

doi:10.19306/j.cnki.2095-8110.2024.03.009

基于改进注意力机制 CNN-ATT 的 区域性 ZTD 预测模型

韦廖军¹, 莫 儒¹, 任晓斌², 任宏权³, 魏二虎⁴

1. 南宁市勘测设计院集团有限公司, 南宁 530000;
2. 奥克兰大学, 新西兰 奥克兰市 1010;
3. 自然资源部第一地形测量队, 西安 710054;
4. 武汉大学测绘学院, 武汉 430079)

摘要: 基于天顶对流层延迟(ZTD)的强时空特征, 提出了一种融合卷积神经网络的改进注意力机制(CNN-ATT)的多站点 ZTD 组合预测模型。该模型首次将多源数据(包括日解算精度、年积日(DOY)和三维坐标)综合运用用于 ZTD 预测任务。通过对南宁市的 5 个参考站(CORS)和 14 个国际 GNSS 服务(IGS)站点共 1 501 个年积日的观测数据进行研究, 选取传统 BP 模型、GPT2w 模型和 ATT 模型作为基线模型进行实验对比分析。研究结果显示, 在预测精度方面, 改进的 CNN-ATT 模型与 BP 模型相比其均方误差(MSE)和平均绝对误差(MAE)分别减少了 5.5 mm 和 4.4 mm, 预测精度分别提高了 41.4% 和 67.8%; 与 ATT 模型相比, CNN-ATT 模型的预测 MSE 和 MAE 也分别减少了 4.6 mm 和 2.1 mm, 预测精度分别提升了 36.2% 和 50.0%。在定位精度方面, 改进的 CNN-ATT 模型的精度表现优于 SAAS, GPT2w, BP 以及 ATT 模型。并且与传统 SAAS 对流层模型相比, CNN-ATT 模型在 N, E, U 3 个方向的精度提升高达 18.2%, 12.6% 和 31.0%。此外, 研究还发现 CNN-ATT 模型在长预测时间步长中的精度表现更为稳定, 更适合多测站预测任务, 并且其精密单点定位(PPP)收敛速度更快。

关键词: 注意力机制; 对流层延迟; 预测模型; 卷积神经网络

中图分类号: V249.3

文献标志码: A

文章编号: 2095-8110(2024)03-0085-16

A regional ZTD prediction model based on improved attention mechanism CNN-ATT

WEI Liaojun¹, MO Nuo¹, REN Xiaobin², REN Hongquan³, WEI Erhu⁴

1. Nanning Survey and Design Institute Group Co., Ltd., Nanning 530000, China;
2. University of Auckland, Auckland 1010, New Zealand;
3. The First Topographic Surveying Brigade of Ministry of Natural Resources, Xi'an 710054, China;
4. School of Surveying and Mapping, Wuhan University, Wuhan 430079, China)

Abstract: Based on the strong spatiotemporal characteristics of zenith tropospheric delays (ZTD), a multi-site ZTD combination prediction model with an improved attention mechanism based on convolutional neural networks (CNN-ATT) is proposed. The model integrates multiple data

收稿日期: 2023-11-06; 修订日期: 2024-01-17

基金项目: 国家自然科学基金(423740145); 天津市轨道交通导航定位及时空大数据技术重点实验室开放基金(TKL2024B04)

作者简介: 韦廖军(1979—), 男, 高级工程师, 主要从事城市规划测量方面的研究。

通信作者: 魏二虎(1965—), 男, 教授, 博士生导师, 主要从事空间大地学与导航方面的研究。

sources, including daily estimation accuracy, day of the year (DOY), and three-dimensional coordinates, for the first time in ZTD prediction tasks. A study is conducted using observation data from 5 reference stations (CORS) in Nanning and 14 International GNSS Service (IGS) stations, spanning a total of 1 501 DOY. Traditional back propagation (BP) models, global pressure and temperature 2wet (GPT2w) models, and ATT models are selected as baseline models for comparative analysis. The prediction results demonstrate that in terms of prediction accuracy, the improved CNN-ATT model outperforms traditional BP neural network models, with a reduction in mean squared error (MSE) and mean absolute error (MAE) by 5.5 mm and 4.4 mm respectively, leading to an improvement in prediction accuracy by 41.4% and 67.8%. Compared to the ATT model, the improved CNN-ATT model also shows reductions in MSE and MAE by 4.6 mm and 2.1 mm, respectively, resulting in a 36.2% and 50.0% enhancement in prediction accuracy. Regarding positional accuracy, the improved CNN-ATT model outperforms the SAAS, GPT2w, BP, and ATT model. Furthermore, when compared to the traditional SAAS tropospheric model, the CNN-ATT model achieves noteworthy accuracy improvements in the N, E and U directions, with enhancements of 18.2%, 12.6% and 31.0% respectively. Additionally, the research unveils that the CNN-ATT model exhibits a more stable performance in extended prediction time steps, making it particularly suitable for multi-station prediction tasks. Moreover, it manifests a faster convergence rate in precise point positioning (PPP) applications.

Key words: Attention mechanism; Tropospheric delay; Prediction model; Convolutional neural networks

0 引言

在卫星精准定位领域中,天顶对流层延迟(zenith tropospheric delay, ZTD)指的是卫星定位系统(如全球定位系统)中由于大气层中的水汽和电离层等因素引起的信号延迟,通常包括干延迟和湿延迟,它是影响全球导航卫星系统(global navigation satellite system, GNSS)定位精度的重要误差源,它甚至可以造成天顶方向 2 m 的干延迟量^[1]。近年来,随着 ZTD 预报模型的持续改进,大尺度 ZTD 预测的准确性显著提高。然而,面对区域性极端天气的频繁出现,如何提高区域性 ZTD 的预测准确性将对 GNSS 精密定位和大气水汽反演等领域具有重要意义。

目前,对流层延迟误差修正的模型有基于实测气象参数的 Hopfield 模型、Saastamoinen(SAAS)模型和 Black 模型,以及不基于实测气象参数的 UNB(University of New Brunswick)、UNB3(University of New Brunswick 3)、EGNOS(Euro Geostationary Navigation Overlay Service)和 GPT2w(global pressure and temperature 2wet)等模型。其中,基于实测气象参数的模型主要是利用目标位置的实测气象参数反映测站上空的气象变化,然后利用温度、压力及水汽压

等气象参数解算 ZTD,其 ZTD 误差改正精度可以达到分米级,甚至是厘米级^[2]。但由于所需的气象参数通常无法实时获取,从而制约了基于气象参数模型的广泛应用。基于非实测气象数据的模型主要利用全球或局部地区已有的气象观测资料和地理观测资料,通过线性拟合等方式获取相关气象参数和地理参数,并通过建立气象参数与地理参数之间的相关性继而进行 ZTD 的推算^[3-4]。例如,常用的 UNB 模型(UNB1~UNB4)和 EGNOS 模型等。尽管基于非实测气象数据的模型的提出提高了 ZTD 的预测精度。然而,考虑到 ZTD 解算中存在的复杂非线性特征(包括趋势信息和长短周期特征),以及此类模型(例如 UNB)过度依赖纬度参数,导致其在长时间序列预测任务中预测精度水平仍然有较大的提升空间。因此,近年来,许多研究致力于开发适用于区域或全球应用的多维网格对流层延迟模型^[5-6],例如 GPT 模型及其更新版本 GPT2, GPT3 以及 GPT2w。其中, GPT2w 作为更全面的对流层模型,融合了局部气象参数,并提供了气象参数的平均值、年度值以及半年度值,因此其预测精度比 UNB 模型更高^[5]。然而,虽然 GPT2w 作为全球广泛使用的经验模型较传统模型精度有所提升,但 GPT2w 在建模过程中对时间分辨

率和空间分辨率的参数考虑较为有限,因此 GPT2w 模型不能充分捕捉局部对流层信息,所以在 ZTD 预测任务中精度仍然存在较大的提升空间^[6]。

近年来,随着机器学习理论的不断发展和其在各行业的广泛应用,利用神经网络算法充分地挖掘海量数据规律已经成为目前的研究热点,其中基于神经网络模型的 ZTD 预测模型的相关研究也层出不穷^[7-9]。其中,反向传播(back propagation, BP)神经网络模型作为一种采用误差 BP 算法训练的多层前馈神经网络,已被应用于许多领域。在 ZTD 预测任务中,BP 神经网络模型通过将经度、纬度以及高差等数据作为输入来源,将 ZTD 的预测值作为预期输出目标。然后随机初始化权重和偏差参数,进行前向传播,计算损失函数,并通过 BP 和参数更新迭代训练。最后使用独立的验证集评估模型性能,并根据结果进行调优^[10]。BP 模型训练过程通常需要前向传递、BP 和参数更新的多次迭代,基于 BP 模型的 ZTD 预测模型的精度较传统模型显著提高。但是,值得注意的是传统的 BP 模型可能存在数据过拟合的问题^[11]。因此,肖恭伟等^[12]通过改进传统的 BP 模型构建了区域精度对流层模型。该模型的输入参数为经过归一化处理的大地纬度(B)、大地经度(L)以及大地高度(H),而其输出值则为基于 GAMIT/GLOBK 解算的 ZTD 值。与此同时,为了优化 BP 神经网络的建模精度,使用经验公式计算潜在隐藏层数目的范围,并通过大量历史的数据对模型参数进行初始化。最后,该研究将改进后的 ZTD 预测模型与常用的对流层区域实时模型进行了精度对比,结果表明改进后的 BP 神经网络构建的区域精度 ZTD 模型无论在拟合还是预报方面都存在更优的精度表现。此外,长短期记忆(long short-term memory, LSTM)神经网络模型因具有强大的长期记忆能力、防止梯度消失和梯度爆炸的特性以及灵活的建模能力被引入到 ZTD 预测领域^[13-14]。值得注意的是,传统的 LSTM 模型的训练时间较长,其自身的并行性较差,在对不同时间步长的特征提取方面优势不明显^[13]。因此,在输入时间步长较长或包含较多输入长周期项时,LSTM 预测精度无法保持稳定。为了进一步克服单一神经网络算法存在的不足,通过组合使用多种神经网络模型和算法往往可以获得更高的精度。例如,杨旭等^[13]提出了一种基于 BP 神经网络和 LSTM 的区域/单站 ZTD

组合预测模型。该模型以香港连续运行参考站(continuously operating reference stations, CORS)网络中的 18 个监测站的连续 14 天观测数据为例,进行了区域、单站及二者组合 ZTD 预测模型研究。结果表明,组合模型的精度较区域和单站模型分别提升了 17.2%和 18.4%。此外,随着机器学习理论的不断发展和 Bahdanau 等^[14]在 2014 年提出了注意力机制理论。注意力机制是一种模拟人类视觉注意过程的机制,能够使机器学习模型有选择性地关注输入数据的特定部分。它允许模型在处理序列数据时将不同的特征投影到查询(query)、值(value)和键(key)多个维度中,并通过点积计算获取不同时间步长的重要性权重,进而可以并行且高效地筛选出高价值的信息^[14-15]。所以理论上基于注意力机制的方法在长预测步长任务中优势较为明显。然而,如何将注意力机制较好地应用到 ZTD 的预测中仍是研究难点。

综上所述,随着 ZTD 预测模型的不断应用和发展,将神经网络模型更好地应用到 ZTD 精度校正与预测领域具有重要的理论价值和实际意义。为此,本文研究提出了融合卷积神经网络的改进注意力机制(convolutional neural network on attention mechanism, CNN-ATT)模型,从而创新性地将注意力机制应用于 ZTD 预测。提出的 CNN-ATT 模型可以有效地提取输入数据信息,并通过卷积神经网络对提取的输入数据进行多层特征提取与处理,最后通过改进的注意力层实现 ZTD 高精度预测。为了验证所构建 CNN-ATT 模型的有效性和准确性,选用南宁市 5 个 CORS 和 14 个国际 GNSS 服务(international GNSS service, IGS)站点进行区域性 ZTD 精密解算,并选用 BP 模型、GPT2w 模型和 ATT 模型作为基线模型进行对比实验。研究表明,CNN-ATT 模型可以实现更高的精度,并且在长时间序列任务中表现出更好的鲁棒性。

1 改进注意力机制的 CNN-ATT 预测模型

提出的 CNN-ATT 预测模型包括 CNN 特征提取层和改进注意力层,如图 1 所示。在输入层,首先根据 GAMIT/GLOBK 估计的测站观测信息和卫星轨道信息,提取测站的三维坐标、DOY 以及日解算精度均方误差(mean square error, MSE)。随后,将提取的测站信息传递至嵌入层,通过嵌入层将原始数据映射到多维空间中,以便提取隐藏特征。接

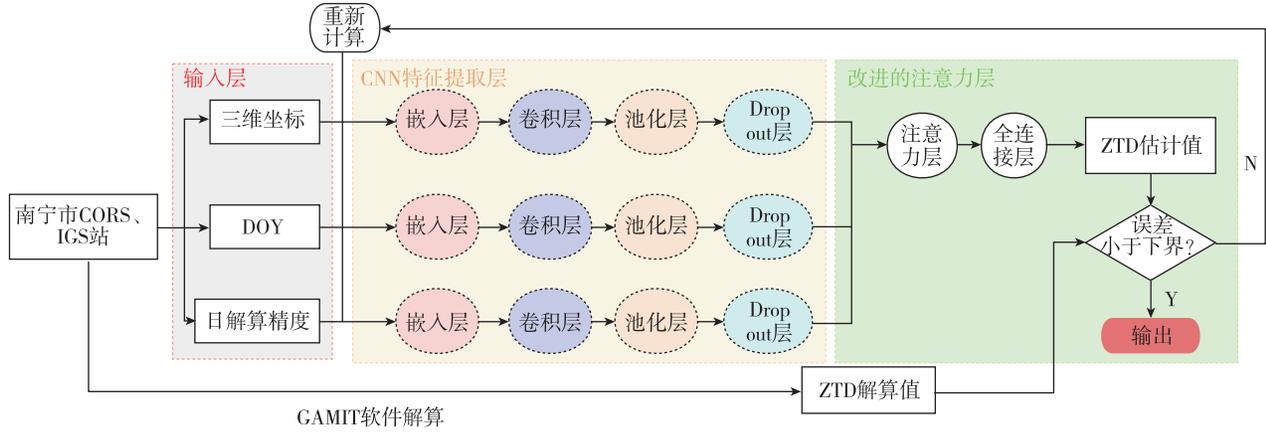


图1 CNN-ATT 模型框架

Fig. 1 The architecture of CNN-ATT framework

着,将初步提取的隐藏特征作为一维 CNN 模型的输入,用于融合多维变量。通过池化层减少特征维度,随后引入 dropout 层进行随机停止,以防止过度拟合并提高模型的泛化能力。然后,便可以将提取的特征引入注意力层,利用多头注意力机制对 dropout 层输出的特征序列进行训练,以获取时间序列数据的信息。通过以上处理可以保证模型能够集中学习关键时间步长的特征,以获取全面的特征信息。最后,通过全连接层输出模型的最终预测结果。如果整体预测误差超过误差下界,将重新进行训练,直至误差收敛。

鉴于 ZTD 预测任务与测站点的经纬度、大地高度以及时间等因素密切相关^[6,12],研究仍然将上述因素作为自变量考虑。同时,考虑到基于 GAMIT 处理得到的单日解算精度 MSE 与 ZTD 之间存在显著相关性,因此将日解算精度也纳入自变量之列。因变量为 ZTD 的预测值 Z_{ZTD} 。所以基于注意力机制的 ZTD 预测如下

$$Z_{ZTD,t+T+1:t+T+l}^{PMS} = \Phi(L_{t:t+T}, B_{t:t+T}, H_{t:t+T}, N_{t:t+T}, \delta_{t:t+T}) \quad (1)$$

其中, Φ 表示基于改进注意力机制的机器学习模型; N 表示年积日; δ 表示每日解算精度,即每日解算 MSE 值; t 表示当前时刻; T 表示时间窗口,指输入的时间序列长度; L 表示预测的时间序列长度。

1.1 输入层

考虑到现有模型仅以模型的三维坐标和 DOY 作为输入,且在区域性实时气象数据较少的情况下,该输入可能会导致模型估计误差较大^[13,15]。因此,利用区域性 CORS 提供的实时精密 GPS 数据的解算精度,构建 CNN-ATT 模型。提出的基于 CNN-ATT 的 ZTD 预测模型所需的输入数据包括站点的三维坐标、DOY 以及日解算精度。为此,利用 GLOBK 软件将包含 IGS 站和 NNCORS 网单日松弛解 h 文件,通过公共站点和卫星进行合并,得到一个包含所有 IGS 测站、NNCORS、极移和卫星参数的松弛解及方差协方差矩阵的 h 文件。将 h 文件代入 GAMIT/GLOBK 进行网平差后,可获得日解算精度,同时也可得到随时间变化的 IGS 测站和 NNCORS 的经纬度和大地高。

1.2 CNN 特征提取层

构建的 CNN 的特征提取层主要包括嵌入层、卷积层、池化层和 Dropout 层,如图 2 所示。其中,嵌入层用于将每个时间步长的特征映射到多个维度 d_{model} ,以供后续注意力层运算使用。然后,将处理后的数据特征作为输入进入卷积层。卷积层可提取时间序列中的局部特征,例如趋势、周期性或者局部噪声特征^[16]。基于 Zhou 等^[17]和 Yu 等^[18]

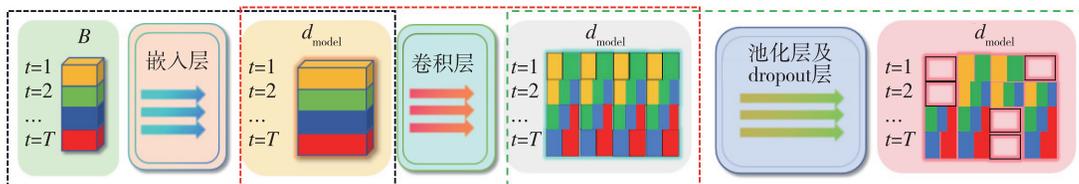


图2 CNN 特征提取层

Fig. 2 CNN feature extraction layer

利用一维卷积进行特征提取的步骤,进行了基于卷积层的特征提取。一维的卷积层通过按列聚合不同时间步长的特征信息,使得输出数据更为平滑。考虑到相邻时间点的信息聚合,经过卷积层提取的数据信息再次进入池化层和 Dropout 层,通过进一步的处理提取数据重要特征,并进一步防止数据过拟合现象的发生。

在 t 时刻, CORS 及 IGS 站的三维坐标为 $P_{t:t+T} = (L_{t:t+T}, B_{t:t+T}, H_{t:t+T})$, DOY 的数据输入格式为 $N_{t:t+T}$, 日解算精度为 $\delta_{t:t+T}$, 则三维坐标、DOY 以及日解算精度经嵌入层处理后的输出分别为

$$P_{\text{emb},t:t+T} = (L_{\text{emb},t:t+T} \in R^{T \times d_{\text{model}}}, B_{\text{emb},t:t+T} \in R^{T \times d_{\text{model}}}, H_{\text{emb},t:t+T} \in R^{T \times d_{\text{model}}})$$

$$N_{\text{emb},t:t+T} \in R^{T \times d_{\text{model}}}$$

$$\delta_{\text{emb},t:t+T} \in R^{T \times d_{\text{model}}} \quad (2)$$

将式(2)的输出结果作为卷积层的数据输入。采用一维的卷积层进行数据处理,卷积核的步幅为 1,通过卷积核实现以局部滑动平均的方式提取局部数据特征。然后,将卷积层的输出结果作为池化层和 dropout 层的输入,数据特征在卷积层被进一步压缩,并生成更重要的特征信息。值得注意的是,通过设置 dropout 层可以在每次更新期间将输入单元的一部分随机设置为零以防止过拟合,有助于通过减少对特定特征的依赖提高模型的泛化能力和鲁棒性,从而极大地提升数据处理效率。通过卷积层、池化层和 dropout 层的依次处理,三维坐标、DOY 以及日解算精度最终的输出结果为

$$H_{\text{POS},t:t+T} = \text{Dropout}(\text{MaxPool}(\text{Conv1d}(P_{\text{emb},t:t+T})))$$

$$H_{\text{DOY},t:t+T} = \text{Dropout}(\text{MaxPool}(\text{Conv1d}(N_{\text{emb},t:t+T})))$$

$$H_{\delta,t:t+T} = \text{Dropout}(\text{MaxPool}(\text{Conv1d}(\delta_{\text{emb},t:t+T}))) \quad (3)$$

式中, Conv1d 表示一维卷积层; MaxPool 表示最大池化层; $H_{\text{POS},t:t+T}$ 表示位置向量的隐藏状态; $H_{\text{DOY},t:t+T}$ 表示 DOY 的隐藏状态; $H_{\delta,t:t+T}$ 表示日解算精度的隐藏状态。

1.3 改进的注意力层

考虑到注意力机制能有效地挖掘存在关联的长时间序列特征,避免了 LSTM 和递归神经网络(recurrent neural network, RNN)等模型在处理超长序列时可能出现的信息丢失问题^[19-21],采用多头注意力机制提取数据特征信息,如图 3 所示。多头

注意力机制允许不同头独立地关注输入序列的不同部分,因此能够并行地处理输入序列,提高计算效率;此外,多头注意力机制模型能够确保同时关注不同的表示子空间,从而捕捉到输入序列中不同层次和粒度的特征信息,从而提高模型对于输入序列复杂关系的建模能力。因此,通过结合多头注意力机制可以更为精准地提取学习特征,达到更好的时间序列分析效果。

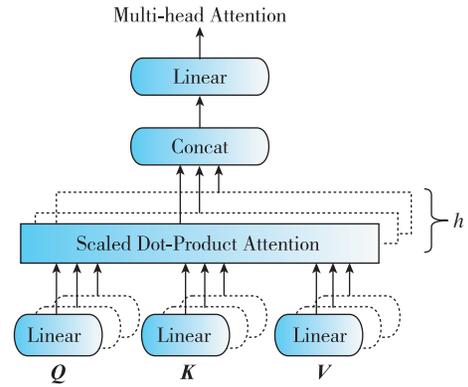


图 3 多头注意力机制结构图

Fig. 3 The structure of multi-head attention mechanism

图 3 显示了注意力机制的基本原理和处理流程。具体而言,本文采用的多注意力机制通过 3 个线性转换将输入向量映射到 3 个维度为 d_{model} 的输出矩阵,包括查询 Q (query)、键 K (key) 以及值 V (value)。改进后的注意力机制的输出矩阵如下

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_{\text{model}}}}\right)V \quad (4)$$

式中, d_{model} 表示特征维度,进行权重缩放,通过 softmax 层映射到 $[0, 1]$ 。

然后,多注意力机制通过 h 个头将时间序列分割成 h 个子空间,各个子空间独立地进行自注意力计算,以增强模型的表示能力。其中,多头注意力中的第 i 个头 m_i 计算表达式如下

$$m_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V) \quad (5)$$

式中, $Q \in R^{T \times d_{\text{model}}}$, $K \in R^{T \times d_{\text{model}}}$, $V \in R^{T \times d_{\text{model}}}$; $W_i^Q \in R^{d_{\text{model}} \times d_{\text{model}}}$, $W_i^K \in R^{d_{\text{model}} \times d_{\text{model}}}$, $W_i^V \in R^{d_{\text{model}} \times d_{\text{model}}}$ 分别表示 Q, K, V 的权重矩阵。

然后,将 h 个头的计算结果进行连接,经过线性变换后得到最终的输出结果 Multihead(Q, K, V),其表达式如下

$$\text{Multihead}(Q, K, V) = \text{Concat}(m_1, m_2, \dots, m_h)W^0 \quad (6)$$

式中, \mathbf{W}^0 为线性变换的权重矩阵; Concat 表示拼接操作。

在 ZTD 的预测任务中, 注意力层的输入向量为融合三维坐标、DOY 和日解算精度后的向量 $\mathbf{X}_{\text{emb},t:t+T}$, 表示如下

$$\mathbf{X}_{\text{emb},t:t+T} = \text{MLP}(\text{Concat}(H_{\text{POS},t:t+T}, H_{\text{DOY},t:t+T}, H_{\delta,t:t+T})), \in \mathbb{R}^{T \times d_{\text{model}}} \quad (7)$$

式中, MLP 由 2 个全连接层构成。经过线性转换后, $\mathbf{Q}, \mathbf{K}, \mathbf{V}$ 分别为

$$\begin{aligned} \mathbf{Q} &= \text{Linear}(\mathbf{X}_{\text{emb},t:t+T}) \\ \mathbf{K} &= \text{Linear}(\mathbf{X}_{\text{emb},t:t+T}) \\ \mathbf{V} &= \text{Linear}(\mathbf{X}_{\text{emb},t:t+T}) \end{aligned} \quad (8)$$

经过上述处理后, ZTD 的最终估计值为

$$Z_{\text{ZTD},t+T+1,t+T+1}^{\text{pre}} = \text{MLP}(\text{Att}(\mathbf{Q}, \mathbf{K}, \mathbf{V})) \quad (9)$$

获得初步的 ZTD 预估值后, 需要进一步判断预估值和真实值之间的误差水平。为此, 使用克拉默-拉奥下界(Cramer-Rao lower bound, CRLB) 方法判断误差是否小于下界, 利用 CRLB 确定误差下界的表达式为

$$\text{Var}(\hat{\theta}) \geq \frac{1}{nI(\theta)} \quad (10)$$

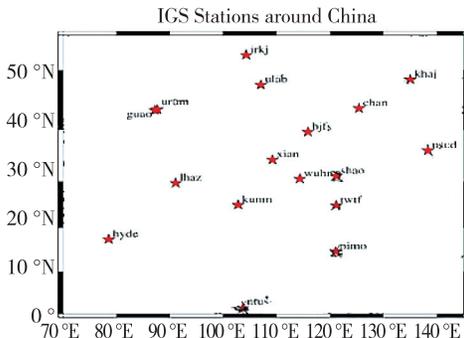
其中, θ 表示 CNN-ATT 中待估计的参数; n 表示观测样本数; $I(\theta)$ 表示 Fisher 信息, 是指参数 θ 关于观测数据的二阶导数的期望值。通过式(10)判断误差的下界是否满足要求, 如果不满足, 则继续训练, 直至收敛; 反之, 则输出最终结果。

2 实验结果与讨论

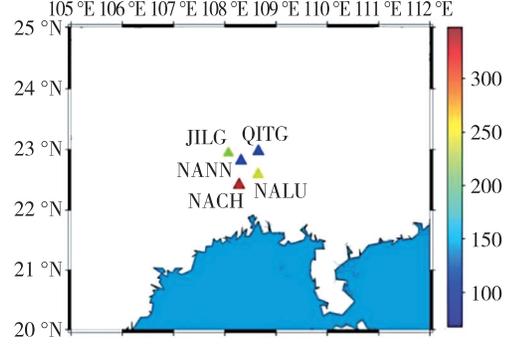
2.1 数据的来源及精度评判标准

2.1.1 实验参数设置

利用南宁市 5 个 CORS(NNCORS) 和 14 个 IGS 站的观测数据作为重要的输入数据, 如图 4 所示。其中, 5 个 CORS 的地理位置信息见表 1。在



(a) IGS 观测站地理分布图



(b) NNCORS 分布图

图 4 IGS 观测站和 NNCORS 分布图

Fig. 4 Distribution of IGS and NNCORS

表 1 CORS 地理位置数据

Tab. 1 The geographic location data of CORS

站点	纬度/(°)	经度/(°)	大地高/m
QITG	22.96	108.64	68.0
NALU	22.61	108.63	235.1
NANN	22.84	108.32	82.6
NACH	22.45	108.28	348.2
JILG	22.92	108.04	212.8

选用 CORS 的基础上, 还利用 14 个 IGS 站的数据增加基线长度, 从而提升 NNCORS 位置的解算精度。然后, 将 14 个 IGS 站和 2 个 CORS(NACH 和 NALU) 的数据用作模型训练, 3 个 CORS(JILG, NANN 和 QITG) 的数据用作模型预测, 从而进行精度评定。

所采用的编程语言为 Python 3.11.4。开发环境为运行在 Windows 10 操作系统下的 PyTorch 1.13.1 框架。硬件配置为 Google Colab 下的 TPU V2。在训练过程中, 损失函数被设定为 MSE, 优化器选择为 Adam。初始化学学习率为 0.1, 输入序列长度设定为 24, 输出序列长度为 12, 注意力头数设置为 8, 特征维度 d_{model} 设置为 64, 卷积核设置为 1, dropout 率设置为 0.05, batch size 设定为 32, epoch 为 30。考虑到参数的复杂调整和初始化, 研究中共涉及 104 个参数, 包括嵌入层权重 θ_1 、卷积层权重 θ_2 、改进的注意力层 ω_1 和 dropout 层 ω_2 。为了获取最佳的参数组合, 采用格网搜索的方法。

2.1.2 GAMIT/GLOBK 数据处理

GAMIT (GPS analysis at MIT)/GLOBK (Global Kalman filter) 是美国麻省理工学院 (Massachusetts Institute of Technology, MIT) 与斯克里普斯海洋研究所 (Scripps Institution of Oceanog-

raphy, SIO) 联合研制的高精度 GPS 定位软件。GAMIT 软件采用双差观测测量及最小二乘算法进行参数估计。GLOBK 软件主要采用多种卡尔曼滤波器进行数据后处理,在长基线解算中精度可达 $10^{-7} \sim 10^{-9}$ 量级。因此,GAMIT/GLOBK 在 ZTD 预测中应用广泛^[22-24]。

基于 GAMIT/GLOBK 数据的处理有两个主要目的:1) 为输入层提供准确的日解算精度,作为 CNN-ATT 模型输入数据的一部分;2) 提供 ZTD 的解算值,作为 CNN-ATT 模型精度检验的真实值。首先对 5 个 NNCORS 和远距离的 14 个 IGS 站的 RINEX 观测数据进行联合基线解算,基线解算的策略如表 2 所示。

表 2 基线解算策略

Tab. 2 The baseline solving strategy

序列	名称	参数
1	解算处理模式	RELAX
2	卫星星历	IGS 精密星历
3	光压模型	BERNE
4	历元间隔	30 s
5	卫星截止高度角	10°
6	观测模式	LC-AUTCLN
7	天顶延迟改正模型	VMF1
8	天顶延迟参数个数	13
9	对流层延迟	SAAS

解算时间跨度为 2009 年年积日 305~2013 年年积日 345,共计 1 501 个年积日。然后,对基线解算结果进行基于最小二乘法的联合网平差,在网平差后分别输出日解算精度 $\delta_{t,t+T}$ 、DOY、三维坐标(NNCORS 的经纬度和大地高)以及 ZTD 的估计值 $Z_{ZTD,t+T+1,t+T+l}^{pre}$ 。详细的 GAMIT/GLOBK 解算步骤如图 5 所示。具体而言,首先通过将原始 RINEX 文件输入 GAMIT 软件,设定 NNCORS 和 IGS 测站以及坐标,按照表 2 的策略创建基线解算的配置文件,进行基线解算。在完成基线解算后,分别执行相位平滑和模糊度固定工作,随后输入 GLOBK 进行参数的强约束。最终,可以输出 N,E,U 3 个方向的日解算精度。此外,为了获得用于模型精度检测的 ZTD 解算,在相位平滑处理后,还需要选择合适的大气模型并建立对流层投影函数映射,最终可以得到基于 GAMIT/GLOBK 的 ZTD 估计值^[24-25]。

按照图 5 步骤处理后,可以得到 19 个站点的日解算精度。其中,IGS 站点之一 BJFS 和 NNCORS 之一 NANN 的结果如图 6 所示。

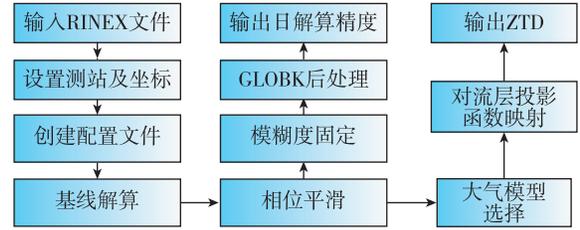
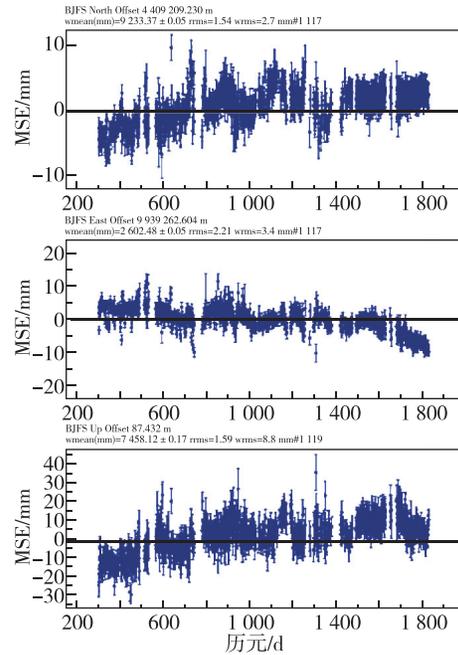
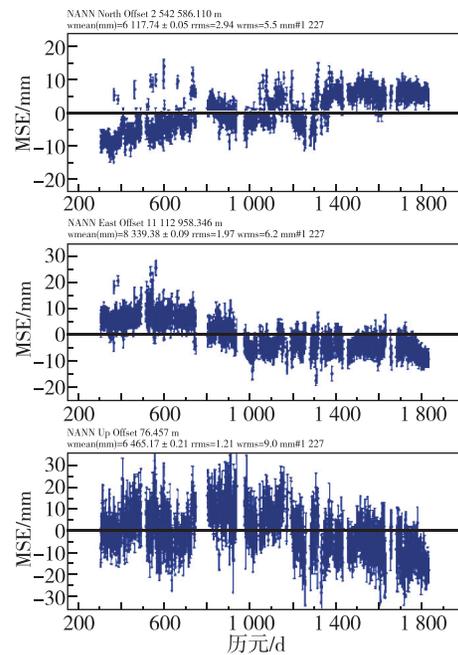


图 5 GAMIT/GLOBK 解算日解算精度步骤

Fig. 5 The procedure of daily solution accuracy results on GAMIT/GLOBK



(a) IGS 站点-BJFS



(b) NNCORS-NANN

图 6 BJFS 与 NANN 的日解算精度结果

Fig. 6 Daily solution accuracy results of BJFS and NANN

由图6可知,平差后NANN站点的平均精度在N水平方向优于3 mm,E方向优于2 mm,U方向优于1.3 mm;BJFS站点的平均精度在N水平方向优于1.6 mm,E方向优于2.3 mm,U方向优于1.6 mm。为了进一步对比解算后的日解算精度水平,统计了19个站点在N,E,U3个方向的平均解算精度,结果如表3所示。由此可知,NNCORS的平均精度在E,U方向均优于2 mm,N方向优于2.5 mm;IGS站点的平均精度则在N,E,U方向均优于2 mm。

表3 站点的平均解算精度结果

Tab. 3 Average solution precision results of stations

年份	NNCORS站平均解算精度			IGS站平均解算精度		
	N	E	U	N	E	U
2009	2.13	1.89	1.94	0.84	0.98	1.09
2010	1.98	1.96	1.92	0.94	0.99	1.12
2011	2.06	1.93	1.98	0.87	1.21	1.10
2012	2.12	1.83	1.93	0.92	1.16	1.12
2013	2.08	1.97	1.87	0.89	1.14	1.15
平均	2.07	1.92	1.93	0.89	1.10	1.12
R^2	0.66	0.68	0.62	0.53	0.54	0.49

由于目前现有模型并未将日解算精度考虑在ZTD预测中,为了进一步确保经GAMIT/GLOBK处理后的日解算精度可以作为构建的CNN-ATT模型的输入,对表3中各站点的平均解算精度和基于GAMIT/GLOBK解算的ZTD估计值 $Z_{ZTD,t+T+1,t+T+1}^{pre}$ 进行了相关性验证。若相关性较强,则表明经GAMIT/GLOBK处理后的各站点日解算精度可以很好地代表原始数据中的主要特征和信息,因此可以将日解算精度作为重要的输入数据代入CNN-ATT模型中进行模型的训练与测试;反之,若相关性较弱,则说明处理后的日解算精度不适合作为模型的输入数据。从表3可以看出,相关性 R^2 处于0.5~0.8的区间,表明经GAMIT/GLOBK处理后的站点平均日解算精度与ZTD估计值存在中度相关性。因此,将GAMIT/GLOBK处理后的日解算精度作为模型的重要输入来源之一是合理的。

2.1.3 精度评定准则

选取了平均偏差(bias)、平均绝对误差(mean absolute error, MAE)和MSE作为精度评价的标准。其中,bias反映了整体的系统误差情况,而

MAE和MSE则反映了预测值与真实值之间的差距程度。3个误差参数的计算公式为

$$\begin{aligned}
 e_{bias} &= \frac{1}{N} \sum_{i=1}^N (Z_{ZTD,i}^{pre} - Z_{ZTD,i}^{true}) \\
 e_{MSE} &= \frac{1}{N} \sum_{i=1}^N (Z_{ZTD,i}^{pre} - Z_{ZTD,i}^{true})^2 \\
 e_{MAE} &= \frac{1}{N} \sum_{i=1}^N |Z_{ZTD,i}^{pre} - Z_{ZTD,i}^{true}| \quad (11)
 \end{aligned}$$

其中, e_{bias} 、 e_{MSE} 、 e_{MAE} 分别表示平均偏差、MSE和MAE; N 代表预测样本数量; $Z_{ZTD,i}^{pre}$ 和 $Z_{ZTD,i}^{true}$ 分别代表ZTD的预测值和GAMIT解算ZTD的值。

2.2 不同ZTD预测模型精度分析

2.2.1 模型拟合精度

为了充分验证所构建的CNN-ATT模型的可靠性和有效性,选择了GPT2w,BP以及ATT模型作为基线模型。在模型拟合过程中,假定GPT2w模型的拟合是无误差的。不同ZTD预测模型的拟合结果如表4所示。由此可知,在拟合偏差方面,CNN-ATT的拟合偏差较ATT模型和BP模型分别减少了25%和50.9%,这表明CNN-ATT模型的拟合精度更高。此外,在拟合MSE和MAE方面,CNN-ATT模型的拟合MAE和MSE与BP模型相比分别减少了7.1 mm和5.5 mm;与ATT模型相比,CNN-ATT模型的拟合精度也提高了35.1%和32.3%。为了进一步降低算法的运行误差,对3个模型重复运行5次,并统计了3种误差的标准差。结果显示,CNN-ATT模型的误差标准差较BP模型降低了44.4%,较ATT模型降低了37.5%。由此可知,CNN-ATT模型的拟合精度表现明显优于BP模型和ATT模型。

表4 3种模型的拟合结果

Tab. 4 Fitting results of three models

	BP	ATT	CNN-ATT
Training bias	-5.5±0.2	-3.6±0.2	-2.7±0.1
Training MAE	11.9±0.4	7.4±0.3	4.8±0.2
Training MSE	9.2±0.3	5.5±0.3	3.7±0.2

图7为14个IGS站和2个NNCORS中的平均拟合MAE和MSE的结果。为了便于观察,该图截取了1501个时间步长中的前100个时间步长进行了可视化。从图7(a)可以看出,BP模型和ATT模型的拟合MAE的均值均高于CNN-ATT模型,且拟合的方差也高于CNN-ATT模型,甚至超过

30 mm。相比之下, CNN-ATT 模型在 0~100 个时间步长内的 MAE 始终未超过 5 mm, 模型的精度表现和鲁棒性表现更佳。这主要是由于改进后的 CNN-ATT 模型设置了多个卷积层和嵌入层, 同时还引入了改进的误差判断层。这使得模型能够在多个维度上对输入数据进行卷积, 从而不仅能够最大程度地提取有效的数据特征, 同时也能够学习到多维度输入数据与 ZTD 之间的复杂映射关系。相比之下, BP 神经网络模型则无法有效学习复杂的映射关系, 因此其对观测数据的拟合效果较差。

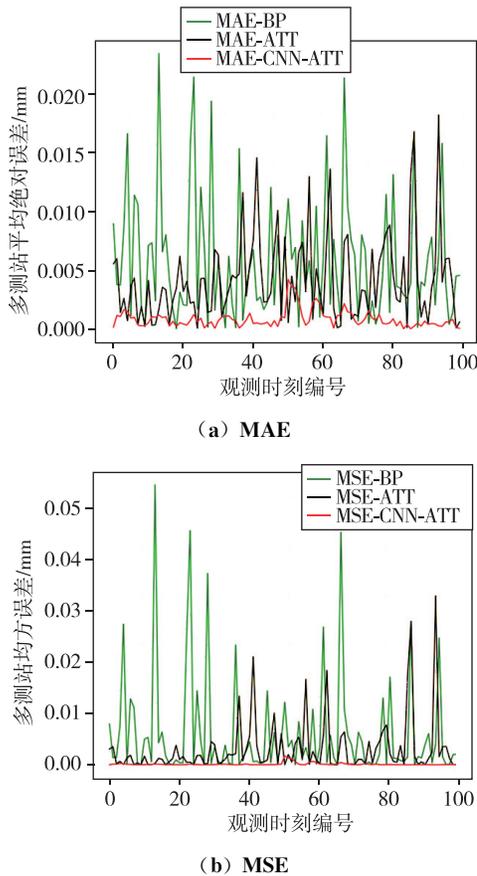


图 7 模型拟合精度结果

Fig. 7 Results of model fitting accuracy

2.2.2 模型预测精度

图 8 展示了在预测步长为 100 时, JILG, NANN 和 QITG 站点的 MAE 和 MSE 的分布结果。从图 8 可以看出, GPT2w 模型的误差水平较高且波动较大。这主要是因为 GPT2w 模型作为误差基线模型无需训练, 因此其模型性能与原始数据的质量密切相关。此外, 可以发现在观测时间步长为 20~40 时, GPT2w 模型的 MAE 和 MSE 波动较大, 精度水平较低。造成这种波动性的原因可能是: 1) GPT2w 模型的先验公式本身就无法准确拟合复杂的气象

参数^[23]; 2) 观测时间步长为 20~40 时, 该时间范围属于夏季, 导致 ZTD 的变化幅度较大, 从而使得经验模型 GPT2w 的预测精度较其他模型较低^[24-25]。相比之下, 在研究预测步长为 0~100 时, CNN-ATT 模型相对于 BP 神经网络模型和 ATT 模型表现出了较强的鲁棒性, 其误差水平也更低。这表明 CNN-ATT 模型可以实现更高的预测精度。

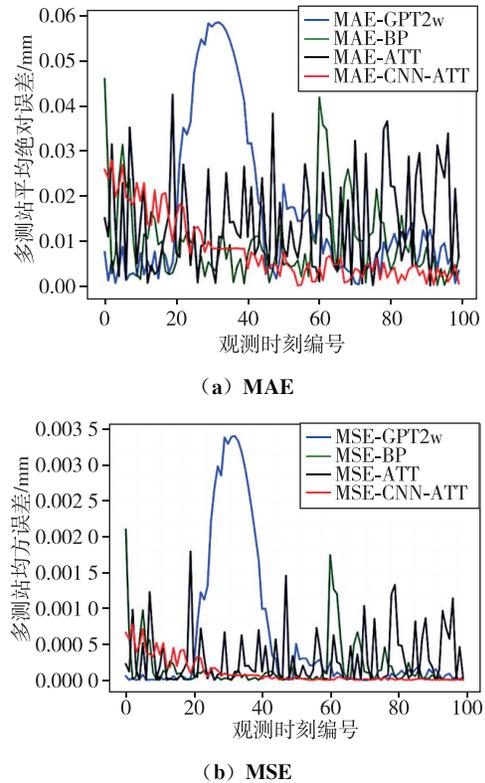


图 8 模型预测精度结果

Fig. 8 Model prediction accuracy results

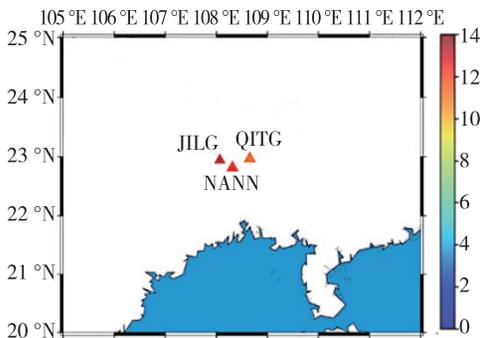
为了更加可靠地对比不同 ZTD 预测模型的预测精度水平, 统计了 4 个 ZTD 预测模型在 1 501 个时间步长下的平均偏差、MAE 和 MSE。为了进一步消除潜在的计算误差, 模型的预测结果为模型运行 5 次后的误差的平均值, 结果如表 5 所示。与 GPT2w 模型相比, CNN-ATT 模型在 MAE 和 MSE 方面的精度分别提升了 47% 和 93%; 与 BP 模型相比, CNN-ATT 模型的预测 MAE 和 MSE 分别降低了 5.5 mm 和 4.4 mm, 精度分别提升了 41.4% 和 67.8%。而相对于 ATT 模型, CNN-ATT 在结合多种卷积层和改进的误差判断层后, MAE 和 MSE 分别降低了 4.6 mm 和 2.1 mm, 精度分别提升了 36.2% 和 50.0%。此外, 在预测偏差方面, CNN-ATT 模型的表现达到了 -3.6 mm, 较 BP 模型和 ATT 模型精度分别提升了 59.1% 和

32.1%。但是 GPT2w 模型的预测偏差优于 BP 模型, ATT 模型以及 CNN-ATT 模型。这主要是因为 GPT2w 模型无需训练样本, 因此模型的输出结果和输入数据之间的偏差较小。此外, 进一步统计了 3 个误差(MAE、MSE 和预测偏差)的标准差, 结果显示, 与 BP 模型和 ATT 模型相比, CNN-ATT 模型的误差标准差分别降低了 71.4% 和 33.3%。所以整体而言, CNN-ATT 模型可以实现更为理想的预测精度。

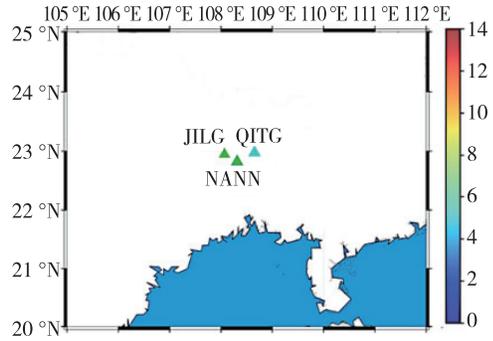
表 5 4 种模型的预测结果

Tab. 5 Prediction results of four models mm				
	GPT2w	BP	ATT	CNN-ATT
预测 bias	-0.098	-8.8±0.7	-5.3±0.4	-3.6±0.2
预测 MAE	14.9	13.4±0.7	12.7±0.3	7.9±0.2
预测 MSE	30.6	6.5±0.6	4.2±0.3	2.1±0.1

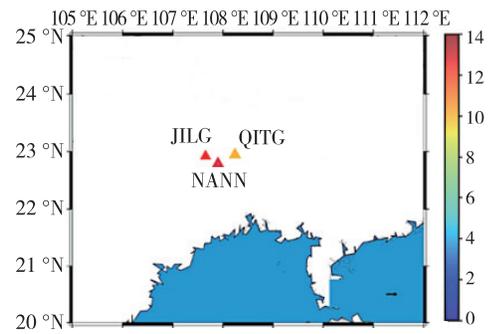
为了进一步对比不同的 ZTD 预测模型预测精度的差异, 统计了 CNN-ATT 模型和传统的 BP 模型的预测精度在空间上的分布情况, 如图 9 所示。如图 9 (a)和(c)所示, 在模型训练精度的空间分布中, 基于 BP 神经网络的预测中, JILG 站的预测精度 MAE 和 MSE 相较 NANN 和 QITG 站点更差。这可能是因为 JILG 站的大地高高于其他 2 个站点, 并且该地区的气候条件更为复杂, 导致该站点的预测精度较低。相比之下, 图 9(b)和(d)显示了 CNN-ATT 模型在不同纬度和大地高条件下, 各测站的预测精度 MAE 和 MSE 均在 8 mm 以下。其中, 与 BP 模型相比, CNN-ATT 模型预测下的 NANN 站的 MAE 和 MSE 分别提升了 42.9% 和 37.8%, JILG 站的 MAE 和 MSE 分别提升了 30.4% 和 32.9%, QITG 站的 MAE 和 MSE 分别提升了 34.7% 和 37.2%。由此可知, 相较于传统的 BP 神经网络模型, CNN-ATT 模型的预测精度在空间分布上也表现较好。



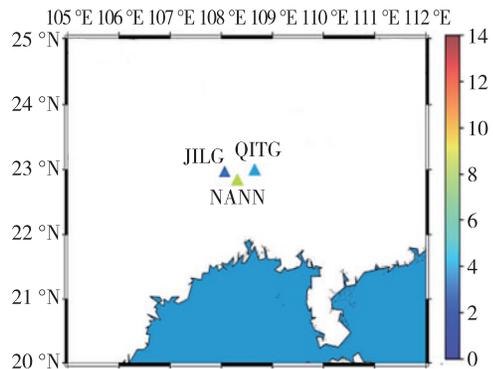
(a) 基于 BP 模型的 MAE 结果



(b) 基于 CNN-ATT 模型的 MAE 结果



(c) 基于 BP 模型的 MSE 结果



(d) 基于 CNN-ATT 模型的 MSE 结果

图 9 模型预测精度的空间分布

Fig. 9 Spatial distribution of model prediction accuracy

此外, 为了验证 CNN-ATT 在多时间步长预测任务中的表现, 以 MAE 为例验证 BP 模型和 CNN-ATT 模型在不同预测时间步长(24 h 和 48 h)下的精度差异。如图 10 所示, 随着预测步长从 24 h 增加至 48 h, 基于 BP 模型的 NANN 站的预测准确性从 20.2 mm 下降至 24.5 mm, 预测精度下降了约 21.3%。同时, JILG 站和 QITG 站的预测精度也分别下降了 13.4% 和 8.2%。相比之下, 基于 CNN-ATT 模型的预测中, NANN 站、JILG 站和 QITG 站的预测精度分别下降了 10.5%、7.5% 和 6.8%。由此可知, 构建的 CNN-ATT 模型在长预测时间步长中的精度表现更为稳定。

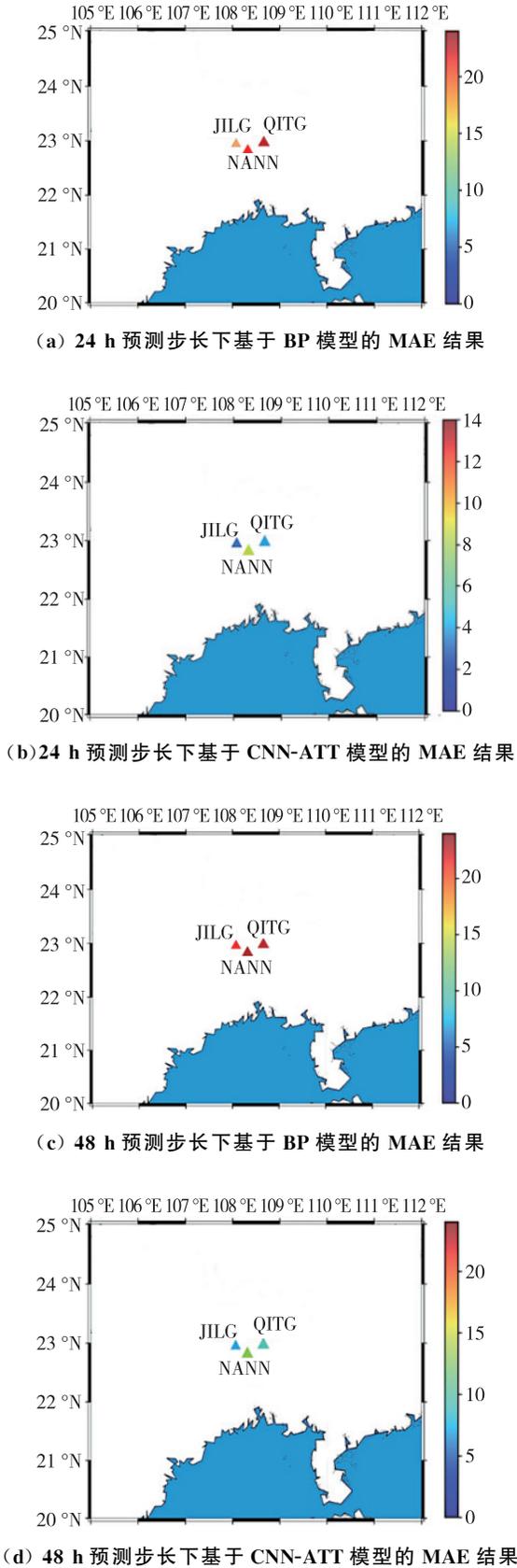


图 10 不同预测长度下的模型预测精度
 Fig. 10 Model prediction accuracy under different input length

此外,进一步探索了不同预测站点数量对模型整体预测精度的影响。对所有 NNCORS 进行了预测,并记录了预测站点数量为 2,3,4,5 时模型的预测精度。如图 11 所示,随着预测站点数量的增加,预测平均精度均呈现出一定程度的提升。值得注意的是,当预测站点数量为 5 时,即将所有 NNCORS 作为预测站点时,同一模型的预测精度优于其他站点数量设置下的精度。此外,当预测站点数量为 2 时,同一模型的预测精度较其他站点数量设置下的精度较差。这可能是因为随着多个测站加入机器学习模型,系统误差得到进一步消除,而仅考虑单一测站则无法消除系统误差,所以模型预测精度降低。此外,由图 11 可知,在同等测站站点数量条件下,CNN-ATT 模型的 MAE 和 MSE 明显低于其他基线模型,模型的预测精度更高,这也表明 CNN-ATT 模型可以更好地适应多测站预测任务。

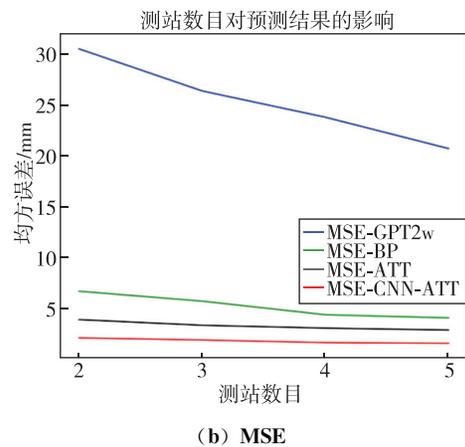
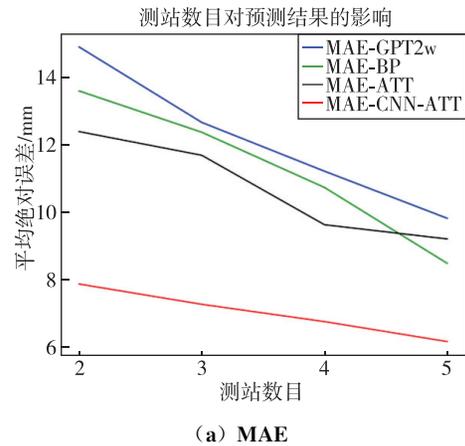


图 11 不同测站数量下的模型预测精度
 Fig. 11 Model prediction accuracy under different number of stations

2.2.3 模型定位精度

不同 ZTD 预测模型的定位精度也是衡量不同模型性能的重要指标。为此,首先需要确定不同 ZTD 预测模型的参数的初始值。考虑到电磁波的传播路径并非沿着天顶方向,因此需要将天顶方向的延迟量映射至某个倾斜的传播方向中,这一映射过程可通过映射函数(mapping function, MF)对干湿度分量进行加权求和解算。ZTD 可以分解为静力学延迟(zenith hydrostatic delay, ZHD)和湿延迟(zenith wet delay, ZWD)。其中,ZHD 占总延迟量的 90%,可以通过实测气压进行估算。ZHD 的值通过传统的 SAAS 模型计算可得,如式(12)所示

$$Z_{\text{ZHD}} = \frac{0.002\ 276\ 8\ P_s}{f(\varphi, h)} \quad (12)$$

相对而言,ZTD 初值的精确计算较为困难,通常将其作为未知数进行求解^[26]。ZTD 的初始值可以通过经典的 SAAS 模型计算,如式(13)所示

$$Z_{\text{ZTD}} = \frac{0.002\ 276\ 8}{f(\varphi, h)} \left[P_s + \left(\frac{1\ 255}{T_s} + 0.05 \right) e_s \right]$$

$$f(\varphi, h) = 1 - 0.002\ 66 \cdot \cos(2\varphi) - 0.000\ 28h$$

$$e_s = e_s \cdot 6.11 \cdot 10^{7.5(T_s - 273.15)/T_s} \quad (13)$$

式中, P_s 表示 GPS 接收机天线处的大气压力; e_s 为高度角; $f(\varphi, h)$ 为纬度和高度的函数,反映了重力加速度随着不同的纬度和大地高而变化, φ 表示纬度(rad), h 代表测站的大地高(m)。

在获得 ZTD 初始值后,ZWD 的值可以通过式(14)计算得到^[27]

$$Z_{\text{ZWD}}(e) = \frac{(Z_{\text{ZTD}}(e) - Z_{\text{ZHD}}(e) \cdot m f_w(e))}{m f_h(e)} \quad (14)$$

式中, $Z_{\text{ZTD}}(e)$ 为总延迟量; $f_h(e)$ 和 $f_w(e)$ 分别表示干湿映射函数。

在确定 $f_h(e)$ 和 $f_w(e)$ 时,常用的常映射函数包括 NMF (Neill mapping function), VMF1 (Vienna mapping function 1) 和 GMF(global mapping function) 等。根据 Boehm 等的经验法则,在高度角大于 10° 时,测站高程方向的误差将不大于 1 mm ^[28]。鉴于此,选用 VMF1 作为映射函数计算参数 (a_h, b_h, c_h) 及 (a_w, b_w, c_w)。因此, $f_h(e)$ 和 $m f_w(e)$ 可通过式(15)确定。

$$m f_h(e) = \frac{1 + \frac{a_h}{1 + \frac{b_h}{1 + c_h}}}{\text{sine} + \frac{a_h}{\text{sine} + \frac{b_h}{\text{sine} + c_h}}}$$

$$m f_w(e) = \frac{1 + \frac{a_w}{1 + \frac{b_w}{1 + c_w}}}{\text{sine} + \frac{a_w}{\text{sine} + \frac{b_w}{\text{sine} + c_w}}} \quad (15)$$

为了进一步对比不同的 ZTD 预测模型在定位精度方面的表现,将经典的 SAAS 模型、GPT2w 模型、BP 模型以及 ATT 模型与 CNN-ATT 进行对比实验。这些对流层模型的 ZTD、ZHD 以及 ZWD 初值的确定方法如表 6 所示。如前所述,ZHD 值采用式(12) SAAS 方法确定。BP 模型、ATT 模型和 CNN-ATT 模型的 ZTD 的初始值由模型算法估算可得,ZWD 则通过式(14)和式(15)确定。

表 6 不同 ZTD 预测模型参数初值确定
Tab. 6 The determination of different ZTD values for different prediction models

ZTD 预测模型/ 估算参数	ZTD	ZHD	ZWD	投影函数
SAAS	SAAS	SAAS	SAAS	
GPT2w	GPT2w	SAAS	GPT2w	
BP	BP	SAAS	ZTD(BP)- ZHD(SAAS)	VMF1
ATT	ATT	SAAS	ZTD(BP)- ZHD(SAAS)-	
CNN-ATT	CNN-ATT	SAAS	ZTD(CNN-ATT)- ZHD(SAAS)	

通过表 6 确定各基线模型的 ZTD、ZHD 和 ZWD 的初始值后,基于 GAMIT/GLOBK 进行模型精度对比实验。实验的基本设置与前述相同,选用了 5 个 CORS(NNCORS)和 14 个 IGS 站的观测数据。解算时间范围为 2009 年年积日 305~2013 年年积日 345,共计 1 501 个年积日。与表 2 相同,采用 RELAX 解算方法,卫星星历选用 IGS 精密星历,光压模型为 BERNE。然后,统计了不同 ZTD 预测模型在 5 个年积日(138,157,178,205,243)时刻 N,E,U 3 个方向定位的平均偏差和标准差。结果如表 7 所示。

由表 7 可知,与传统的 SAAS 对流层模型^[27]相比,经典的 GPT2w 模型在 N,E,U 3 个方向的精度提升分别为 2%,5%,1%;基于 BP 模型的 ZTD 预测模型定位精度在 N,E,U 3 个方向的精度提升分别为 11%,23%,-1%;基于 ATT 模型在 N,E,U 3 个方向的精度分别平均提高了 16.0%,6%,

表 7 5 种 ZTD 预测模型下的定位平均偏差和标准差

Tab. 7 The average deviation and standard deviation of positioning under five ZTD prediction models mm

方向	SAAS		GPT2w		BP		ATT		CNN-ATT	
	平均偏差	标准差	平均偏差	标准差	平均偏差	标准差	平均偏差	标准差	平均偏差	标准差
N	7.05	0.23	6.94	0.20	6.30	0.16	5.92	0.11	5.76	0.09
E	-4.92	0.16	-4.65	0.14	-3.81	0.12	-4.62	0.13	-4.33	0.11
U	-5.96	0.21	-5.93	0.13	-5.99	0.12	-4.57	0.08	-4.11	0.08

23.3%；而基于 CNN-ATT 模型在 3 个方向的精度分别提高了 18.2%，12.6%，31.0%。由此可知，神经网络模型的结合可以很好地提升 ZTD 预测模型的精度，而通过将卷积神经网络模型与注意力机制算法结合可以进一步提升模型的定位精度。

2.2.4 模型计算效率

在研究不同 ZTD 预测模型的性能时，除了考

虑模型的预测精度和定位精度外，模型的计算效率也是评估其性能的基本指标之一。为此，选择模型的时间复杂度、空间复杂度、每个历元下的训练时间、总测试时间以及 PPP 模型的收敛速度作为衡量计算效率的重要性能指标，对比结果如表 8 所示。

表 8 模型计算复杂度对比结果

Tab. 8 Comparison results of model computational complexity

类型	SAAS	GPT2w	BP ^[29]	ATT ^[30]	CNN-ATT
时间复杂度	—	—	$O(ndml^2)$	$O(n^2d)$	$O(n^2d + kn d^2)$
空间复杂度	—	—	$O(om l^2)$	$O(n^2d)$	$O(n^2d + kn d^2)$
每个历元下的训练时间/s	—	—	0.17	0.31	0.32
总测试时间/s	—	—	0.15	0.53	0.55
PPP 收敛速度 /min	53.25	53.15	52.18	48.39	47.64

表 8 中， n 表示输入样本的数量， d 表示输入数据的维度， m 表示全连接层数量， l 表示神经元的数量， o 表示输出数据的维度， k 表示 CNN 中的卷积核。根据文献[29]可知，BP 模型的时间复杂度为 $O(ndml^2)$ ，空间复杂度为 $O(nml^2)$ 。由此可知，BP 神经网络模型的计算复杂度极大程度上取决于神经元的数量和层数。相比之下，ATT 和 CNN-ATT 模型的计算复杂度主要与样本的数量和输入数据的维度相关^[30]。在选择使用较为简单的多层感知机 (multi-layer perceptron, MLP) 搭建 BP 神经网络时，BP 模型的时间复杂度和空间复杂度较 ATT 和 CNN-ATT 模型呈现出较大的优势。然而，当样本数量较大且神经网络设计较为复杂时，CNN-ATT 模型的计算优势会更加明显。

在训练时间和总测试时间方面，与 ATT 模型和 CNN-ATT 模型相比，BP 神经网络模型所需时间更短，模型的计算效率更高。然而，需要注意的是，由表 5 和表 7 中模型的预测精度和定位精度的结果可知，BP 模型的精度表现均低于 ATT 模型和

CNN-ATT 模型。此外，由表 8 结果可知，在 GPU/TPU 的支持下，ATT 模型和 CNN-ATT 模型的计算效率非常接近。虽然二者的计算效率相近，但由表 5 可知，CNN-ATT 模型的预测 MAE 和 MSE 相较于 ATT 模型分别降低了 36.2% 和 50.0%。所以，CNN-ATT 模型在综合表现上优于 ATT 模型。值得注意的是，在机器学习和深度学习领域，模型计算效率和精度之间存在一个常见的权衡。通常来说，提高模型的计算效率可能会降低一定的预测精度，而提高精度可能导致模型变得更加复杂和计算量更大。所以在实际应用过程中，不同的 ZTD 预测模型的选择需要根据具体的情况和需求来权衡模型计算效率和精度。而提出的 CNN-ATT 模型的重点是提高 ZTD 预测模型的精度，未来的研究可以考虑继续对提出的 CNN-ATT 模型进行算法优化，从而确保模型具有更优越的性能表现。

PPP 模型的收敛速度反映了模型从初始状态到获得最终解过程中所需的时间。较快的收敛速度通常被视为模型性能优越的指标，它表明模型能

够迅速适应观测数据并提供准确的解^[31]。因此,本文进一步探究了不同的 ZTD 预测模型在 PPP 收敛速度方面的性能差异。由表 8 可知,经典 SAAS 模型和 GPT2w 模型的收敛速度接近,分别为 53.25 min 和 53.15 min。与 SAAS 模型相比,BP 模型、ATT 模型和 CNN-ATT 模型的收敛速度分别减少了 2%,9.1%,10.5%。并且 ATT 和 CNN-ATT 模型的收敛速度也明显优于 BP 神经网络模型。由此可知,随着神经网络模型算法的优化与改进,模型的收敛速度也不断提升。值得注意的是,ATT 模型和 CNN-ATT 模型的收敛速度仍然相近,但结合表 5 模型的预测精度可知,相较于 ATT 模型,CNN-ATT 模型的 ZTD 预测精度更高。因此综合来看,CNN-ATT 模型的整体性能表现更佳。此外,通过对不同 ZTD 预测模型 PPP 收敛速度的对比实验可知,PPP 的收敛速度可能与 ZTD 的初始值密切相关,ZTD 初始值越精准,模型的解算效率越高。

3 结论

随着机器学习理论在 ZTD 预测领域的不断应用与发展,本研究将注意力机制理论应用到 ZTD 预测任务,建立了基于 CNN-ATT 的 ZTD 预测模型。与传统的 ZTD 预测模型相比,提出的 CNN-ATT 模型利用区域性 CORS 提供的实时精密 GPS 数据进行 ZTD 解算。具体如下:

1) 为获得更精准的 ZTD 预测结果,提出的 CNN-ATT 模型首次将日解算精度、年积日以及三维坐标作为重要的输入数据来源。为有效提取数据特征并防止过拟合现象,该模型采用卷积神经网络对输入数据进行多层特征提取和处理。最终,通过改进的注意力层和 CRLB 误差判断模块获得高精度的 ZTD 预测结果。

2) 选取南宁市的 5 个 CORS 和 14 个 IGS 站点进行区域性 ZTD 精密解算,并将 BP 模型、GPT2w 模型和 ATT 模型作为基线模型进行对比实验。结果显示,在拟合精度方面,与 GPT2w、BP 以及 ATT 模型相比,CNN-ATT 模型的系统偏差、MAE 以及 MSE 的误差水平更低,模型的拟合精度更高。在预测精度方面,相对于 BP 模型和 ATT 模型,CNN-ATT 模型的误差标准差也分别降低了 71.4% 和 33.3%。此外,在定位精度方面,CNN-ATT 模型的定位精度也明显优于其他的基线模型,并且与

ATT 模型相比,改进后的 CNN-ATT 模型的定位精度在 N, E, U 3 个方向上分别提高了 18.2%, 12.6%, 31.0%。整体而言,CNN-ATT 模型具有更为理想的预测精度表现。

3) 最后,对比了不同 ZTD 预测模型在计算效率方面的性能差异。研究结果表明,在训练时间和总测试时间方面,BP 模型表现最佳,ATT 模型和 CNN-ATT 模型表现接近。在 PPP 模型收敛速度方面,CNN-ATT 模型较其他基线模型收敛速度更快。但考虑到 CNN-ATT 模型具有较优的定位精度和预测精度,因此综合来看,CNN-ATT 模型的整体性能表现较好。

值得注意的是,区域对流层模型的精度受到多个因素的影响,包括测站分布的密集程度、测站之间的高度差异、建模区域的大小以及实验数据的实时性等。这些因素仍需要进一步地研究和探讨。此外,尽管提出的 CNN-ATT 模型能够实现更高的预测精度和定位精度,但在计算效率方面仍然存在改进的空间。未来的研究可以考虑进一步优化 CNN-ATT 模型,以实现更高的计算效率,使其能够更好地适应多场景任务。

参考文献

- [1] 戴祺,章浙涛,文援兰,等. 不同气候条件下几种对流层模型性能分析[J]. 导航定位学报, 2023, 11(4): 38-48.
DAI Qi, ZHANG Zhetao, WEN Yuanlan, et al. Performance analysis of several tropospheric models under different climate conditions[J]. Journal of Navigation and Positioning, 2023, 11(4): 38-48(in Chinese).
- [2] 徐天河,李耸,王帅民,等. 顾及气象数据的中国区域对流层延迟 RBF 神经网络优化模型[J]. 测绘学报, 2022, 51(8): 1690-1707.
XU Tianhe, LI Song, WANG Shuaimin, et al. Optimization model of regional tropospheric delay RBF neural network in China considering meteorological data[J]. Journal of Surveying and Mapping, 2022, 51(8): 1690-1707(in Chinese).
- [3] ZHANG H, YUAN Y, LI W, et al. Assessment of three tropospheric delay models (IGGtrop, EGNOS and UNB3m) based on precise point positioning in the Chinese region[J]. Sensors, 2016, 21(1): 122.
- [4] 黄良珂,刘立龙,文鸿雁,等. 亚洲地区 EGNOS 天顶对流层延迟模型单站修正与精度分析[J]. 测绘学

- 报, 2014, 43(8): 808-817.
- HUANG Liangke, LIU Lilong, WEN Hongyan, et al. Single station correction and accuracy analysis of EGNOS zenith tropospheric delay model in the Asian region[J]. *Journal of Surveying and Mapping*, 2014, 43(8): 808-817(in Chinese).
- [5] LAGLER K, SCHINDELEGGER M, BÖHM J, et al. GPT2: empirical slant delay model for radio space geodetic techniques[J]. *Geophysical Research Letters*, 2013, 40(6): 1069-1073.
- [6] CHEN J, JIANG Y, FAN Y, et al. Comprehensive analysis of the global zenith tropospheric delay real-time correction model based GPT3[J]. *Atmosphere*, 2023, 14(6): 946.
- [7] LIAN D, HE Q, LI L, et al. A novel method for monitoring tropical cyclones' movement using GNSS zenith tropospheric delay[J]. *Remote Sensing*, 2023, 15(13): 3247.
- [8] AKAR A U, INAL C. Prediction of zenith tropospheric delay in GNSS observations using support vector regression[J]. *Advances in Space Research*, 2023, 71(11): 4659-4680.
- [9] YANG X, LI Y, YU X, et al. Regional/single station zenith tropospheric delay combination prediction model based on radial basis function neural network and improved long short-term memory [J]. *Atmosphere*, 2023, 14(2): 303.
- [10] DING S, SU C, YU J. An optimizing BP neural network algorithm based on genetic algorithm[J]. *Artificial Intelligence Review*, 2011, 36(2): 153-162.
- [11] 王勇, 张立辉, 杨晶. 基于 BP 神经网络的对流层延迟预测研究[J]. *大地测量与地球动力学*, 2011, 31(3): 134-137.
- WANG Yong, ZHANG Lihui, YANG Jing. Research on tropospheric delay prediction based on BP neural network[J]. *Geodesy and Geodynamics*, 2011, 31(3): 134-137(in Chinese).
- [12] 肖恭伟, 欧吉坤, 刘国林, 等. 基于改进的 BP 神经网络构建区域精密对流层延迟模型[J]. *地球物理学报*, 2018, 61(8): 3139-3148.
- XIAO Gongwei, OU Jikun, LIU Guolin, et al. Constructing a regional precise tropospheric delay model based on an improved BP neural network[J]. *Journal of Geophysics*, 2018, 61(8): 3139-3148(in Chinese).
- [13] 杨旭, 何祥祥, 王媛媛, 等. 一种基于机器学习算法的区域/单站 ZTD 组合预测模型[J]. *全球定位系统*, 2022, 47(1): 98-102+126.
- YANG Xu, HE Xiangxiang, WANG Yuanyuan, et al. A regional/single station ZTD combination prediction model based on machine learning algorithms[J]. *Global Positioning System*, 2022, 47(1): 98-102+126(in Chinese).
- [14] BAHDANAU D, CHO K, BENGIO Y. Neural machine translation by jointly learning to align and translate[J]. arXiv:1409.0473, 2014.
- [15] MANGHI R L, BERNACCHIA D, CASAJUS L G, et al. Tropospheric delay calibration system performance during the first two BepiColombo solar conjunctions[J]. *Radio Science*, 2023, 58(2): 1-15.
- [16] SSENUNZI C R, ANDIMA G, AMABAYO B E, et al. Performance of ray-traced VMF3 products in retrieving zenith tropospheric delay over the African tropical region[J]. *Journal of Atmospheric and Solar-Terrestrial Physics*, 2023, 243: 106014.
- [17] ZHOU H, ZHANG S, PENG J, et al. Informer: beyond efficient transformer for long sequence time-series forecasting[C]// *Proceedings of AAAI Conference on Artificial Intelligence*, 2021: 11105-11115.
- [18] YU F, KOLTUN V, FUNKHOUSER T. Dilated residual networks[C]// *Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu: IEEE, 2017: 636-644.
- [19] 孟怡悦, 郭迟, 刘经南. 基于注意力机制和奖励塑造的深度强化学习视觉目标导航方法[J/OL]. *武汉大学学报(信息科学版)*: 1-9 [2023-09-04]. <https://doi.org/10.13203/j.whugis20230193>.
- MENG Yiyue, GUO Chi, LIU Jingnan. A visual target navigation method for deep reinforcement learning based on attention mechanism and reward shaping [J/OL]. *Journal of Wuhan University (Information Science Edition)*: 1-9 [2023-09-04]. <https://doi.org/10.13203/j.whugis20230193>(in Chinese).
- [20] 黄悦华, 张子豪, 陈庆, 等. 一种基于多头注意力机制的 CNN-BiLSTM 高海拔多因素输电线路可听噪声预测方法: 202310550858[P]. 2023-11-05.
- HUANG Yuehua, ZHANG Zihao, CHEN Qing, et al. A method for predicting audible noise of high altitude multifactor transmission lines using CNN-BiLSTM based on multi head attention mechanism: 202310550858[P]. 2023-11-05(in Chinese).
- [21] 袁红春, 张文凤. 融合 SimAM 注意力机制和双向 ConvLSTM 的异常检测方法[J]. *小型微型计算机系统*, 2023, 44(8): 1777-1784.
- YUAN Hongchun, ZHANG Wenfeng. Anomaly detection method combining SimAM attention mechanism and bidirectional ConvLSTM[J]. *Small Micro Computer*

- System, 2023, 44(8): 1777-1784(in Chinese).
- [22] YANG F, GUO J, ZHANG C, et al. A regional zenith tropospheric delay (ZTD) model based on GPT3 and ANN[J]. Remote Sensing, 2021, 13(5): 838.
- [23] 魏民, 余学祥, 杨旭, 等. 基于随机森林和反向传播神经网络机器学习方法的区域 ZTD 建模精度分析[J]. 大地测量与地球动力学, 2023, 43(7): 755-760.
- WEI Min, YU Xuexiang, YANG Xu, et al. Accuracy analysis of regional ZTD modeling based on random forest and back propagation neural network machine learning method[J]. Journal of Geodesy and Geodynamics, 2023, 43(7): 755-760(in Chinese)
- [24] 鄂栋臣, 詹必伟, 姜卫平, 等. 应用 GAMIT/GLOBK 软件进行高精度 GPS 数据处理[J]. 极地研究, 2005, 17(3): 173-182.
- E Dongchen, ZHAN Biwei, JIANG Weiping, et al. High-precision GPS data processing by GAMIT/GLOBK[J]. Polar Research, 2005, 17(3): 173-182 (in Chinese).
- [25] 张永林, 蔡昌盛. 基于再分析资料 ERA5 的对流层延迟估计方法及精度评估[J]. 大地测量与地球动力学, 2020, 40(1): 62-65.
- ZHANG Yonglin, CAI Changsheng. A tropospheric delay estimation method and accuracy evaluation based on reanalysis data ERA5[J]. Geodesy and Geodynamics, 2020, 40(1): 62-65(in Chinese).
- [26] 何力维, 陈义, 周亚军. 水汽辐射计与 PPP 估计天顶湿延迟对比研究分析[J]. 全球定位系统, 2015, 40(6): 68-71.
- HE Liwei, CHEN Yi, ZHOU Yajun. Comparison between zenith wet delay by water vapor radiometer and PPP[J]. GNSS World of China, 2015, 40(6): 68-71(in Chinese).
- [27] 陈猛, 陈俊平, 胡丛玮. 对流层模型评估及其在 GNSS 精密单点定位中的应用[J]. 大地测量与地球动力学, 2016, 36(3): 229-234.
- CHEN Meng, CHEN Junping, HU Congwei. Performance evaluation of troposphere models and application in GNSS precise point positioning[J]. Geodesy and Geodynamics, 2016, 36(3): 229-234(in Chinese).
- [28] BÖHM J, WERL B, SCHUH H. Troposphere mapping functions for GPS and VLBI from ECMWF operational analysis data[J]. Journal of Geophysical Research, 2006, 111(B02): B02406.
- [29] LARSSON F, KARLSSON A. Evaluation of machine learning algorithms to reduce paging signalling in a telecom network[D]. Linköping: Linköping University, 2017.
- [30] KELES F D, WIJewardena P M, HEGDE C. On the computational complexity of self-attention[C]// Proceedings of 34th International Conference on Algorithmic Learning Theory. Singapore: PMLR, 2023, 201: 597-619.
- [31] 孙大伟, 艾孝军, 贾小林, 等. BDS/GNSS 精密单点定位性能分析 [J]. 大地测量与地球动力学, 2022, 42(11): 1111-1116+1127.
- SUN Dawei, AI Xiaojun, JIA Xiaolin, et al. Performance analysis of BDS/GNSS precise point positioning[J]. Geodesy and Geodynamics, 2022, 42(11): 1111-1116+1127(in Chinese).

(编辑:黄利华)