练的稳定性
Transformer	扁平化输入→线性投影到相对应的向量维度→若干Transformer编码器层→回归头	~500 K	(1)强表征与建模长程依赖；(2)能捕捉复杂的全局交互；(3)对训练样本量与噪声敏感
CNN-Transformer	每通道用小CNN提取token→token序列+位置token→Transformer编码器→池化回归	~600 K	(1)融合局部CNN与全局Transformer的优点；(2)适合跨通道复杂交互场景；(3)模型复杂度中高
SRResNet	SRResNet风格网络：头部卷积→若干残差块→全局池化+全连接回归	~500 K	(1)残差堆叠稳定训练；(2)全局池化汇聚局部信息；(3)在保守参数预算下改善细节学习
EDSR	EDSR变体(残差块、去批归一化设计)→全局池化+全连接回归，强调残差路径保留细节	~600 K	(1)去除批归一化、宽残差路径以保留细节；(2)在重建高频成分上优于常规残差设计；(3)对噪声敏感，需要正则或鲁棒损失

空洞卷积以扩大感受野，避免频繁下采样而丢失分辨率，旨在更好地捕获中-大尺度地形背景同时保留局部细节。实验中涉及的超分神经网络变体模型信息汇总见表4。

3.2　残差U-Net注意力背景场超分辨率优化模型有效性验证

本实验旨在验证基于窗口位移的多源数据融合背景场超分辨率优化方案中，残差U-Net注意力模型的有效性，验证其相比于经典神经网络模型的优越性，并基于模型误差系统地分析其性能。

3.2.1　30″分辨率的海底地形背景场超分辨率优化

表5总结了各种深度学习基础模型在30″分辨率下对海底深度预测的定量评估结果，对比了模型与多波束探测真值数据以及GEBCO地形数据集的关键指标，包括MAE、RMSE、平均值(Mean)、标准差(STD)、最大值(Max)和最小值(Min)。总体而言，残差U-Net注意力模型在两片实验海域均表现出色，其MAE和RMSE值显著低于其他模型和GEBCO背景场。在西沙海域，该模型的MAE为23.48 m、RMSE为65.74 m，远优于GEBCO的MAE(85.82 m)和RMSE(131.29 m)，RMSE的下降比率为49.9%。在夏威夷-皇帝海山链海域，此模型的MAE为52.62 m、RMSE为141.75 m，依然远优于GEBCO的MAE和RMSE，下降比率分别为10.5%和55.5%。这表明该模型在复杂地形中成功对背景场进行了优化并显著提高了高分辨率深度预测精度。RMSE作为对大值误差更为敏感的指标，它相较于GEBCO的下降比率显著更低，这体现出残差U-Net注意力模型重建的30″高分辨率海底地形在容易出现大误差的区域表现更稳定。

在上述参与对比的基础神经网络模型中选取了几种效果相对较好的基础神经网络模型对于30″分辨率的海底地形背景场超分辨率优化的直观效果如图5所示。可以看出残差U-Net注意力模型在两片海域的地形重建效果均最接近多波束探测数据真值的效果。这初步反映出残差U-Net注意力模型的有效性。

为更加直观反映出模型之间的对比效果，对于不同区域不同模型的实验绘制了模型绝对误差分布的误差热图(图6)以及更能直观反映出数值拟合情况的散点拟合图(图7)。差值分布图的颜色越浅表明模型输出和多波束探测的差别越小，散点拟合图的点越集中贴近直线表明模型对真值的拟合效果越强。可以看到，无论是西沙海域还是皇帝海山链海域，残差U-Net注意力模型背景场超分辨率优化绝对误差图最接近浅灰色(0值)。散点拟合图同样表现出此模型的点相对更紧凑，且没有明显的大偏差点聚集的情况，具有最佳的真值拟合效果。

图8表示了不同模型在不同海域的一维绝对误差分布状况。从图8中可以清晰地看到棕色线条的残差U-Net注意力模型相比GEBCO背景有着更低的误差，这表明背景场优化模型可以有效地将背景场分辨率提升2倍的同时兼顾更高的精度；此模型也具备和其他所有模型相比更少的极端大值误差，有着更平稳精准的总体输出效果。相比之下，结构较为简单的MLP和CNN基础通道注意力模

表4　超分神经网络变体模型概况说明

模型名称	架构描述	参数量	主要特点
残差U-NetASPP	在残差U-Net瓶颈处加入ASPP模块以获取多尺度上下文，并在跳跃连接处加入通道门控模块	~1 M	(1)ASPP扩大感受野、捕获多尺度上下文；(2)跳跃处SE提升通道选择性与鲁棒性；(3)计算中等、实现稳健，适合多尺度地形
门控注意力残差U-Net	在残差U-Net中对每级编码输出加入通道门控(SE)，并在瓶颈处保留自注意力(轻量注意力融合)	~2.5 M	(1)通道门控抑制噪声通道并增强有用通道；(2)与残差U-Net结合后在多源噪声场景更鲁棒；(3)实现代价较低，易集成
膨胀卷积残差U-Net	在编码器使用不同膨胀率的膨胀卷积(dilated conv)以在不降分辨率情况下扩大感受野，瓶颈保留残差结构与可选注意力	~1 M	(1)更大感受野用于捕获中大尺度地形趋势；(2)在保持空间分辨率下有利于长程依赖建模；(3)需防止膨胀引入的格点效应，慎选膨胀率
SRResNet大型模型	SRResNet的放大版本，增大通道数与残差块数量以提升容量与细节重建能力	~5 M	(1)更大容量有利于拟合复杂映射；(2)对数据量与正则化更敏感；(3)训练/推理资源需求显著上升
EDSR大型模型	EDSR的放大版本，扩展通道与残差块数，保留去批归一化设计以强化高频重建	~8 M	(1)在高频细节恢复上潜力大；(2)需更稳健训练策略(小lr、混合精度、梯度裁剪)；(3)易受噪声与过拟合影响

表5　不同海域各模型重要指标统计表(30″分辨率)

区域	模型名称	平均绝对误差 MAE(m)	均方根误差 RMSE(m)	深度平均值 Mean(m)	深度标准差 STD(m)	深度最小值 Min(m)	深度最大值 Max(m)
西沙海域	多波束探测(真值)	–	–	1226.3	483.48	0	2604.28
	GEBCO背景	85.82	131.29	1178.97	502.98	0	2621
	多层感知机	117.41	162.04	1186.08	368.64	0	2337.88
	卷积神经网络	149.66	231.81	1123.19	326.46	0	1859.49
	基础通道注意力模型	160.07	222.45	1112.34	393.51	0	2034.13
	残差U-Net注意力模型	23.48	65.74	1228.65	475.08	0	2670.88
	Transformer	194.49	278.97	1124.67	273.6	174	1663.86
	CNN-Transformer	292.67	406.35	1032.70	192.96	160.23	1261.97
	SRResNet	84.38	126.74	1176.59	397.24	0	2297.79
	EDSR	102.67	154.79	1154.23	377.29	0	2224.31
夏威夷-皇帝海山链海域	多波束探测(真值)	–	–	5219.11	1154.25	280.30	7654.39
	GEBCO背景	58.82	318.60	5209.22	1189.62	283.00	7671.00
	多层感知机	229.97	398.60	5347.48	867.75	1116.77	6602.38
	卷积神经网络	192.12	378.01	5347.5	886.05	819.91	6835.82
	基础通道注意力模型	121.76	331.72	5294.95	1028.713	220.77	8497.93
	残差U-Net注意力模型	52.62	141.75	5226.29	1136.019	346.91	7473.52
	Transformer	755.02	1371.60	5963.93	621.47	1977.47	10972.59
	CNN-Transformer	619.28	1095.37	5798.66	312.22	3901.67	6405.44
	SRResNet	377.74	667.01	5581.36	669.93	2027.98	7473.87
	EDSR	350.33	520.67	5559.02	936.18	925.91	7775.18

型在两片海域的背景场超分辨率优化表现均不理想，可见对于这种错位移动的背景场超分辨率优化任务，上述模型不具备足够强大的提取数据空间相关性的能力，这两种基础神经网络架构均在30″分辨率背景场超分辨率优化任务中有着劣于GEBCO数据的精度指标，这说明模型不具备背景场优化能力。而参与了对比但是未在图中展示的Transfomer模型和CNN-Transformer模型具备较大误差，可见Transformer骨干网络结构难以在此类背景场超分辨率优化任务中发挥作用，经典的超分神经网络SRResNet有着优于GEBCO的效果但是相比于残差U-Net注意力模型还是有着较大的精度差距。以上现象表明残差U-Net骨干网络在此类滑窗海底地形超分任务中相比其他神经网络骨干具备优越性。

3.2.2　15″分辨率的海底地形背景场超分辨率优化

表6展示了在15″分辨率下，各种深度学习模型对海

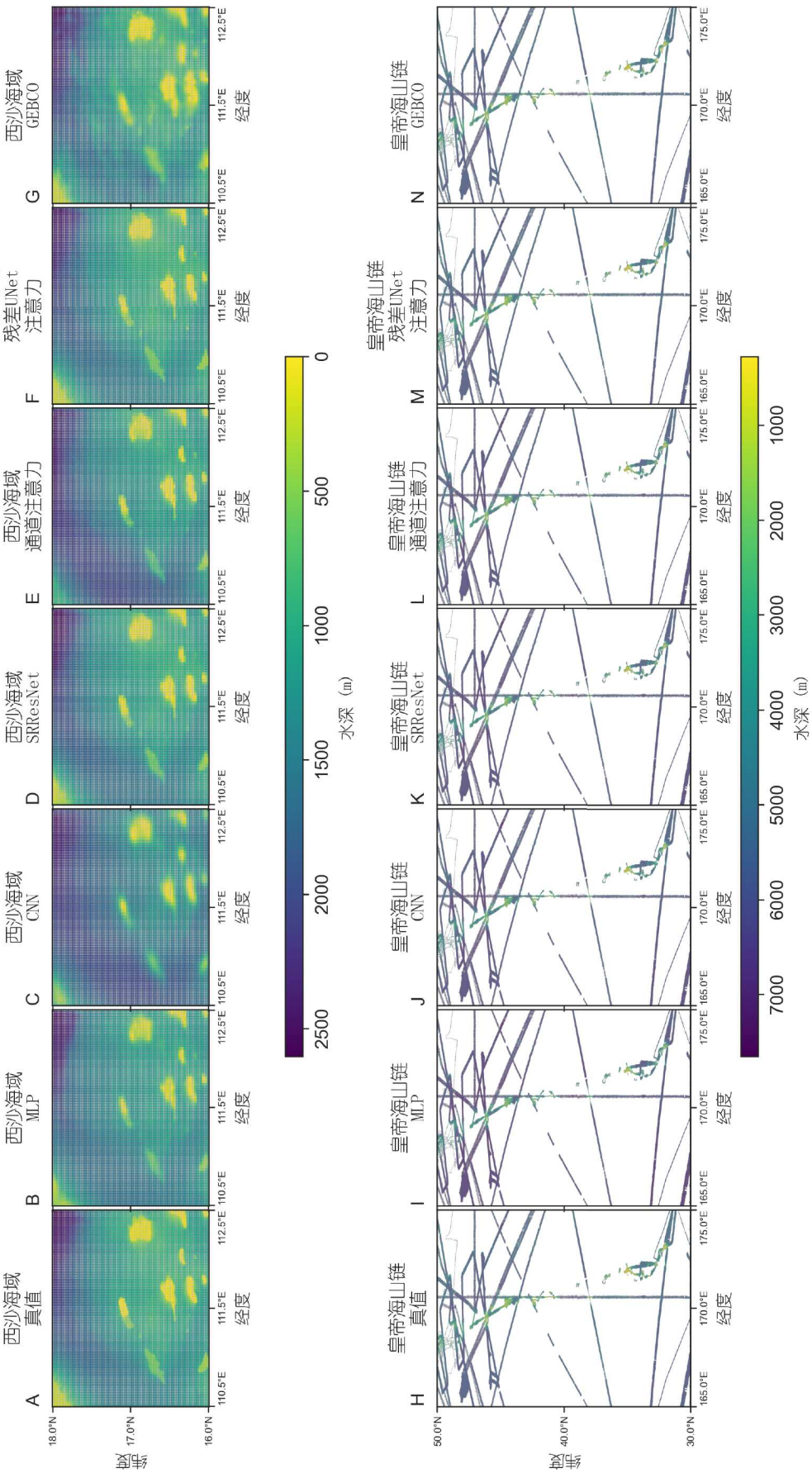


图5　不同海域海底地形超分辨率重建效果对比图(30″分辨率)。(A–G) 西沙海域的重建结果；(H–N) 皇帝海山链区域的重建结果。从左至右依次为多波束测量真值(Ground Truth)、MLP、CNN、SRResNet、通道注意力模型、残差U-Net注意力模型以及GEBCO基线数据

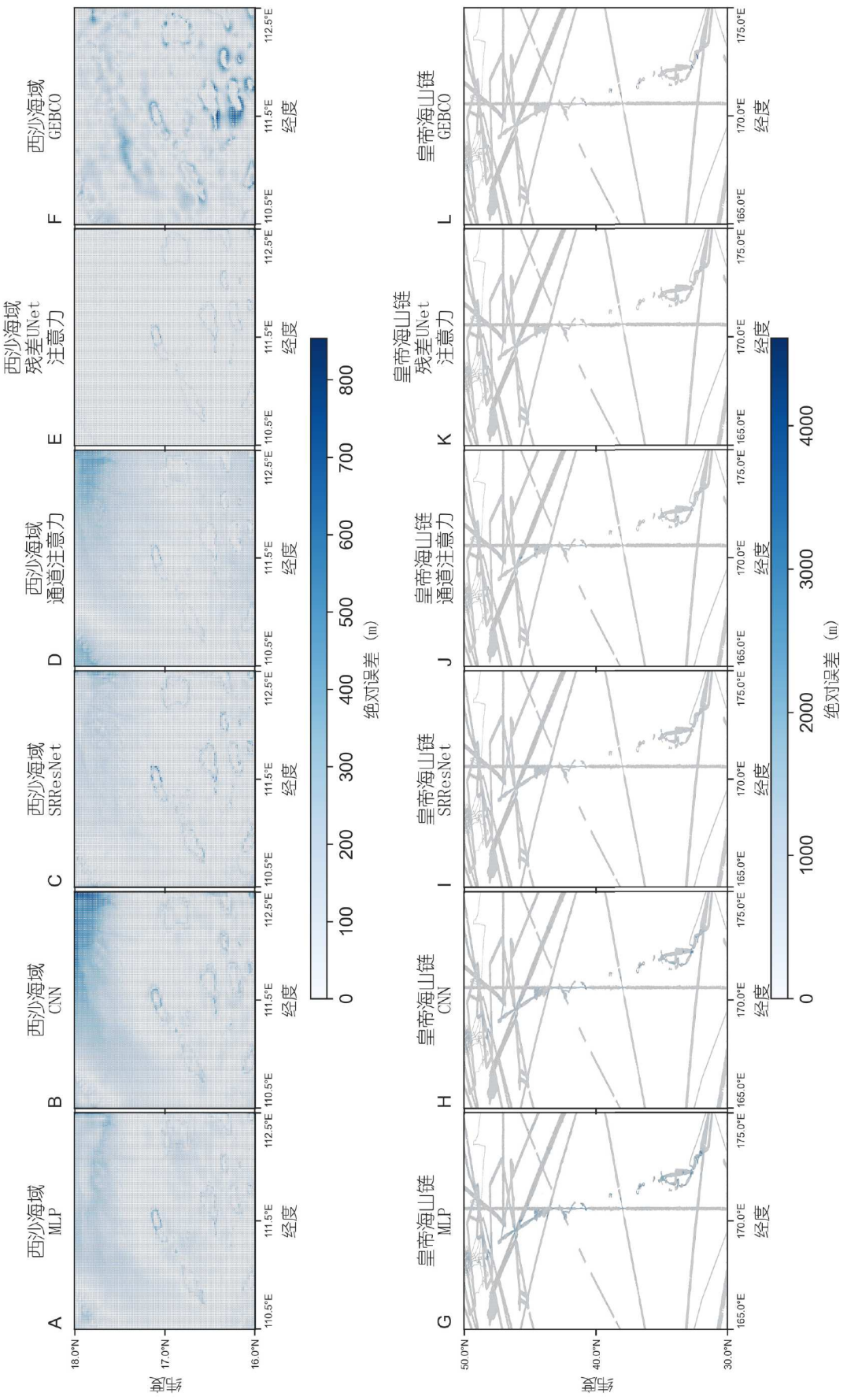


图 6　不同海域各模型预测结果的绝对误差分布热图(30″分辨率)。(A–F) 西沙海域；(G–L) 皇帝海山链区域。从左至右依次为MLP、CNN、SRResNet、通道注意力模型、残差U-Net注意力模型与GEBCO数据的空间误差分布。图例颜色越深代表模型预测值与真值之间的绝对误差越大

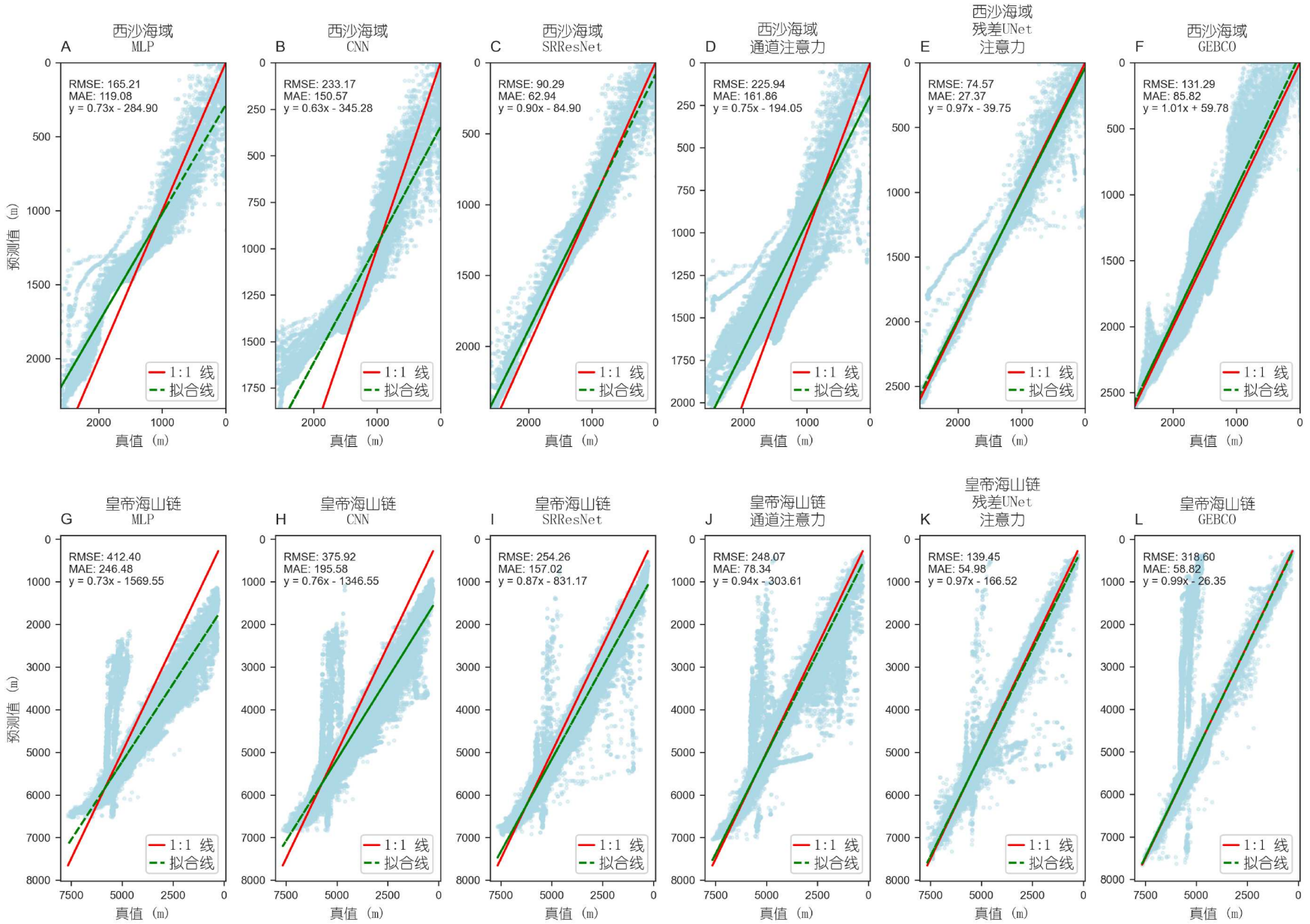


图 7 不同海域各模型预测深度的散点拟合图(30″分辨率)。(A–F) 西沙海域；(G–L) 皇帝海山链区域。从左至右依次为MLP、CNN、SRResNet、通道注意力模型、残差U-Net注意力模型与GEBCO的散点图。其中，红色实线代表1：1完美预测参考线，绿色虚线代表预测数据的线性拟合线。每个子图左上角附有该模型的RMSE与MAE指标

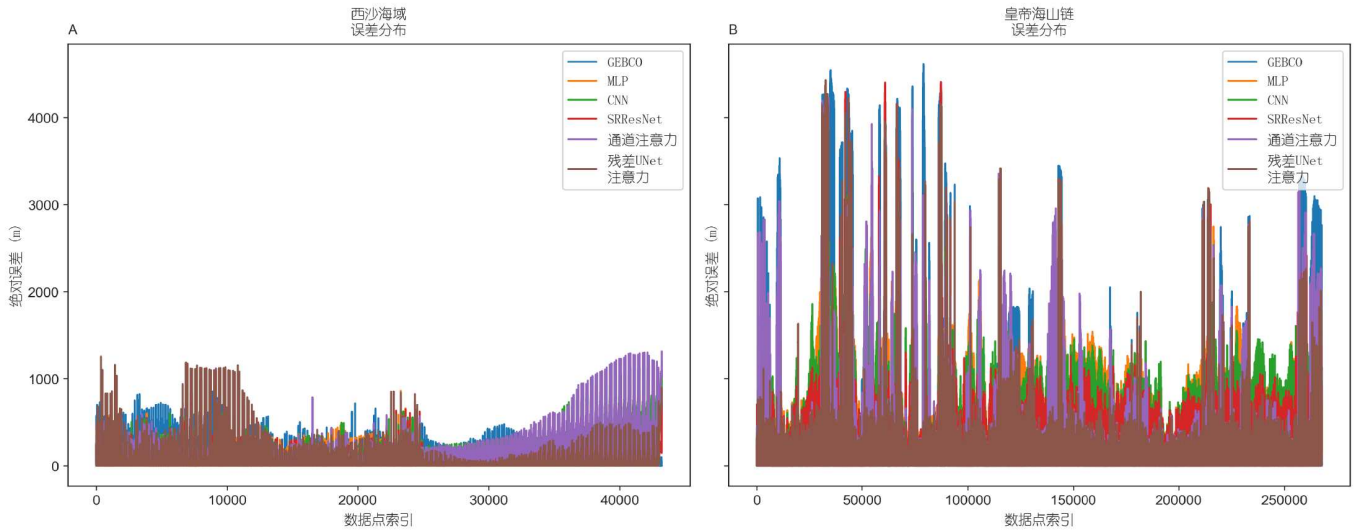


图 8　不同海域各模型的一维绝对误差分布折线图(30″分辨率)。(A) 西沙海域内有效数据点的绝对误差波动情况；(B) 皇帝海山链区域的绝对误差波动情况。不同颜色的曲线代表了不同神经网络模型与GEBCO基线数据的误差分布，横坐标为展平后的有效数据点索引

表 6　不同海域各模型重要指标统计表(15″分辨率)

区域	模型名称	平均绝对误差 MAE(m)	均方根误差 RMSE(m)	深度平均值 Mean(m)	深度标准差 STD(m)	深度最小值 Min(m)	深度最大值 Max(m)
西沙海域	多波束探测(真值)	–	–	1224.12	486.48	0	2604.28
	GEBCO背景	88.17	145.45	1179.99	504.40	0	2621
	多层感知机	118.97	169.43	1186.42	369.57	0	2337.88
	卷积神经网络	151.68	237.38	1123.35	326.64	0	1859.49
	基础通道注意力模型	161.27	226.18	1112.75	394.04	0	2034.26
	残差U-Net注意力模型	24.49	86.84	1229.59	476.09	0	2670.88
	Transformer	196.76	284.28	1124.82	273.74	173.85	1665.39
	CNN-Transformer	294.23	408.65	1032.63	193.02	159.86	1261.97
	SRResNet	85.86	135.77	1177.08	397.73	0	2297.83
	EDSR	104.24	162.78	1154.64	377.88	0	2224.35
夏威夷-皇帝海山链海域	多波束探测(真值)	–	–	5218.83	1155.41	279.19	7710.41
	GEBCO背景	58.77	318.32	5208.68	1190.62	280.00	7696.00
	多层感知机	229.56	397.98	5347.16	868.51	1116.72	6602.38
	卷积神经网络	191.75	377.42	5347.38	886.67	801.52	6837.34
	基础通道注意力模型	119.10	328.86	5294.33	1029.80	220.70	8574.28
	残差U-Net注意力模型	44.94	126.34	5225.36	1137.43	346.91	7504.72
	Transformer	755.66	1373.79	5964.65	622.73	1957.21	10986.58
	CNN-Transformer	619.59	1096.05	5798.66	312.54	3900.89	6405.53
	SRResNet	155.99	250.56	5355.75	1012.95	506.25	6940.37
	EDSR	350.07	520.18	5559.10	936.54	925.91	7778.63

底地形背景场超分辨率优化的定量评估结果。在西沙海域，残差U-Net注意力模型以24.49 m的最低MAE和86.84 m的RMSE表现出色，相比GEBCO的MAE(88.17 m)和RMSE(145.45 m)分别下降了72.2%和40.3%，而其他基础神经网络模型在15″分辨率背景场超分辨率优化任务上的精度指标表现远远弱于残差U-Net注意力模型，这表明残差U-Net注意力模型对15″分辨率超分辨率任务具有优异效果。在夏威夷-皇帝海山海域，模型的表现差异显著。残差U-Net注意力模型再次表现优异，MAE为44.94 m，RMSE为126.34 m，远远低于GEBCO在此海域的MAE(58.77 m)和RMSE(318.32 m)，在数值上分别下降23.5%和60.3%。其他基础神经网络模型在15″分辨率背景场超

分辨率优化任务上依然有着较弱的表现，有着较大的误差并且效果均弱于GEBCO模型。

选取了几个效果较好的模型对15″分辨率的海底地形背景场超分辨率优化的实验结果数据如图9所示，不同的子图表示了不同模型在不同区域地形重建的视觉效果。可以看到对于西沙海域只有注意力模型和残差U-Net注意力模型具有和多波束探测真值相当的视觉效果，而GEBCO数据明显有一些数据噪点。在夏威夷-皇帝海山链海域残差U-Net注意力模型和GEBCO数据的颜色更接近真值。初步显示出残差U-Net注意力模型在15″分辨率背景场优化任务中的有效性。

为直观反映出模型之间的误差对比情况，本文对于不同区域不同模型的实验绘制了反映模型绝对误差分布的误差热图(图10)。通过误差热图可以看到，15″超分辨率任务中残差U-Net注意力模型的绝对误差分布几乎为完全的浅灰色，这表明了其具有最低数值的绝对误差分布，而误差相对较大的注意力模型和GEBCO都有表示较大误差的浅色斑块。

同时我们也绘制了能直观展示出数值拟合情况的散点拟合图，如图11所示。在西沙海域，残差U-Net注意力模型的数据点分布明显更接近直线，更加紧凑，而GEBCO分布较散，这表示有一个较高的平均误差水平。在皇帝海山链区域，虽然总体上残差U-Net注意力模型和GEBCO两者之间的点线拟合效果类似，但是GEBCO出现了一块更为密集的离群误差点，这与上面的误差热图中发现的零星深蓝色点相对应，显示了大值误差的存在。

图12表示了不同模型在不同海域的15″分辨率超分辨率任务的一维绝对误差分布情况，可以清楚地看到残差U-Net注意力模型具备更少的极端误差，具有更平稳的背景场超分辨率优化效果。在不同海域的分析表明，残差U-Net注意力模型在15″分辨率的海底地形背景场超分辨率优化的表现依然始终优于其他模型，有着最低的MAE和RMSE指标。同时值得注意的是，在两片海域的所有实验数据表明，残差U-Net注意力模型在15″分辨率的背景场超分辨率优化精度指标未因分辨率的提高而得到明显衰减。基于残差U-Net注意力模型的窗口错位背景场超分辨率优化方案可以在提升目标海域背景场分辨率的同时，保证其相对于真实标签的精度依然可靠。这得益于它强大的空间特征提取能力和基于注意力机制的灵活多源数据整合能力。

3.3 残差U-Net注意力模型超分辨率误差分析

上述实验结果验证了基于窗口位移的海底地形背景场超分辨率优化方案下的残差U-Net注意力模型在不同分辨率情况下优秀的性能。本小节内容将结合空间域和频域的分析进一步阐释此方案在多分辨率下背景场优化性能的稳定性。图13和14分别展示了在30″和15″分辨率情况下具有代表性的基础神经网络模型超分辨率误差点数值位于不同区间的占比统计。在西沙海域，随着分辨率的提升误差大于300 m数据点的比例几乎持平(残差U-Net注意力模型的大于300 m误差点比率从1.3%降至1.1%，相对稳定)，增大分辨率后，有效点基数的增加使得大值误差的数量大幅度增加，这造成了对大值误差敏感的RMSE指标显著上升；在夏威夷-皇帝海山链区域，随着分辨率的提升误差大于300 m的数据点比例有小幅下降(残差U-Net注意力模型的大于300 m误差点比率从1.9%降低到了1.3%)，这造成了此海域的RMSE指标随着分辨率的提高有所下降。残差U-Net注意力模型的大值误差数量占比相较于其他模型一直处在最低的水平，对大值误差的有效减少也是其RMSE指标效果最好的根本原因。

为了更全面地对模型效果进行评估和说明，本文对二维深度场进行了局部径向平均功率谱密度(PSD)分析以比较真值与模型预测的尺度特征，为了确保可复现性，详细说明频谱分析步骤如下：(1)在每个区域内以固定大小(默认128×128像素)的滑动子窗口采样，保留真值与预测窗口内有效像素比例高于阈值的窗口；(2)对每个窗口先用基于有效像素的二维插值填补缺失值并去均值，再乘以二维Hanning窗以减少泄露，FFT后将功率谱除以窗能量进行补偿；(3)根据像素实际物理间距(考虑纬度方向的经度缩放)构建空间频率网格，将二维谱按半径分箱并做径向平均得到一维频谱(单位：cycles km^{-1})；(4)对同一区域/模型在不同窗口上得到的PSD在每一频率区间的集合上取样本均值作为主曲线；(5)当可用窗口过少时，通过小幅随机平移和镜像增强单窗样本以估计谱波动。 图15和16分别为在30″分辨率和15″分辨率重建任务中各个模型在不同海域的功率谱密度图。

从30″分辨率的功率谱密度图(图15)可以看出，在西沙区域通道注意力模型和残差U-Net注意力模型的曲线与真值曲线重合度最高，具备较优的真值拟合效果。在空间频率为0.1 cycles km^{-1}的中等频率特征区间处，残差U-Net注意力模型可以与真值保持最佳的贴合效果，这体现了在30″分辨率超分任务中此模型对中等频率地形特征的重建具备突出优势。在皇帝海山链区域残差U-Net注意力模型依然相比于其他的基础神经网络模型有着更优的拟合度。

从15″分辨率的功率谱密度图(图16)依然可以看到，相比于其他基础神经网络模型，残差U-Net注意力模型有着最优的拟合效果，且对于中等频率地形有着突出优势。

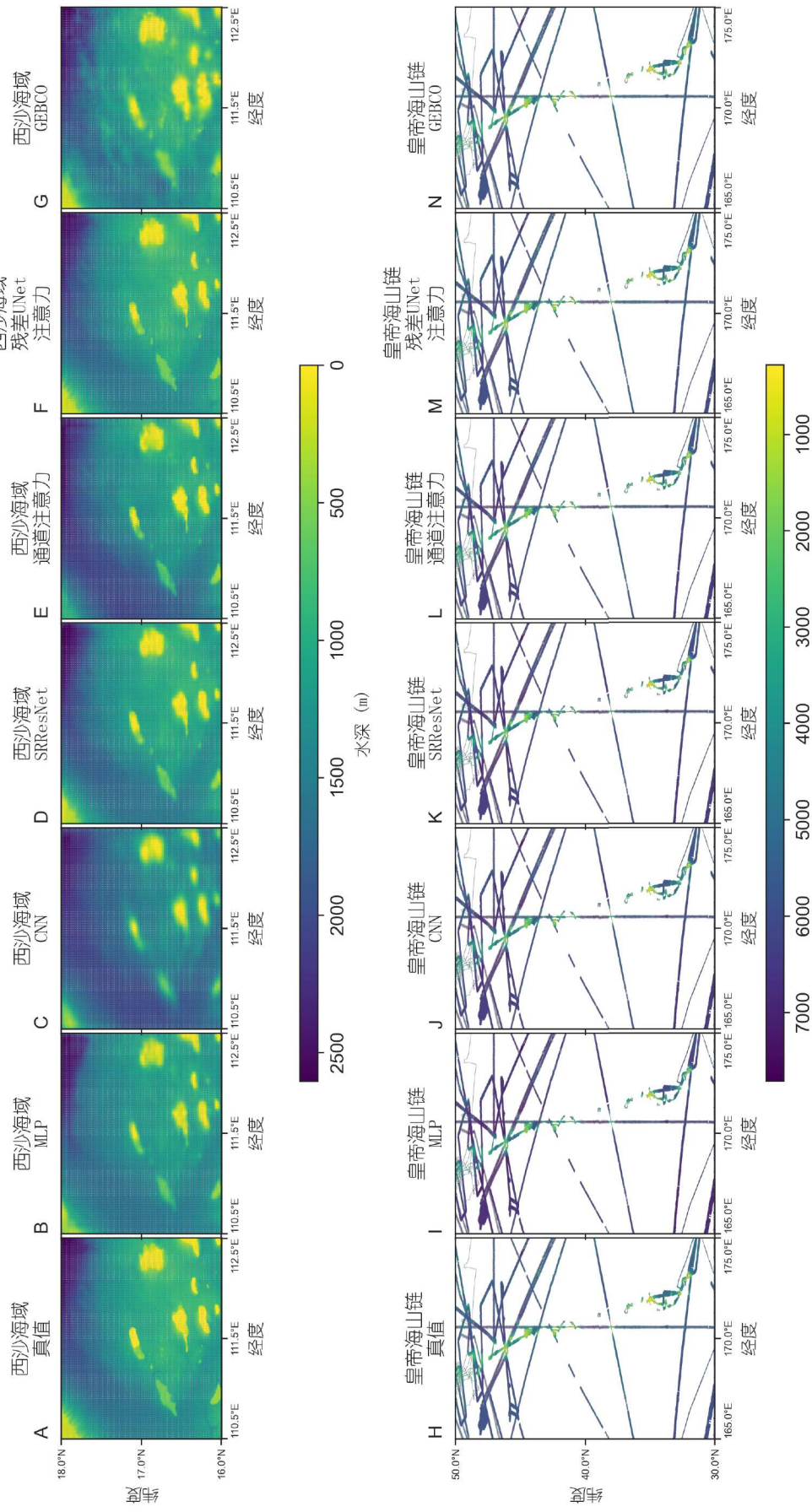


图 9　不同海域海底地形超分辨率重建效果对比图(15″分辨率)。(A–G) 西沙海域的重建结果；(H–N) 皇帝海山链区域的重建结果。从左至右依次为多波束测量真值(Ground Truth)、MLP、CNN、SRResNet、通道注意力模型、残差U-Net注意力模型以及GEBCO基线数据

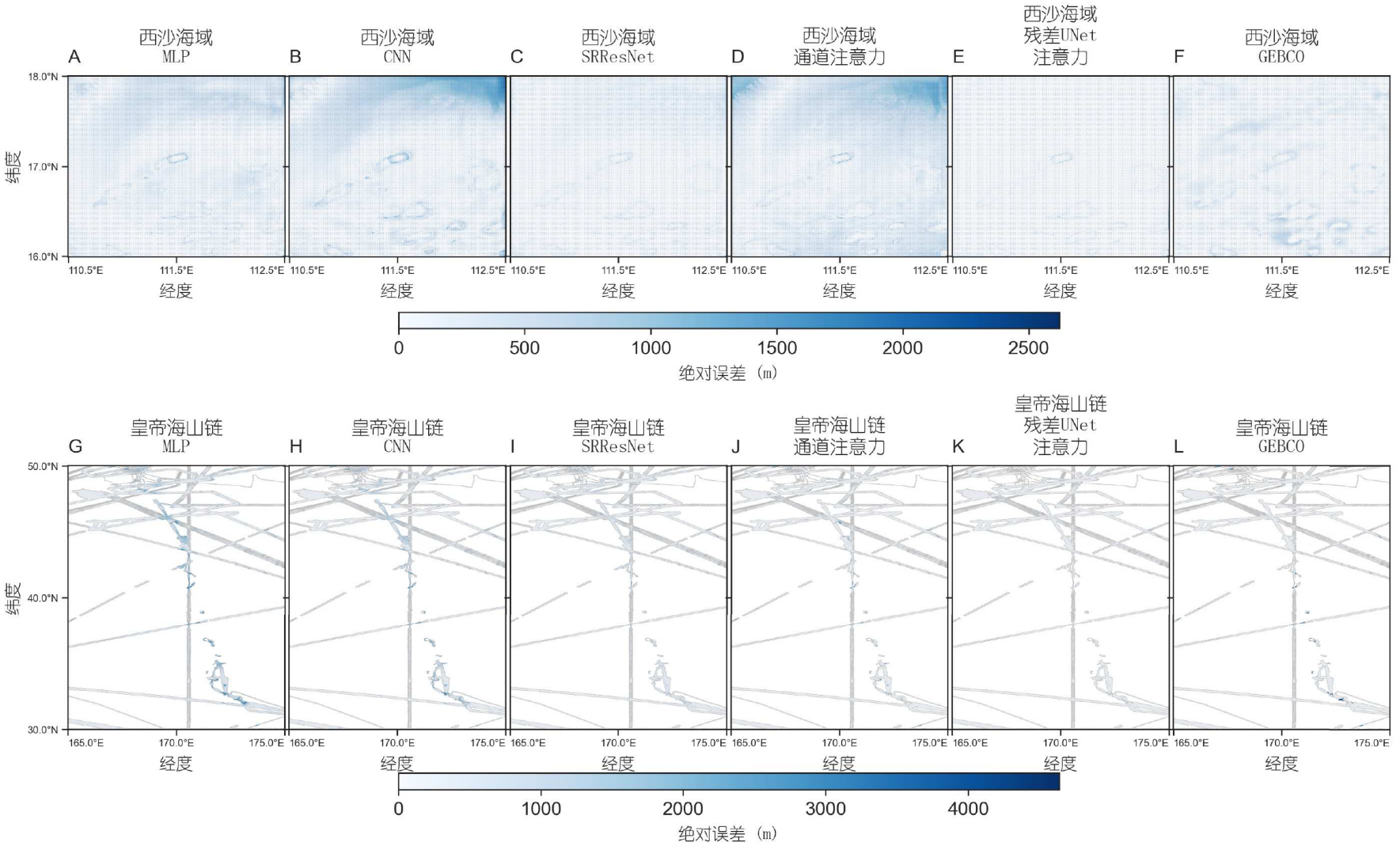


图 10 不同海域各模型预测结果的绝对误差分布热图(15″分辨率)。（A–F）西沙海域；（G–L）皇帝海山链区域。从左至右依次为MLP、CNN、SRResNet、通道注意力模型、残差U-Net注意力模型与GEBCO数据的空间误差分布。图例颜色越深代表模型预测值与真值之间的绝对误差越大

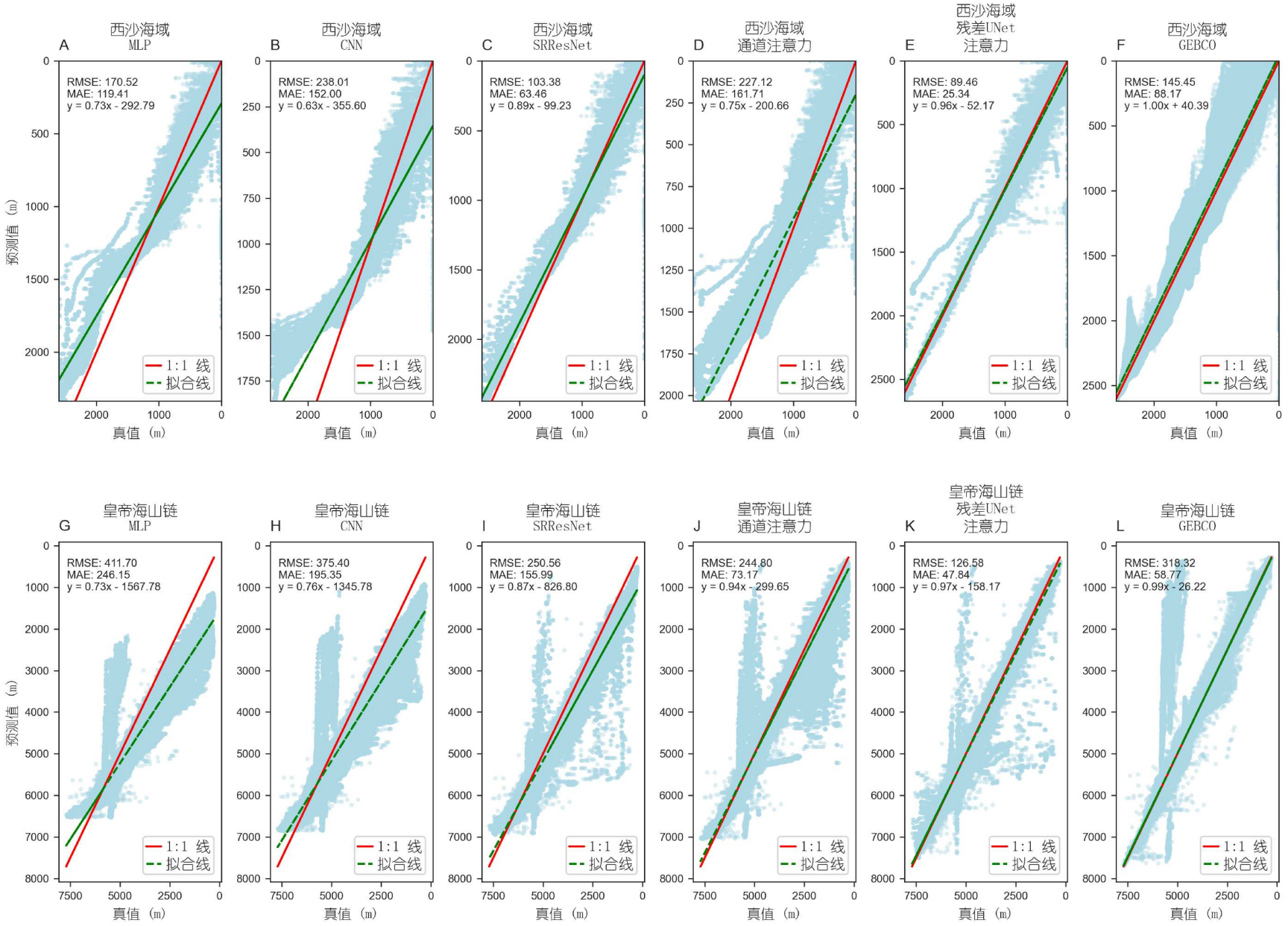


图 11 不同海域各模型预测深度的散点拟合图(15″分辨率)。（A–F）西沙海域；（G–L）皇帝海山链区域。从左至右依次为MLP、CNN、SRResNet、通道注意力模型、残差U-Net注意力模型与GEBCO的散点图。其中，红色实线代表1：1完美预测参考线，绿色虚线代表预测数据的线性拟合线。每个子图左上角附有该模型的RMSE与MAE指标

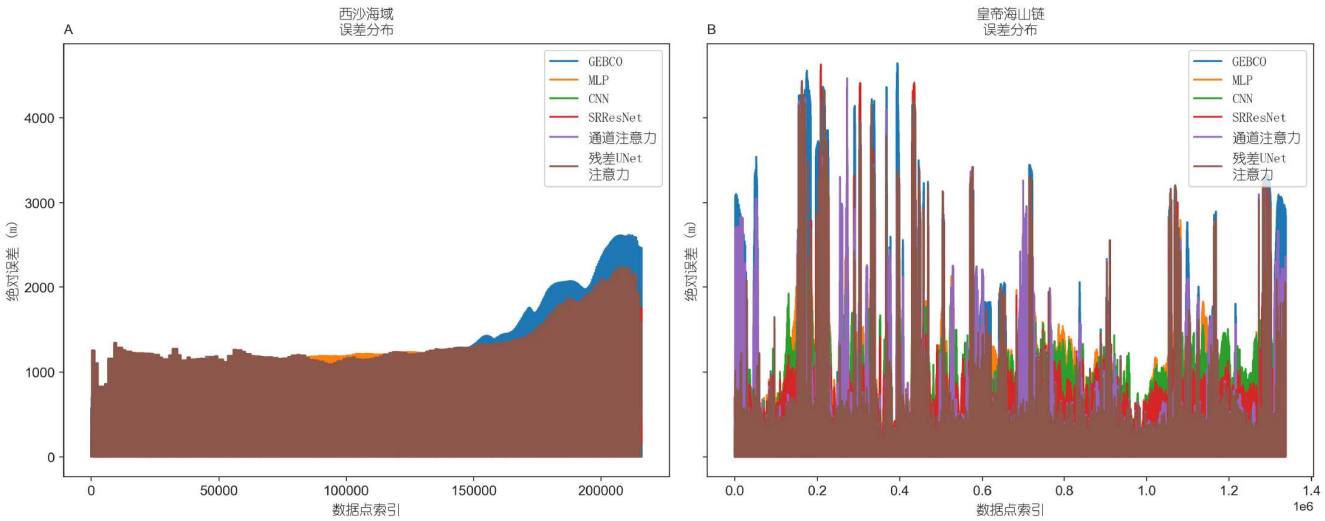


图 12 不同海域各模型的一维绝对误差分布折线图(15″分辨率)。(A) 西沙海域内有效数据点的绝对误差波动情况；(B) 皇帝海山链区域的绝对误差波动情况。不同颜色的曲线代表了不同神经网络模型与GEBCO基线数据的误差分布，横坐标为展平后的有效数据点索引

更高的分辨率代表着对模型细节重建能力有着更高的要求，从图16中还可以发现GEBCO数据在0.1–1 cycles km^{-1}频段产生了较大的偏差，而残差U-Net注意力模型在中高频区间和真值依然有着更好的拟合效果，这表明残差U-Net模型在更高分辨率的地形重建任务中有着更强的高频细节重建效果。虽然西沙海域模型输出的PSD曲线总体最贴近真值数据，但是在高频频段曲线仍有明显偏离。该偏差可能侧面反映出模型在某些高频尺度上过度恢复或放大噪声，但是不影响背景场优化的整体精度统计指标。

3.4 对于海底地形背景场超分辨率优化网络结构的更多探究

本节旨在系统性地探索残差U-Net骨干网络的多种优化路线对海底地形背景场超分辨率优化性能的影响。具体目标不是重复基线网络之间的横向对比，而是在以残差U-Net为基础的统一框架下，逐项引入并评估若干代表性改进模块(包括ASPP用于并行多尺度上下文聚合、膨胀卷积以扩大不下采样情形下的感受野、通道注意力以增强多源通道选择性)，同时考察超分神经网络单纯扩大超分神经网络容量(SRResNet大型模型、EDSR大型模型)所带来的性能增益。为保证结论的公正与可比性，所有试验在相同的数据预处理流程、相同的训练超参数与统一的评估掩码下进行：训练集均采用全部1′分辨率的多源观测点与对齐的GEBCO标签；测试集使用独立未参与训练的30″与15″多波束真实观测点。评价维度包括重建精度(RMSE、MAE)。通过该分步消融与对比实验，我们既旨在揭示各类结构增强相较于容量扩张的相对价值，也期望给出在实际工程部署中针对不同资源约束与数据条件的优化建议。

表7(30″)与表8(15″)的实验结果表明，在统一的数据预处理与评估流程下，基于残差U-Net的模型族总体上在海底地形超分任务中表现最为稳健且精度最高。具体而言，残差U-Net系列(含注意力机制与ASPP增强版本)在两片海域和两档目标分辨率上均取得了最低的MAE与RMSE，且在平坦至中等起伏的西沙海域与地形复杂的夏威夷-皇帝海山链均展现了出色的局部有效性与多地形适应性。ASPP与通道注意力作为结构性增强，在大多数情形下能进一步或微幅改进基础残差U-Net的性能，说明并行多尺度上下文聚合与通道选择性对同时恢复广域趋势与局部细节均有积极作用。膨胀卷积在扩大感受野、改善中大尺度结构建模方面也表现出正面效果，但其收益通常低于融合ASPP或通道注意力的变体，且对膨胀率的选择较为敏感。

相比之下，超分领域经典的残差堆叠模型(SRResNet、EDSR)地形重建质量优于GEBCO基线，但在地形复杂或采样不规则的海域中稳定性不足，尤其是小型实现与某些大型变体在复杂海域上出现较高的RMSE。尽管将SRResNet/EDSR扩展为大参数量版本在某些场景下有助于降低误差，但从总体结果看，单纯放大模型容量并不能保证在复杂、噪声显著的海域中超越具备多尺度特征聚合与注意力机制的U-Net变体。

4 结论与讨论

本文提出了一种基于窗口错位移动的多源数据融合背景场超分辨率优化方案，并在西沙群岛和夏威夷-皇帝

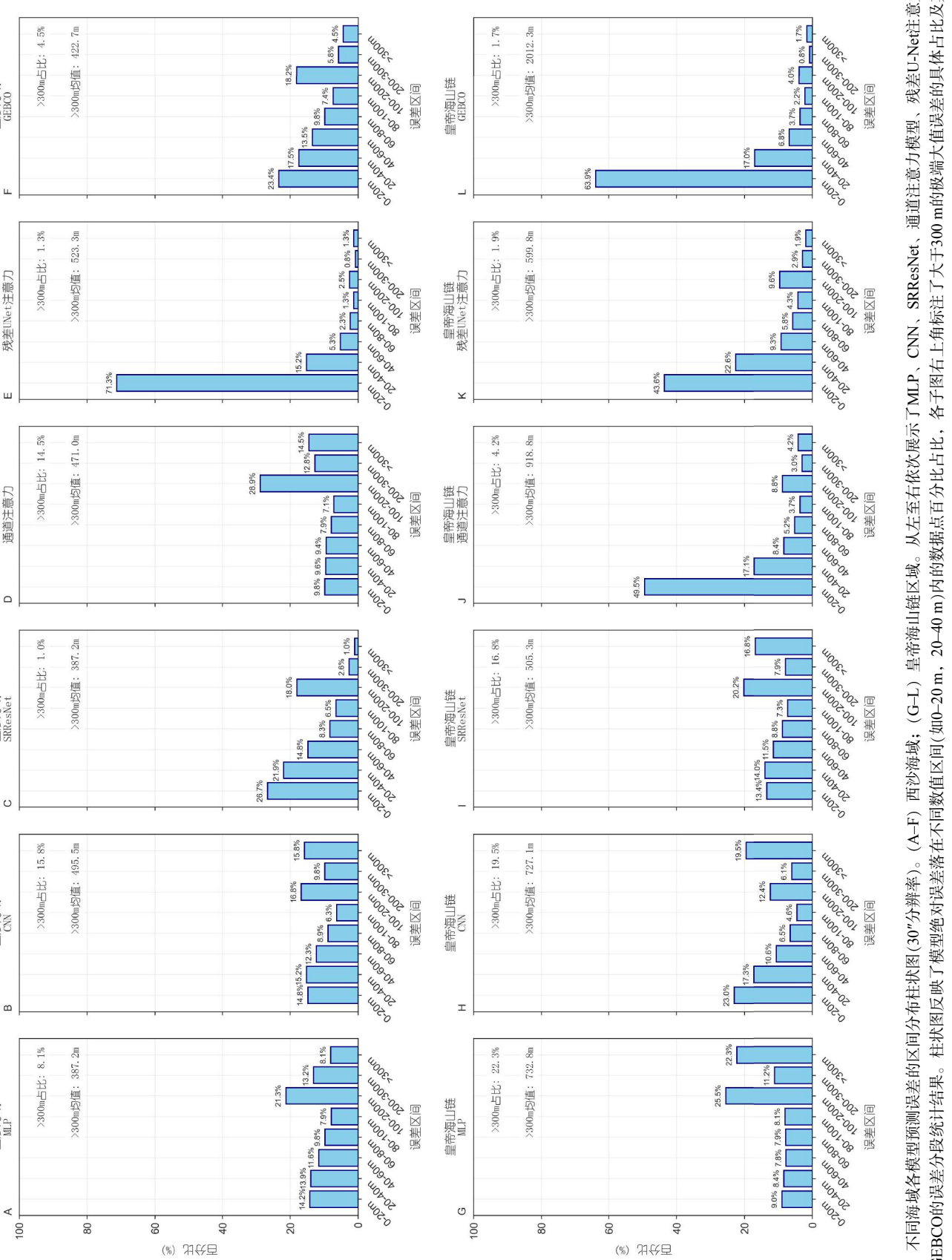


图 13　不同海域各模型预测误差的区间分布柱状图(30″分辨率)。(A–F) 西沙海域；(G–L) 皇帝海山链区域。从左至右依次展示了MLP、CNN、SRResNet、通道注意力模型、残差U-Net注意力模型与GEBCO的误差分段统计结果。柱状图反映了模型绝对误差落在不同数值区间(如0–20 m，20–40 m)内的数据点百分比占比，各子图右上角标注了大于300 m的极端大值误差的具体占比及其均值

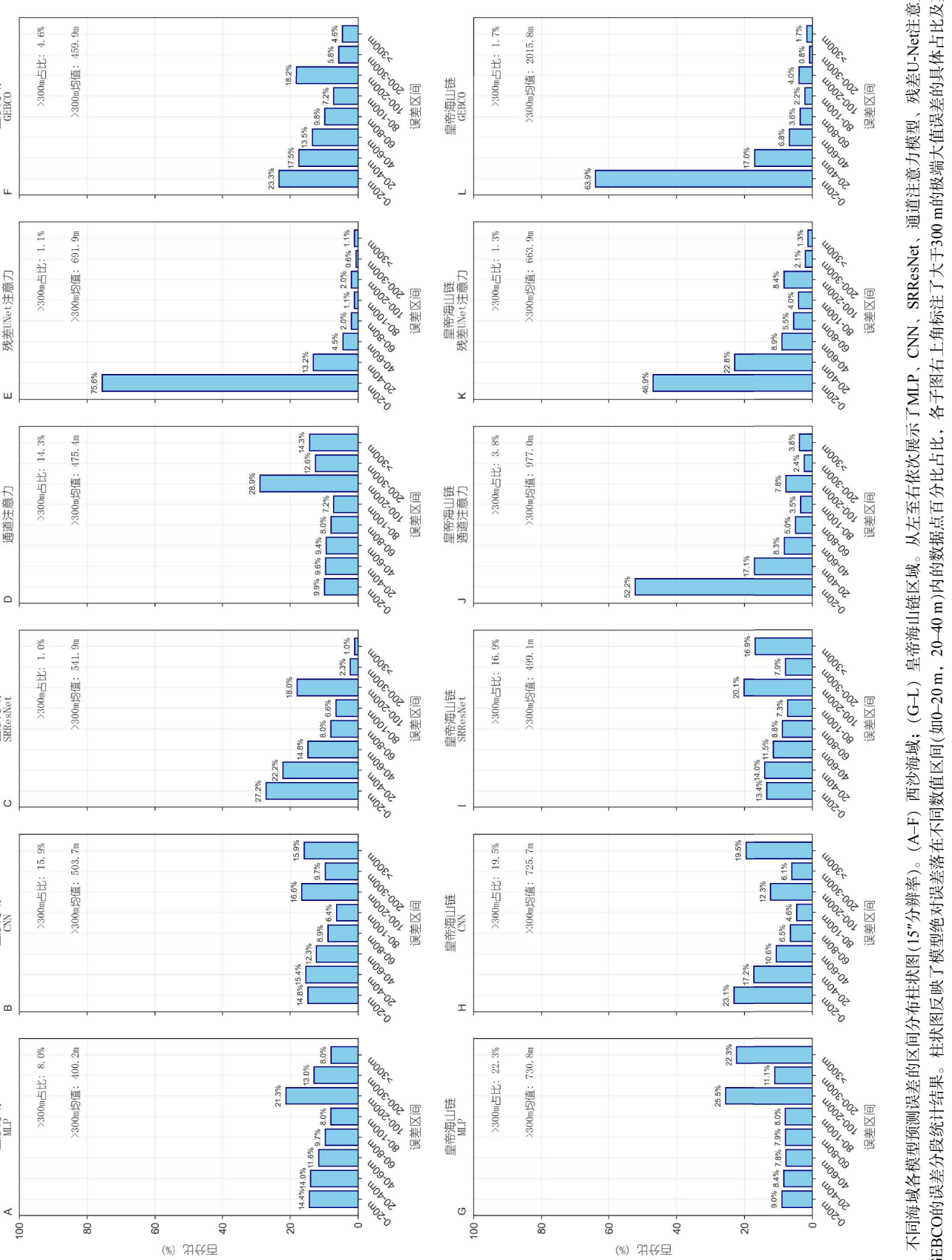


图 14 不同海域各模型预测误差的区间分布柱状图(15″分辨率)。(A–F) 西沙海域；(G–L) 皇帝海山链区域。从左至右依次展示了MLP、CNN、SRResNet、通道注意力模型、残差U-Net注意力模型与GEBCO的误差分段统计结果。柱状图反映了模型绝对误差落在不同数值区间(如0–20 m，20–40 m)内的数据点百分比占比，各子图右上角标注了大于300 m的极端大值误差的具体占比及其均值

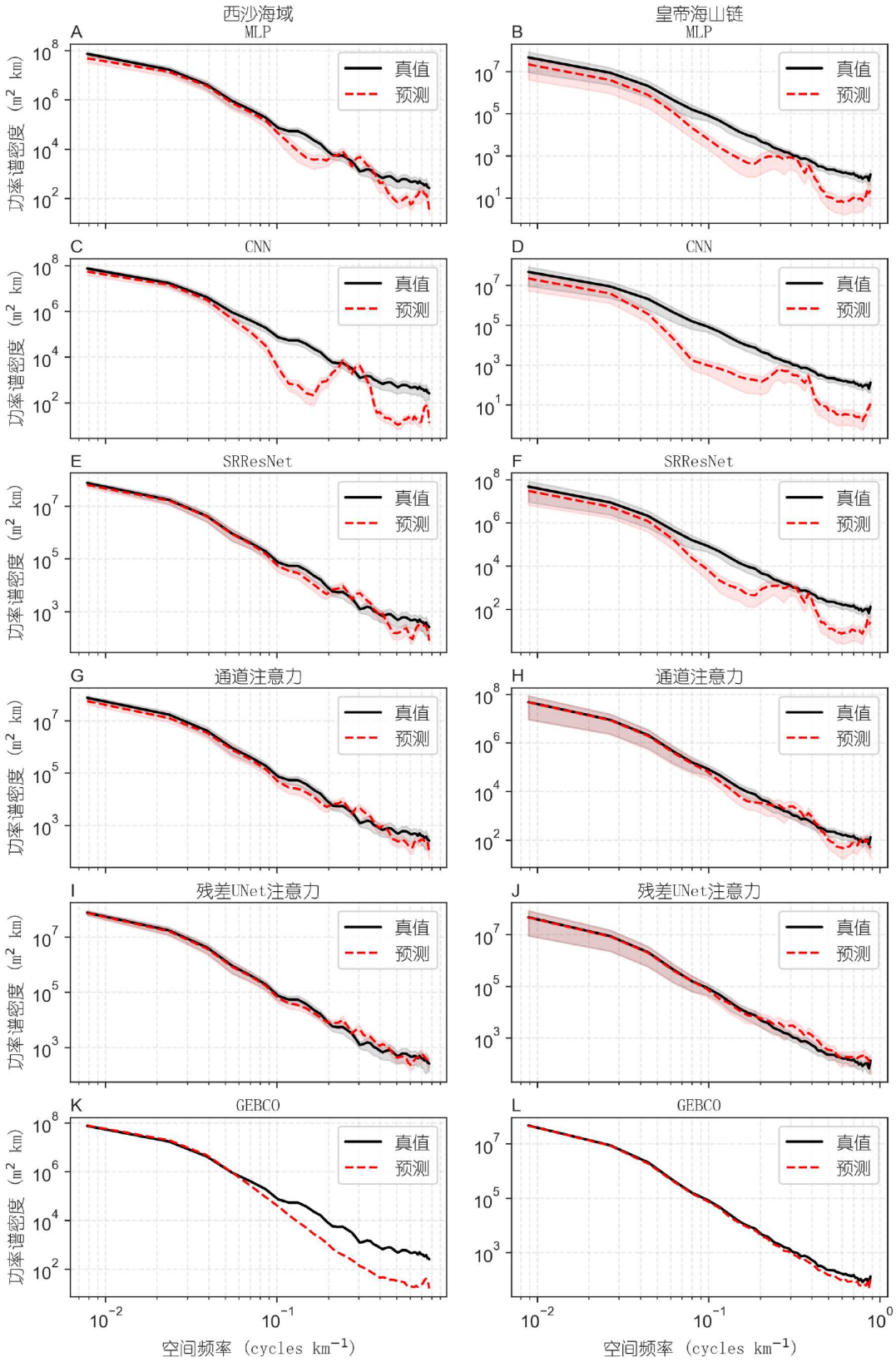


图 15　不同海域各模型重建地形的PSD对比图(30″分辨率)。(A、C、E、G、I和K) 西沙海域；(B、D、F、H、J和L) 皇帝海山链区域。从上至下依次为MLP、CNN、SRResNet、通道注意力模型、残差U-Net注意力模型以及GEBCO的PSD响应曲线。图中黑色实线代表多波束真值频谱，红色虚线代表模型预测频谱。对于神经网络模型，红色/黑色阴影区域分别表示利用多个滑动窗口计算所得频谱波动的标准差

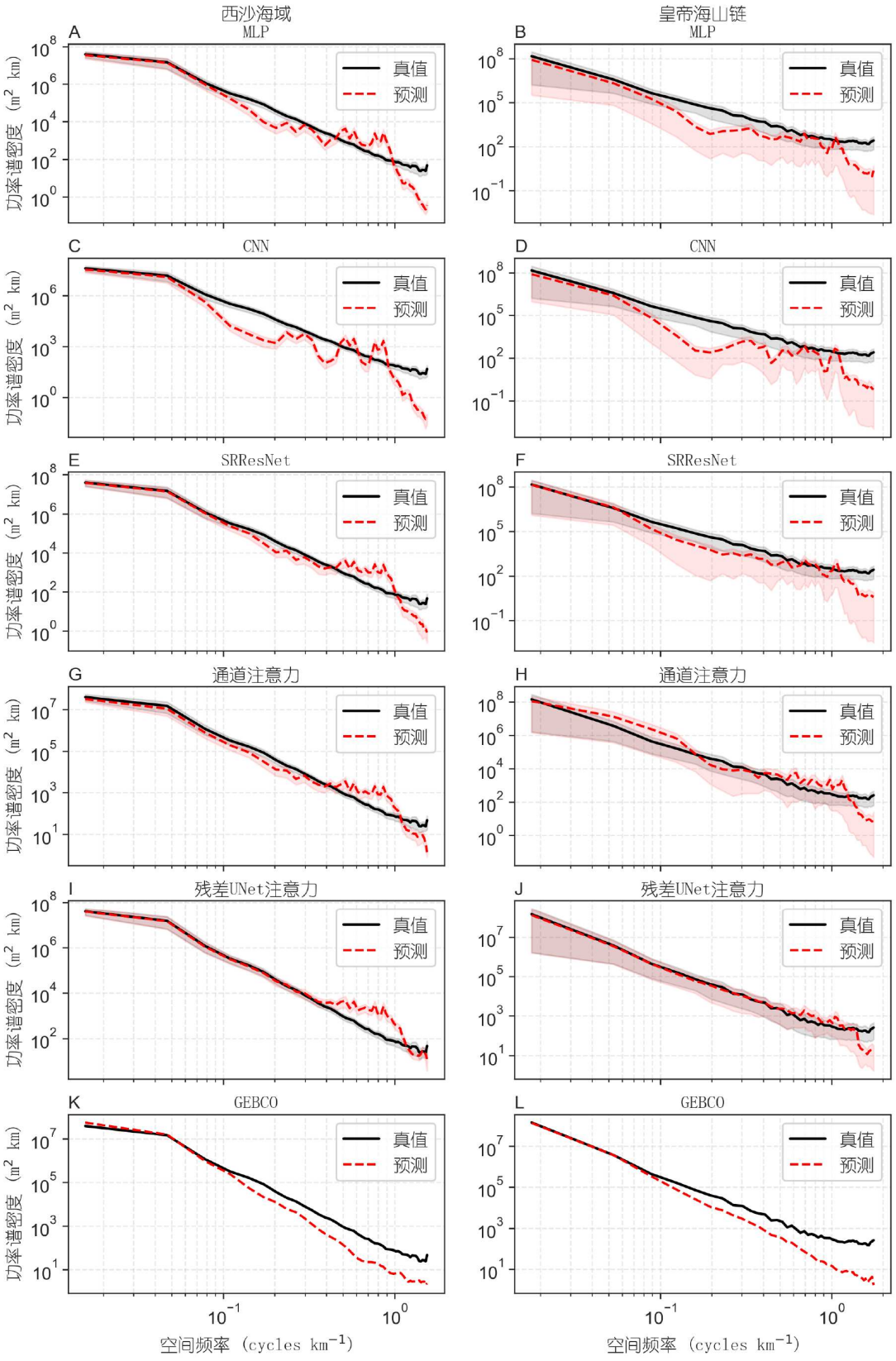


图 16 不同海域各模型重建地形的PSD对比图(15″分辨率)。(A、C、E、G、I和K) 西沙海域；(B、D、F、H、J和L) 皇帝海山链区域。从上至下依次为MLP、CNN、SRResNet、通道注意力模型、残差U-Net注意力模型以及GEBCO的PSD响应曲线。图中黑色实线代表多波束真值频谱，红色虚线代表模型预测频谱。对于神经网络模型，红色/黑色阴影区域分别表示利用多个滑动窗口计算所得频谱波动的标准差

表 7　不同海域各模型重要指标统计表(30″分辨率)

区域	模型名称	平均绝对误差 MAE(m)	均方根误差 RMSE(m)	深度平均值 Mean(m)	深度标准差 STD(m)	深度最小值 Min(m)	深度最大值 Max(m)
西沙海域	多波束探测(真值)	–	–	1226.3	483.48	0	2604.28
	GEBCO背景	85.82	131.29	1179.56	503.76	0	2621.00
	残差U-Net注意力模型	23.48	65.74	1228.65	475.08	0	2670.88
	SRResNet	61.63	87.97	1189.82	440.4	0	2442.41
	EDSR	102.67	154.79	1154.23	377.29	0	2224.31
	SRResNet大型模型	84.38	126.74	1176.59	397.24	0	2297.79
	EDSR大型模型	91.39	132.87	1156.85	398.98	0	2309.58
	残差U-NetASPP模型	21.28	53.21	1227.31	476.58	0	2568.5
	U-Net门控注意力模型	24.28	69.63	1227.94	474.3	0	2796.51
	扩张U-Net模型	23.15	64.9	1228.82	474.28	0	2617.24
夏威夷-皇帝海山链海域	多波束探测(真值)	–	–	5219.11	1154.25	280.30	7654.39
	GEBCO背景	58.82	318.60	5209.22	1189.62	283.00	7671.00
	残差U-Net注意力模型	52.62	141.75	5226.29	1136.02	346.91	7473.52
	SRResNet	157.02	254.26	5356.03	1011.80	518.44	6933.66
	EDSR	350.33	520.67	5559.02	936.18	925.91	7775.18
	SRResNet大型模型	377.74	667.01	5581.36	669.93	2027.98	7473.87
	EDSR大型模型	264.36	409.87	5472.70	921.03	998.80	7107.46
	残差U-NetASPP模型	51.89	145.02	5226.88	1147.41	291.62	7512.57
	U-Net门控注意力模型	57.65	147.94	5244.77	1130.17	361.39	7495.84
	扩张U-Net模型	58.86	148.68	5233.53	1117.48	221.96	7539.73

表 8　不同海域各模型重要指标统计表(15″分辨率)

区域	模型名称	平均绝对误差 MAE(m)	均方根误差 RMSE(m)	深度平均值 Mean(m)	深度标准差 STD(m)	深度最小值 Min(m)	深度最大值 Max(m)
西沙海域	多波束探测(真值)	–	–	1224.12	486.48	0	2604.28
	GEBCO背景	88.17	145.45	1179.99	504.40	0	2621
	残差U-Net注意力模型	24.49	86.84	1229.59	476.09	0	2670.88
	SRResNet	63.1	102.11	1190.38	441.24	0	2442.95
	EDSR	104.24	162.78	1154.64	377.88	0	2224.35
	SRResNet大型模型	85.86	135.77	1177.08	397.73	0	2297.83
	EDSR大型模型	92.96	142.36	1157.4	399.57	0	2309.62
	残差U-NetASPP模型	22.38	75.85	1228.09	477.33	0	2568.5
	U-Net门控注意力模型	25.16	84.82	1228.89	475.11	0	2796.51
	扩张U-Net模型	24.21	84.76	1229.65	475.11	0	2617.24
夏威夷-皇帝海山链海域	多波束探测(真值)	–	–	5218.83	1155.41	279.19	7710.41
	GEBCO背景	58.77	318.32	5208.68	1190.62	280.00	7696.00
	残差U-Net注意力模型	44.94	126.34	5225.36	1137.43	346.91	7504.72
	SRResNet	155.99	250.56	5355.75	1012.95	506.25	6940.37
	EDSR	350.07	520.18	5559.10	936.54	925.91	7778.63
	SRResNet大型模型	377.32	666.38	5581.17	670.90	2027.98	7475.35
	EDSR大型模型	263.85	408.13	5472.77	921.78	998.71	7110.11
	残差U-NetASPP模型	44.20	129.53	5225.83	1148.91	284.36	7566.00
	U-Net门控注意力模型	50.70	134.29	5243.85	1131.71	361.39	7495.84
	扩张U-Net模型	51.83	133.54	5232.76	1118.87	221.96	7539.73

海山链两个地形差异显著的研究区域，以残差U-Net注意力模型为核心，实现了2倍(30″)与4倍(15″)分辨率的海底地形背景场超分辨率优化。本研究的主要结论如下。

(1)显著提升背景场精度：相比原始GEBCO背景场，残差U-Net注意力模型在30″重建任务中使RMSE分别下降约49.9%(西沙)和55.5%(夏威夷-皇帝海山链)；在15″任务中分别下降约40.3%(西沙)和60.3%(夏威夷-皇帝海山链)。结果表明，本方案具备出色的背景场超分辨率优化能力，能够极大程度地逼近多波束实测真值。

(2)有效控制大值误差与高频恢复：通过误差统计与频域分析证实，基于窗口位移的超分方案显著减少了背景场的极端大值误差。以15″分辨率重建任务为例，在西沙海域，残差U-Net注意力模型将大于300 m的绝对误差占比由原始GEBCO的4.6%显著压缩至仅1.1%；而在地形起伏更为剧烈的夏威夷-皇帝海山链区域，GEBCO原始背景场中大于300 m的极端误差其平均偏离幅度高达2015.8 m，本方案成功将该极端误差均值大幅削减至663.9 m。这种在不同地形下对大值误差频率与幅度的双向控制，在统计学上证明了模型有效平抑了虚假的地形突变。此外，PSD评估显示，该方案在0.1–1 cycles km^{-1}频段内表现出更优的拟合特征，成功恢复了地形的高频细节起伏。

(3)明确模型工程优化路径：针对“结构扩展”、“多尺度与上下文增强”和“通道/注意力机制增强”三类路径的对比实验表明：ASPP与瓶颈注意力可同步建模地形的全局趋势与局部细节，高度适配复杂海域；通道注意力能有效抑制噪声，提升多源物理数据融合质量。相比之下，单纯扩大模型参数量在处理高噪、不规则采样的地形任务时收益有限，且徒增计算与存储开销。因此，工程实践中建议优先采用“注意力机制+多尺度增强”的轻量化改进策略。

当前构建的模型虽在特定海域得到了局地有效性与多地形鲁棒性的验证，但尚属局部区域的优化。未来工作将聚焦于模型跨区域泛化能力的提升；同时，针对模型在高频频段仍存在的微小偏差，计划在训练拟合过程中引入更为严格的频域物理理论公式(如Parker公式)作为物理约束条件，以期实现更加符合地球物理规律的全球海底地形高分辨率背景场优化。

致谢 斯克里普斯海洋研究所(Scripps Institution of Oceanography，SIO)提供SWOT卫星重力数据，广州海洋地质调查局(GMGS)提供西沙地区的多波束船载测深数据，美国国家环境信息中心(NCEI)提供夏威夷-皇帝海山链的多波束船载测深数据，GEBCO提供GEBCO2024海底地形模型作为本研究中使用的基础数据，在此一并衷心感谢。

参考文献

Amante C, Eakins B W. 2009. ETOPO1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis. Version 1.0. NOAA National Centers for Environmental Information

Andersen O B, Knudsen P. 1998. Global marine gravity field from the ERS-1 and Geosat geodetic mission altimetry. J Geophys Res, 103: 8129–8137

Andersen O B, Knudsen P. 2008. The DTU10 global Mean sea surface and Bathymetry. EGU-2008

Bahdanau D, Cho K, Bengio Y. 2015. Neural machine translation by jointly learning to align and translate. arXiv: 1409.0473

Cai W, Liu Y, Chen Y, et al. 2023. A seabed terrain feature extraction transformer for the super-resolution of the digital bathymetric model. Remote Sens, 15: 4906

Chen L C, Papandreou G, Kokkinos I, et al. 2017. Rethinking atrous convolution for semantic image segmentation. arXiv: 1706.05587

Fan D, Li S, Feng J, et al. 2022. A new global bathymetry model: STO_IEU2020. Remote Sens, 14: 5744

Ge B, Guo J, Kong Q, et al. 2025. Seafloor topography inversion from multi-source marine gravity data using multi-channel convolutional neural network. Eng Appl Artif Intell, 139: 109567

GEBCO Compilation Group. 2023. GEBCO 2023 grid: General bathymetric chart of the oceans

Harper H, Sandwell D T. 2024. Global predicted bathymetry using neural networks. Earth Space Sci, 11: e2023EA003199

He K, Zhang X, Ren S, et al. 2016. Deep residual learning for image recognition. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE. 770–778

Hu J, Shen L, Sun G. 2018. Squeeze and excitation networks. In: Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE. 7132–7141

Hu M, Li J, Li H, et al. 2015. A program for bathymetry prediction from vertical gravity gradient anomalies and ship soundings. Arab J Geosci, 8: 4509–4515

Ibrahim A, Hinze W J. 1972. Mapping Buried Bedrock Topography with Gravity. Groundwater, 10: 18–23

Irisawa N, Iiyama M. 2024. High-resolution bathymetry by deep-learning based point cloud upsampling. IEEE Access, 12: 4387–4398

Jayne S, St. Laurent L, Gille S. 2004. Connections between ocean bottom topography and Earth's climate. Oceanography, 17: 65–74

Lecun Y, Bottou L, Bengio Y, et al. 1998. Gradient-based learning applied to document recognition. Proc IEEE, 86: 2278–2324

Ledig C, Theis L. Huszár F. et al. 2017. Photo-realistic single image super-resolution using a generative adversarial network. In: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE. 4681–4690

Lim B, Son S, Kim H, et al. 2017. Enhanced deep residual networks for single image super resolution. arXiv: 1707.02921

Ma Y, Wu S, Lu F, et al. 2011. Seismic characteristics and development of the Xisha carbonate platforms, northern margin of the South China Sea. J Asian Earth Sci, 40: 770–783

Mayer L, Jakobsson M, Allen G, et al. 2018. The Nippon Foundation—GEBCO seabed 2030 project: The quest to see the world's oceans completely mapped by 2030. Geosciences, 8: 63

Nair V, Hinton G E. 2010. Rectified Linear Units Improve Restricted Boltzmann Machines. In: Proceedings of the 27th International Conference on Machine Learning. Israel: ICML. 807–814

Nettleton L L. 1939. Determination of density for reduction of gravimeter

observations. Geophysics, 4: 176–183
Parker R L. 1973. The rapid calculation of potential anomalies. Geophys J Int, 31: 447–455
Ronneberger O, Fischer P, Brox T. 2015. U-Net: Convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells W, et al., eds. Medical Image Computing and Computer-Assisted Intervention —MICCAI 2015. Cham: Springer. 234–241
Rumelhart D E, Hinton G E, Williams R J. 1986. Learning representations by back-propagating errors. Nature, 323: 533–536
Sandwell D T, Smith W H F. 1997. Improved resolution of bathymetry from multibeam sonar and satellite altimetry. Geophys Res Lett, 24: 119–122
Sandwell D T, Müller R D, Smith W H F, et al. 2014. New global marine gravity model from CryoSat-2 and Jason-1 reveals buried tectonic structure. Science, 346: 65–67
Sandwell D T, Goff J A, Gevorgian J, et al. 2022. Improved bathymetric prediction using geological information: SYNBATH. Earth Space Sci, 9: e2021EA002069
Smith W H F. 1993. On the accuracy of digital bathymetric data. J Geophys Res, 98: 9591–9603
Smith W H F, Sandwell D T. 1994. Global sea floor topography from satellite altimetry and ship depth soundings. Science, 277: 1956–1962
Sonogashira M, Shonai M, Iiyama M, et al. 2020. High-resolution bathymetry by deep-learning-based image superresolution. PLoS ONE, 15: e0235487
Srivastava N, Hinton G, Krizhevsky A, et al. 2014. Dropout: A simple way to prevent neural networks from overfitting. J Mach Learn Res, 15: 1929–1958
Tozer B, Sandwell D T, Smith W H F, et al. 2019. Global bathymetry and topography at 15 Arc Sec: SRTM15+. Earth Space Sci, 6: 1847–1864
Vaswani A, Shazeer N, Parmar N, et al. 2017. Attention is all you need. NeurIPS, 2017: 6000–6010
Wan X, Annan R F, Ziggah Y Y. 2023. Altimetry-derived gravity gradients using spectral method and their performance in bathymetry inversion using backpropagation neural network. J Geophys Res Solid Earth, 128: e2022JB025785
Watts A B. 1978. An analysis of isostasy in the world's oceans 1. Hawaiian-Emperor Seamount Chain. J Geophys Res, 83: 5989–6004
Wölfl A C, Snaith H, Amirebrahimi S, et al. 2019. Seafloor mapping—The challenge of a truly global ocean bathymetry. Front Mar Sci, 6: 283
Yang L, Liu M, Liu N, et al. 2023. Recovering bathymetry from satellite altimetry-derived gravity by fully connected deep neural network. IEEE Geosci Remote Sens Lett, 20: 1502805
Yu F, Koltun V. 2016. Multi-scale context aggregation by dilated convolutions. arXiv: 1511.07122
Yu Y, Sandwell D T, Dibarboure G. 2024. Abyssal marine tectonics from the SWOT mission. Science, 386: 1251–1256
Zhang B, Xiong W, Ma M, et al. 2022. Super-resolution reconstruction of a 3 arc-second global DEM dataset. Sci Bull, 67: 2526–2530
Zhou S, Liu X, Gao J, et al. 2023. Bathymetry of the Gulf of Mexico predicted with multilayer perceptron from multisource marine geodetic data. IEEE Trans Geosci Remote Sens, 61: 1–11
Zhou S, Liu X, Sun Y, et al. 2024. Predicting bathymetry using multisource differential marine geodetic data with multilayer perceptron neural network. Int J Digital Earth, 17: 2393255
Zhou S, Guo J, Zhang H, et al. 2025. SDUST2023BCO: A global seafloor model determined from a multi-layer perceptron neural network using multi-source differential marine geodetic data. Earth Syst Sci Data, 17: 165–179

（责任编委：姜卫平）