

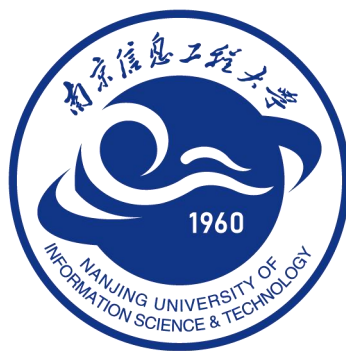
分类号: _____

单位代码: _____

密 级: _____

学 号: _____

硕士专业学位论文



论文题目: 基于深度学习的葵花 8 卫星亮温资料
降水反演研究

申请人姓名: 王瑞

指导教师: 刘玉宝教授

类别名称: 资源与环境

领域名称: 气象工程

培养学院: 大气物理学院

提交时间: 2023年5月27日

二〇二三年五月

目 录

第一章 绪论.....	1
1.1 研究背景.....	1
1.2 国内外研究现状.....	3
1.2.1 单通道法反演降水算法现状.....	3
1.2.2 多通道阈值法反演降水算法研究现状	4
1.2.3 机器学习算法反演降水现状.....	5
1.3 研究内容.....	7
1.4 章节安排.....	8
第二章 资料与方法	10
2.1 资料介绍.....	10
2.1.1 静止卫星数据.....	10
2.1.2 降水基准数据.....	12
2.1.3 葵花 8 数据与 GPM IMERG 降水数据的比较.....	12
2.1.4 数据集构建.....	15
2.2 深度学习模型简介.....	16
2.2.1 U-Net 模型介绍.....	16
2.2.2 ConvMixer 模型介绍	17
2.2.3 pix2pixGAN 模型介绍	18
2.3 反演结果的评估方法.....	19
第三章 不同卫星观测通道对降水反演的影响	21
3.1 不同卫星观测单通道的降水反演	21
3.2 不同卫星观测多通道的降水反演	27
3.3 本章小结.....	30
第四章 不同卫星观测时刻对降水反演的影响	31
4.1 不同卫星观测单时刻的降水反演	31
4.2 不同卫星观测多时刻的降水反演	35
4.3 本章小结.....	37
第五章 不同反演模型对降水反演的影响及其改进	38
5.1 不同反演模型对降水反演的影响比较	38
5.2 基于两步预报法的模型改进	42
5.3 本章小结.....	43
第六章 结论与展望	44
6.1 结论.....	44
6.2 创新点.....	45

6.3 不足与展望.....	45
参考文献.....	46

摘要

大气中不同尺度的降水过程在水和能量循环中扮演着关键的角色，同时也对人类社会的生产活动有重要影响。降水及其相关的极端天气和次生地质灾害会带来巨大的生命财产和经济损失。因此对于降水全面、及时和准确的监测，包括降水强度和范围精确观测和判别，不但能够为降水发展过程机理研究提供很好的观测资料，同时对减灾防灾、保障人民日常生活和社会稳定发展具有重要意义。

目前，基于静止卫星红外通道遥感观测的降水反演算法在业务气象中有广泛的应用，同时，深度学习技术的发展给基于卫星红外通道遥感观测的降水反演带来了一个机遇，产生了基于机器学习或深度学习的定量降水强度反演技术。目前，虽然该技术已取得了不错的效果，但仍存在诸多问题，如反演的降水强度低估、降水区误报、降水区轮廓模糊化等。本研究利用华南地区夏季的 Himawari-8 静止卫星成像仪观测资料和 GPM 多卫星降水估计值资料（IMERG）构建一个深度学习降水反演模型，基于 Himawari-8 的辐射亮温资料输入，反演降水。GPM IMERG 是 30 分钟的平均降水，为目前全球降水估算最佳产品，但其需要等待 3.5 个月才可获取，而 Himawari-8 的资料可近实时获取。本研究基于先进深度学习算法，构建了葵花 8 静止卫星红外通道观测的降水反演模型，将 GPM IMERG 作为降水基准数据，使用 Himawari-8 数据进行格点对格点反演，以产生高分辨率、及时、快速和有效的卫星降水数据。基于 Himawari-8 反演降水，可以实现逐 10 分钟的 30 分钟的降水强度，一定程度上填补了地面雨量计稀少地区的空白。此外，本研究开展了针对输入卫星不同观测通道、时间和时段及不同深度学习模型的敏感性研究，并建立了一个新的深度模型结构，有效提升了降水强度和范围的反演精度。

本文构建了 2016 至 2019 年 Himawari-8 静止卫星上 9 个红外通道亮温数据和 GPM IMERG 半小时平均降水量的机器学习训练集，然后将其划分为训练集（共 5352 个样本）、验证集（1784）和测试集（1784）。首先，选用目前深度学习技术中使用最为广泛、稳定可靠的 U-Net 模型，分析了 Himawari-8 卫星的不同红外通道对降水估计的贡献，比较了 9 个单通道和多通道的反演，结果表明采用通道 13 能获得相对更好的反演效果。随后，本文比较了多个单时刻以及多时刻组合的降水反演精度，发现采用多时刻观测输入，并未能进一步改进反演精度。最后，为进一步提升降水反演效果，本文比较了三个深度模型的降水反演能力，包括 U-Net、pix2pixGAN 和 ConvMixer 等。结果表明，pix2pixGAN 反演的降水分布最优，比其他模型结果更具有鲜明详实的结构，但仍然存

在少量降水漏报、低估强降水的问题。为进一步优化 pix2pixGAN 的反演效果，本文基于两步预报法的 pix2pixGAN 模型构建了 B-pix2pixGAN 模型，改进了降水反演的精度高。

关键词：降水反演，U-Net，CGAN，深度学习，IMERG，Himawari-8 卫星

Abstract

Precipitation at various scales in the atmosphere is crucial for the water and energy cycle, and it also has an important impact on human activities, particularly the productivity of societies. Extreme weather and secondary geological hazards triggered by precipitation can lead to significant losses in terms of life, property, and economy. Therefore, it is essential to monitor precipitation activity comprehensively, accurately, and in a timely manner. Precipitation observation and identification, including the intensity and extent, can not only provide critical information for studying precipitation development mechanisms but also aid in disaster mitigation and prevention, ensuring stable social development.

Currently, precipitation retrieval algorithms based on the infrared channel of geostationary satellites has a wide range of applications in operational meteorology. With the continuous development of artificial intelligence technology, new precipitation intensity retrieval methods and quantitative precipitation estimation models are proposed based on machine learning (ML) or deep learning (DL). However, many studies have shown that although ML/DL technologies have achieved good results, several problems still exist, such as underestimation of precipitation intensity in the retrieval, incorrect prediction of precipitation zones and blurring of the output contours. In this study, Himawari-8 geostationary imager observations and GPM multi-satellite precipitation estimate (IMERG) information for the summer season in southern China are used. The target dataset of GPM IMERG needs to wait for 3.5 months to be obtained, while the data of Himawari-8 can be obtained in real-time. GPM data is essentially a 30 minute precipitation intensity, which is the average precipitation within 30 minutes. Himawari-8 inversion of precipitation advantages can achieve precipitation intensity of 30 minutes per 10 minutes, filling the gap in surface rainfall gauge data to some extent. The Himawari-8 satellite data has higher horizontal resolution and can provide a more refined description of precipitation activities. Although GPM IMERG precipitation products are close to real precipitation, the timeliness and resolution of the data are difficult to meet the needs of existing precipitation observation operations and research. Therefore, GPM IMERG is used as the baseline data of precipitation in this study, and Himawari-8 data is used for grid to grid point reflection to generate high-resolution, timely,

fast and effective satellite precipitation data. Based on the advanced deep learning algorithm, a precipitation inversion model for the infrared channel observations of the Sunflower 8 geostationary satellite was constructed. Sensitivity studies for different observation channels, times and periods of input satellites and deep learning models were carried out. And a new depth model structure is established to effectively improve the inversion accuracy of precipitation intensity and extent.

In this paper, a machine learning training set of 9 infrared channels of bright temperature data and GPM IMERG half-hourly average precipitation on the Himawari-8 geostationary satellite from 2016 to 2019 was constructed and then divided into a training set (5352 samples in total), a validation set (1784) and a test set (1784). The U-Net model, the most widely used and stable deep learning technique, is chosen to analyze the contribution of different IR channels of the Himawari-8 satellite to the precipitation estimation. The retrieval results of single and multiple channels are compared, and it is found that the Channel 13 yields better results. Subsequently, the accuracy of model outputs with the input of single-moment observation and the input of multiple-moment combinations is compared, and it is found that the latter did not improve the accuracy of the inversions. Finally, to further improve the precipitation prediction, the accuracy of three different DL models is compared, including U-Net, pix2pixGAN and ConvMixer. The results show that pix2pixGAN is the best model, but its output still suffers from omission and underestimation. In order to further optimize the inversion effect of pix2pixGAN, in this paper, we construct a B-pix2pixGAN model based on the pix2pixGAN model of the two-step forecast method to improve the accuracy of precipitation inversion.

Keywords: Precipitation retrieval, U-Net, CGAN, Deep learning, GPM IMERG, Himawari-8 satellite

第一章 绪论

1.1 研究背景

地球大气中的各类降水过程在地球系统中水、能量和物质等循环中扮演着重要的角色，不同尺度和强度的降水活动也会对人类活动以及社会经济发展也有着一定程度的影响。我国华南地区属于东亚季风区，具有显著降水季节性差异性，且在近年来气候变化影响下极端降水灾害频发，对当地气象综合防灾减灾提出了严峻的挑战。因此进一步发展降水观测和预报技术在农林业、交通运输和海洋监测等社会公共服务领域始终是不可或缺的需求^{[1][2]}。对于降雨活动频繁的华南地区，尤其是人口密集的城市地区，及时准确的对降水活动的发生、发展和消停进行即时监测有着非常重要的社会经济价值。对于降水观测和预报技术和相关研究，一方面关乎社会秩序和人民生命财产的安全，另一方面也对研究地区乃至全球气候变化具有重要意义。

目前，已有许多研究基于各类降水观测数据对于我国季风区部分降水活动的季节和年际变化等特征进行长期的统计和分析^[3-5]。虽然这些研究得到了我国降水活动气候的一般特征，但是也显示出了当前长期降水观测研究的局限性。传统的降水观测以地基观测为主，其针对降水过程的水平范围十分有限，存在一定的观测系统盲区，难以反映出水平尺度较大的降水活动，其获取的数据资料也存在着时空分辨率较低等问题。其中地面雨量计和地基天气雷达是目前降水主要地基观测手段，但是两者收集的降水观测数据的空间分辨率在许多地区非常低，不能有效地描述这些地区的降水演变过程^[6]。此外，地基天气雷达雨区信号距离衰减、反射因子与降水关系不稳定和地形影响等导致降水估计存在显著不确定性^[7]。因此综合来看，卫星观测技术对弥补目前各类尺度降水过程快速和即时的监测、跟踪、预报，及长期的降水气候统计分析研究，都是非常重要的意义。

近年来，随着空间遥感技术和卫星通信技术的迅速发展，基于气象卫星的降水遥感观测和相应的反演技术也随之兴起。卫星观测能够提供直观的遥感图像信息，在一定程度上弥补传统地面观测手段的不足^[8]。相较于地基观测，卫星以“自上而下”的观测方式观测，观测目标区域在水平范围和分辨率等方面更具优势。尤其在缺乏地基观测的无人区、海域等区域，卫星降雨观测是对现有地基观测提供很好的补充。目前，卫星云图在实际业务应用中也已经成为了获取高质量、高分辨率卫星降水资料的重要手段，并且在监测暴雨和台风等气象灾害中发挥着十分重要的作用^{[9][10]}。研究表明获取高精度高时

间分辨率的卫星数据并进行降水过程相关的数据反演,对更进一步提高降水预报的准确度有着重大意义^{[12][20]}。

卫星降水反演产品是基于卫星微波、可见光/红外遥感技术和多传感器联合遥感技术等获取的资料的数据处理,获得降水估计。其中,微波遥感是用卫星搭载的微波设备来探测、接收被测物体在微波波段的电磁辐射和散射特性,以识别远距离物体的技术。该观测技术与可见光、红外遥感技术相比,不易受气象条件和日照水平的影响具有全天候昼夜工作能力。它们可以穿透云层探测降水云团内部的垂直结构,能够以较高精度反演降水。此类在轨的卫星目前包括欧洲 METOP 气象卫星、我国的风云三号卫星 (FY-3) 以及美国 NOAA 系列卫星等。不过,微波遥感探测技术多搭载于极轨卫星,其时间采样精度不高^[11],因此该类数据在区域性精细化的降雨反演技术中的应用仍然存在相当大的局限性。

在基于卫星观测的降水反演算法中,静止卫星数据可以反演得到高分辨率的卫星降水产品,其中可见光和红外通道观测可以兼顾降雨的区域范围和精细化监测的要求,并且可以提供大气中水汽的垂直信息,还能够获得丰富的大气、云和降水各类宏微观信息,如云厚、云高、云顶温度、降水的强度和分布、以及各类微物理变量等。这些要素的准确和即时的观测和监测可以很好的描述不同尺度降水结构特征及其演变过程。因此,利用卫星可见光和红外通道数据进行降水反演可以作为非地基降水活动观测的一个重要手段。

新一代地球静止气象卫星都搭载了可见光、红外和水汽传感器,此类卫星有风云四号 (FY-4)、MSG (Meteosat Second Generation)、葵花 8 号卫星 (Himawari-8) 以及 GOES (Geostationary Operational Environmental Satellite) 等。这些静止卫星上的可见光、红外遥感技术能够提供的各类云层和降水活动相关的红外遥感信息,并且具有较高的时间和空间分辨率,对于短时间内迅速变化的天气现象可以很好的捕捉。因此,基于该数据反演的卫星降水资料具有较好的连续性,其时空范围和分辨率上也存在独特的优势,因此目前基于地球同步卫星的多通道反演算法是反演大范围的高分辨率降水数据的热门方向之一^{[12][13]}。

上述多传感器联合反演类降水产品克服了以往地面观测的许多缺点,有效扩展了数据集的时空覆盖度,已经成为在全球范围内被广泛应用的高质量降水数据资料^[11],其中常用的数据产品有 TRMM (Tropical Rainfall Measuring Mission satellite) 卫星降水产品、

PERSIANN (Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks) 以及 CMORPH (Climate Prediction Center Morphing Method)。近年来许多学者展开了针对此类卫星云降水观测资料产品的时空特征、误差以及适用性等方面进行了分析。薛鹏飞等^[14]针对雅鲁藏布江流域的降水过程,对 TRMM 等卫星降水产品进行误差分析与精度评估,其结果表明 TRMM 在整体上对降水发生反演产品具有较好效果。张天宇等^[14]针对三峡水库区域,研究了多源降水产品的时空上的适用性,发现在实际时间尺度上 TRMM 和 CMORPH 产品都具有极大潜力。杨娜等^[15]利用 PERSIANN-CDR 产品识别淮河流域干旱事件以及反演该区域降水时空分布特征,并验证了 PERSIANN-CDR 的降水反演效果和精度。

目前,全球降水观测中最先进的天基平台是全球降水测量卫星 GPM (Global Precipitation Measurement),该卫星搭载了更先进的 GMI 被动微波辐射计和双频雷达传感器 DPR,加强了对微量降水和固态降水的识别能力,为多卫星降水估计提供了精确的校准参考。GPM IMERG 的三级产品 IMERG (Integrated Multi-satellite Retrievals for GPM) 能提供水平分辨率 $0.1^{\circ}\times 0.1^{\circ}$,时间分辨率 30 分钟的降水数据,其滞后处理的 Final-run 版本经过地面站资料的订正后和地面观测的真实降雨量已经十分接近^[16-18]。近年来,该产品已被广泛应用在气候变化、干旱识别和洪水预报等多个方面^[19-21]。此外,GPM 产品在时空上与高分辨率的红外格点数据有着较好的对应关系,具备相互耦合的可对比和可数据分析的特性,这为红外反演降水算法的发展提供了较好的基础,并且可以为相应的机器学习技术提供可靠的模型目标变量。因此,利用当代深度学习技术,基于实时静止卫星数据,反演类似 IMERG 的降水反演结果,对提升降水预报精度有重大意义^{[12][20]}。

1.2 国内外研究现状

1.2.1 单通道法反演降水算法现状

使用静止卫星的红外辐射观测数据反演降水的算法,主要依赖于卫星观测的云属性信息与降水之间的关系。早期气象卫星观测内容主要为单光谱的低分辨率云图。由于早期的技术限制,首先出现的是单通道方法。在大西洋热带实验 GATE (GARP Atlantic Tropical Experiment) 中,Arkin^[24]发现在红外云顶亮温低于 225K 到 255K 的足够大的区域中,冷云覆盖面积与 6h 降水累计量存在高度的相关性^[23],表明了单通道方法对于降水反演的可用性。基于此,Arkin 和 Meisner 等^[24]提出使用 GOES 降水指数 (GOES

Precipitation Index, GPI) 的降水估测算法来对降水的强度和分布进行区域性反演。不过计算 GPI 指数需要足够大的时空尺度。他们也发现在大尺度时空条件下, 基于单通道卫星数据估算中小尺度降水的准确率会因此受到限制^[25]。针对该问题, Todd 等^[26]结合被动遥感中的瞬时值与高频次的红外图像, 将 Margin Infused Relaxed Algorithm, 即 MIRA 算法引入, 使用概率匹配的方法优化红外和降水的关系, 最终实现了对降水估计所需的时空尺度的缩小和优化。此后 Adler 与 Negri 提出了基于红外通道的“由内到外”的对流层状技术 CST (Convective-Stratiform Technology) 算法^[27]。该方法将局部亮温最小值视作对流中心, 并将降水率和降水面积分配给周围区域各点, 同时利用雷暴云砧的亮温阈值法来判断层状降水信息。Petersen 等^[28]观测到在某些天气情况下闪电密度与对流降水之间存在很强的正相关性, 因此 Grecu 等^[29]将闪电信息作为补充变量, 结合到 CST 方案中来优化性能。Reudenbach 等^[30]则加入了水汽通道信息来区分对流云及卷云, 从而进一步提升降水识别效果。Ebert 和 Fruh 等^{[31][32]}发现此算法在反演中纬度热带气旋降水效果时表现并不理想, 仅局限于适用热带深对流系统。

1.2.2 多通道阈值法反演降水算法研究现状

随着地球同步卫星光谱数据分辨率的提高, 针对降水反演的多通道阈值法算法也在不断发展。Ba 与 Nauss 等^{[33][34]}从卫星的多光谱数据中提取了丰富的云辐射特性和云相关物理参数, 例如云顶高度、云的光学厚度、云相等, 从而使得相应的降水反演算法得以进一步提升。Ebert 等^[35]发现增强对流和层状降水分类的效果有助于提高卫星降水估计的准确度。因此, 一些降水反演算法将重点放在了对流-层状区域的分类上。Torricella 和 Cattani 等^{[36][37]}研究发现云顶高度在对流降水强度上有一定的意义, 且冰云中云的有效粒子半径与降水强度成正相关。Thies 等^[38]使用 MSG-SEVIR (Meteosat Second-Generation-Spinning Enhanced Visible and InfraRed Imager) 数据通过计算降水概率矩阵识别了对流与层状降水过程。该方法的主要特点之一是使用亮温及亮温差代表云参数信息, 比如将 $8.7\mu\text{m}$ 与 $10.8\mu\text{m}$, $12.1\mu\text{m}$ 与 $10.8\mu\text{m}$ 的亮温差作为云相信息, 将 $6.2\mu\text{m}$ 与 $10.8\mu\text{m}$, $7.3\mu\text{m}$ 与 $12.1\mu\text{m}$ 的亮温差及 $10.8\mu\text{m}$ 通道的红外亮温数据视为云顶高度信息, 从而提高对于云水资源的反演精度。Feidas 等^[39]在对对流和层状降水分类时, 也采用了类似的多光谱红外数据结合的阈值法, 并取得了一定的效果。Lazri 等^[40]提出适用于地中海地区方案 CS-RADT, 利用多通道信息区分对流云区和层状云区, 由瞬时雷达

数据校准来确定阈值，并且结合雨量站资料计算得出了各种云区的降水率。

然而，多通道阈值算法虽然比单通道方法改进了许多，但是仍然存在诸多问题。如多通道阈值算法实际上严重依赖于云的物理属性与降水过程之间的参数关系，其使用的物理变量和数据集的大小也有限，并且遥感信息与降水之间实际是非线性关系，这些都是制约该方法的进一步深入发展的重要问题。此外，随着气象数据量的激增，更多的输入特征被输入模型，传统的概念模型假设逐渐变得难以适用。因此，目前以多通道阈值算法提升反演准确度的途径已经遇到了一定的瓶颈，是目前卫星云降水资料反演的产品释用中急需解决的关键问题。

1.2.3 机器学习算法反演降水现状

近几年来，机器学习算法逐渐成为连结多通道亮温数据与降水信息的有力工具。已有多个研究将不同来源的气象产品包括数值预报产品、卫星遥感资料、地面观测站资料以及卫星降水产品资料相结合，利用机器学习的方法估测降水。传统的机器学习方法比如支持向量机、人工神经网络以及梯度提升树等都被成功地用来预估降水，为模拟高度非线性关系提供了新的方向。

多位学者进行了此方面研究，一些算法和模型的应用已颇具成果。Bellerby 等^[41]将高分辨率的 GOES 多光谱图像信息作为模型输入变量，输入至人工神经网络中进行训练，发现反演 TRMM 产品的效果优于 GPI 降水指数。师春香等^[42]从 GMS 卫星红外图像中挑选出 12 个云图特征量用于反向传播神经网络模型训练，得到了相似的结论。而熊秋芬等^[43]利用 4 个 GMS4 卫星通道资料及经纬度信息估算地表降水量，得到的 BP 神经网络效果优于单通道阈值法。Hong 等^[44]提出使用 PERSIANN 云分类系统 (CCS) 估算降水，先将云图分为不同的云块，基于自组织映射根据几何纹理特征对云块进行聚类分析，并建立不同类别的亮温-降水关系用于估算降水量。Hamidi 等^[45]使用支持向量机模型，并发现模拟得到的月累计降水量效果优于人工神经网络。Ma 等^[46]使用葵花 8 号静止卫星的多通道亮温及亮温差数据，在考虑了到可见光的影响后，根据太阳高度角分别在三个时段建立梯度提升树算法子模型反演水量站降水，提升识别降水区域准确度。这些算法突破了传统阈值法的单一性，有效避免了同步多光谱卫星反演降水的精度限制，将降水反演算法带领到新的思路和方向上去，对卫星降水资料反演准确度的提升起到了重大推动作用。

随机森林算法 (Random Forests, RF) 也被广泛应用于降水反演模型中, 并取得较好的结果。由于随机森林通过输入多个与降水相关的特征变量, 来建立多个决策树以达到分类和回归预测的目的, 因此该方法与其他算法相比不易过拟合, 从而在应用时通常优于单个决策树算法。Kühnlein 等^[47]利用 MSG SEVIRI 多光谱数据使用随机森林方法在不同日变化阶段成功地分类了对流和层状降水区域, 进一步的提升了降水估计的精度。DAS 等^[48]拓宽了随机森林算法的使用数据源。除红外亮温数据, 他们还使用大气温度及湿度信息作为随机森林模型输入, 成功对对流降雨进行反演, 因此该算法在预报应用中可以应对不同来源的数据的情况。

Min 等^[12]使用葵花 8 数据和 GFS 模式数据通过随机森林方法分离了降水与非降水区域, 发现在此基础上使用随机森林回归方法模拟的降水强度相比 GPM IMERG 降水数据有一定低估。他们发现对于集中在 0.5mm/h 以下的小雨区域的反演效果很好, 对于降水强度大于 3mm/h 的区域反演效果较差, 不过整体上仍取得了较好的结果。Turini 等^[49]基于 MSG1 静止卫星数据以及云覆盖产品使用随机森林算法反演云区、降水区域及降水强度, 对高于 4mm/h 以上降水强度的反演精度较差, 表明使用机器学习模型反演中雨和大雨区降水强度的能力还有待提升。

上述研究结果发现: 在区分降水区域时, 命中率和虚警率平均分别达到 0.6 和 0.5 左右; 而在模型反演降水强度时, 由于降水样本分布不平衡的问题, 机器学习算法对于降水高值区存在严重低估的现象, 对识别中雨区的精度较低, 并且几乎反演不出大雨区域。虽然, 机器学习算法在降水反演技术中发挥了重要的作用, 但是目前的机器学习模型仍具有一定的局限性。主要有以下几个方面问题: 1) 在估测降水强度时, 样本的不平衡往往导致对较高强度降水的低估; 2) 将反演算法应用于同一时段的训练集和验证集已经能获得不错的结果, 但模型还未能应用于不同时段的测试集, 还未能充分证明模型的普适性。

深度学习是从人工神经网络延伸出来的算法, 通常具有多层结构, 经过逐层转换从初始特征中获取层次化的深层特征, 但由于输入图像的维度较高, 往往对施展深度学习的计算平台有着较高的硬件需求。近年来, 我国自动观测系统的逐步建设与推广, 气象逐渐信息化和集约化, 气象数据量空前激增。由于擅长从海量的数据集中处理和提取出丰富的特征信息, 深度学习算法已然作为一种新兴方法被引入气象领域。深度学习 (Deep Learning, DL) 由 Hinton 提出, Hinton 和 Ruslan 在 2006 年解决了如何消除深层网络训

练中梯度消失的问题，至此深度学习的浪潮兴起^[50]。2012年，Hinton证明了深度学习与传统的机器学习相比具有功能强大且灵活的优越性^[51]，也为提升降水反演效果指明了一条新方向。

国内外许多学者使用了深度卷积神经网络（Convolutional Neural Networks, CNN）等以及它们衍生的变体来构建降水反演模型。Wang等^[52]基于CNN建立红外反演算法，使用了与PERSIANN-CCS相同的输入变量但模型性能优于云分类系统，体现了CNN多层网络结构单元对于降水高维特征的捕捉能力。在对降水的时空序列预测研究中，Shi等^[53]提出了卷积长短期记忆神经网络ConvLSTM模型，同时具备提取空间和时间特征的能力；之后又出现了改进的变体TrajGRU、PredRNN以及MIM等提高预测准确率；生成对抗网络（Generative Adversarial Nets, GAN）也被提出进一步改善模型精度和图像模糊问题^[54]。Sadeghi等^[55]在CNN模型中加入了U-Net结构，建立近实时红外反演模型，有利于恢复自编码器在特征传递过程中丢失的信息。Trebing等^[56]在U-Net结构中加入了注意力机制和深度可分离卷积模块，在保持模型性能的前提下大大降低了模型参数，提升了训练效率。

目前深度学习模型大多被应用于计算机视觉和自然语言处理领域，在气象领域应用相对较少，其中在降水反演领域更加缺乏应用。因此，将不同类型的深度学习模型应用于降水反演，分析反演效果，探究各模型的适配性，也是本文研究和加强卫星云降水资料反演技术的重要研究内容。

1.3 研究内容

前述分析表明，虽然现有基于机器学习或深度学习的定量降水估计模型的降水强度卫星反演已经有所发展，但是仍存在较严重的降水强度低估、降水区误报、降水区轮廓模糊化的问题。利用可见光、红外不同通道数据对于降水反演的效果贡献尚不明晰。尚缺乏对不同的深度学习模型，特别是近期发展的生成对抗神经网络（GAN）的效果的评估比较。因此，本文通过对2016至2019年共4年葵花8静止卫星9个红外通道亮温数据和GPM IMERG降水量的机器学习，针对降雨过程，比较和分析卫星观测通道和时刻对降水反演的影响，不同深度学习模型的反演能力，并在此基础上对其现有的模型进行调试和改进。

本文选择使用日本的葵花8号静止卫星高分辨率多通道数据，作为训练模型的输入

特征,把 GPM IMERG 降水资料数据作为模型的目标变量。基于本文发展的降水深度学习模型,对包含广东省、海南省的矩形区域的降水量进行反演估计,以常用的降水预报评价指标评估深度学习模型的降水反演效果。

本文的主要研究内容为:1)分析和对比葵花 8 不同通道的数据作为模型输入对降水反演精度的影响;2)分析和对比葵花 8 中不同时刻的数据作为模型输入对降水反演精度的影响;3)根据已经得到的最优通道和时刻数据作为模型输入,比较不同深度学习模型之间的降水反演效果;4)在上述两步结论上进一步改进模型,提出了基于两步预报法的 B-pix2pixGAN 模型。

1.4 章节安排

本文组织结构如下:

第一章 绪论

主要阐述卫星反演降水的研究背景及意义,国内外基于单通道法卫星反演降水与基于多通道法反演降水算法的研究现状,基于机器学习与深度学习算法反演降水的研究现状。并对于文章的研究内容进行简要介绍。

第二章 资料与方法

介绍本研究使用的葵花 8 静止卫星数据资料和 GPM IMERG 卫星降水资料,使用的深度学习模型及降水评价指标。并且介绍本研究的卫星资料和 GPM IMERG 降水资料按气象条件的筛选方法和深度学习的训练样本的构建方法。

第三章 不同通道对降水反演的影响

基于 U-Net 深度学习模型分别用单通道和多通道组合作为模型输入的降水反演实验,分析其对降水反演效果及其特征。计算降水评价指标定量统计不同通道反演效果,选出最优的葵花 8 降水反演红外通道,为下一步反演试验提供最优通道方案。

第四章 不同时刻卫星观测对降水反演的影响

基于 U-Net 深度学习模型分别用不同时刻组合的降水反演实验,并且分析了其反演效果及其特征,根据降水评价指标定量统计不同时刻反演效果,选出最优的葵花 8 降水反演时刻,为下一步反演试验提供最优时刻方案。

第五章 比较不同深度学习模型的对降水反演的影响

比较不同深度学习模型,包括 U-Net、ConvMixer 和 pix2pixGAN,对降水反演的效

果。根据降水评价指标定量统计不同模型反演效果，选出最优的模型并分析其存在的问题，在此基础上改进对最优模型，分析和比较改进模型的反演效果。

第六章 结论与展望

对本研究的工作内容和结论进行总结，针对其中存在的不足进行讨论和提出解决思路，以及提出对未来工作的进一步展望。

第二章 资料与方法

本章介绍本文研究使用的葵花 8 卫星数据以及作为目标变量的 GPM IMERG 数据，并且简单地比较葵花 8 和 GPM IMERG 降水数据的一致性。然后，介绍资料预处理方案，包括对数据进行数据清洗和整理，并构建数据集。接着，我们对于文中使用的模型结构、特征等进行介绍。最后，介绍所使用的反演降水的评价指标。

我国东南沿海地区经济相对发达，是全球人口最为稠密的地区之一。因受热带/亚热带季风影响，其降水季节差异大，尤其夏季，强对流易发生发展，且频遭台风影响。该地区对于降水观测信息的精度和要求较高，因此本研究区域选择以广东、海南省为主，针对该区域展开葵花 8 号静止卫星降水反演技术的研究。

2.1 资料介绍

2.1.1 静止卫星数据

本研究采用葵花 8 号静止卫星资料（Himawari-8，下简称葵花 8）作为降水卫星数据反演的基础数据集。葵花 8 由日本气象厅于 2014 年 10 月发射，并于 2015 年 7 月正式启用，探测范围覆盖了包括中国在内的东亚和西太平洋区域（ $60^{\circ}\text{N}-60^{\circ}\text{S}$ ， $80^{\circ}\text{W}-160^{\circ}\text{W}$ ）。葵花 8 定位于东经 140.7° 的轨道上，其时间分辨率为 10 分钟，水平分辨率为 $0.05^{\circ}\times 0.05^{\circ}$ ，最低可达 0.5-2km，共拥有 16 个通道，包含 3 个可见光通道，3 个近红外通道和 10 个红外通道。卫星数据由日本宇宙航空研究机构（JAXA）发布，注册后在其官方网站上即可获得实时更新的数据。

本研究使用的葵花 8 号的各红外通道的信息见表 2-1。

表 2-1 葵花 8 卫星 AHI 传感器通道信息

类别	通道	中心波长 (μm)	空间分辨率 (km)	应用
水汽	8	6.2	2	对流层上层水汽、降水、高层水含量估计
	9	6.9	2	对流层中层水汽
	10	7.3	2	对流层中低层水汽
长波红外	11	8.6	2	云相态、薄卷云、SO ₂
	12	9.6	2	臭氧、大气湍流、风切变
	13	10.4	2	低层水汽、云、表面特性
	14	11.2	2	晴空大气水含量、SST、云、降水
	15	12.4	2	晴空大气水含量、火山灰、SST
	16	13.3	2	云高、云量、大气温度

相比于以往的静止气象卫星，葵花 8 号搭载了先进的气象观测遥感器可视红外线辐射计 AHI (Advanced Himawari Imager)，能实现对大气云降水活动的更高频率的观测，同时在传感器的波长种类和数据分辨率等硬件性能方面也得到大幅度增强^[56]。葵花 8 采用了大量先进技术，包括高分辨率成像技术、高性能数字处理技术、可靠的卫星控制技术，这些技术使得葵花 8 拥有高精度的观测能力，可以实现高分辨率、高精度的同步观测，能够捕捉气象图像中细节和变化，因此是当前较为先进可靠的气象监测卫星。目前葵花 8 气象观测数据和产品已被广泛应用于天气预报、气象灾害监测预警、农业生产管理、城市规划与建设、环境保护等方面。由于该卫星上 AHI 传感器提供的红外信息能够很好的捕捉到高时空分辨率时空变化的降水过程，因此非常适合本研究中的降水反演技术。

本研究建立的降水反演模型主要采用葵花 8 号卫星 3 个水汽通道（通道 8、通道 9、通道 10）和 6 个长波红外通道（通道 11-16）的亮温数据。其中三个水汽通道分别对应对流层上、中、低层水汽，而六个长波红外通道包含云相态、低层水层、晴空水汽含量、云高、云量等信息。

2.1.2 降水基准数据

针对降水观测较为先进的天基平台则是全球降水测量卫星 GPM，全球卫星降水计划是一项国际卫星任务，由 NASA 和 JAXA 合作开展，利用多传感器多卫星多算法结合卫星网络和雨量计反演得到更高精度的降水数据，其能够提供全球范围基于微波的 3h 以内以及基于微波红外的半小时的雨雪数据产品，范围延伸至南北极圈。该卫星搭载了更先进的 GMI 被动微波辐射计和双频雷达传感器 DPR，加强了对微量降水和固态降水的识别能力，为多卫星降水估计提供了精确的校准参考，因此本研究选取全球卫星降水计划（GPM）的降水产品作为降水基准数据。

GPM IMERG 的数据产品主要分为：Level-1、Level-2 和 Level-3。其中 Level-1 数据是卫星上观测仪器得到的基数据，Level-2 的数据是以 Level-1 为基础在相同的分辨率和位置上处理导出的地球物理变量。Level-3 的数据基于 IMERG(Integrated Multi-satellite Retrievals for GPM)算法生成，该算法结合来自 GPM IMERG 中所有无源微波仪器的数据提供降水估计，在精细时间和空间分辨率上对所有卫星微波降水估计值以及微波校准红外(IR)卫星估计值、降水量分析以及可能的其他降水估计量进行相互校准、合并和插值，并于观测后 4 小时发布早期产品，观测后 12 小时发布后期产品，观测 3.5 月后发布最终产品。Level-3 的最终产品较为完整且一致性高，本研究将其作为降水基准值，并作为模型训练中的目标变量。该数据单位为 mm/h，水平分辨率为 $0.1^{\circ} \times 0.1^{\circ}$ ，时间分辨率为 30 分钟。

2.1.3 葵花 8 数据与 GPM IMERG 降水数据的比较

本研究主要使用的数据是葵花 8 红外亮温数据和 GPM IMERG 降水数据，通过比较和分析我们可以得知，葵花 8 亮温与 GPM IMERG 降水数据具有较强的一致性。其中 GPM IMERG 资料数据的时间分辨率 30 分钟。图 2-1 是起止时间是 2016 年 5 月 3 日 6 时至 2016 年 5 月 3 日 6 时 30 分的葵花 8 卫星红外通道以及个别通道差与 GPM IMERG 降水的比较。可以看到，降水与通道 12-16 具有较好的一致性，且与多个通道组合的通道差均具有较好的一致性。由此可以看出葵花 8 卫星的红外具有较高的质量，具备反演降水概率和强度的能力。

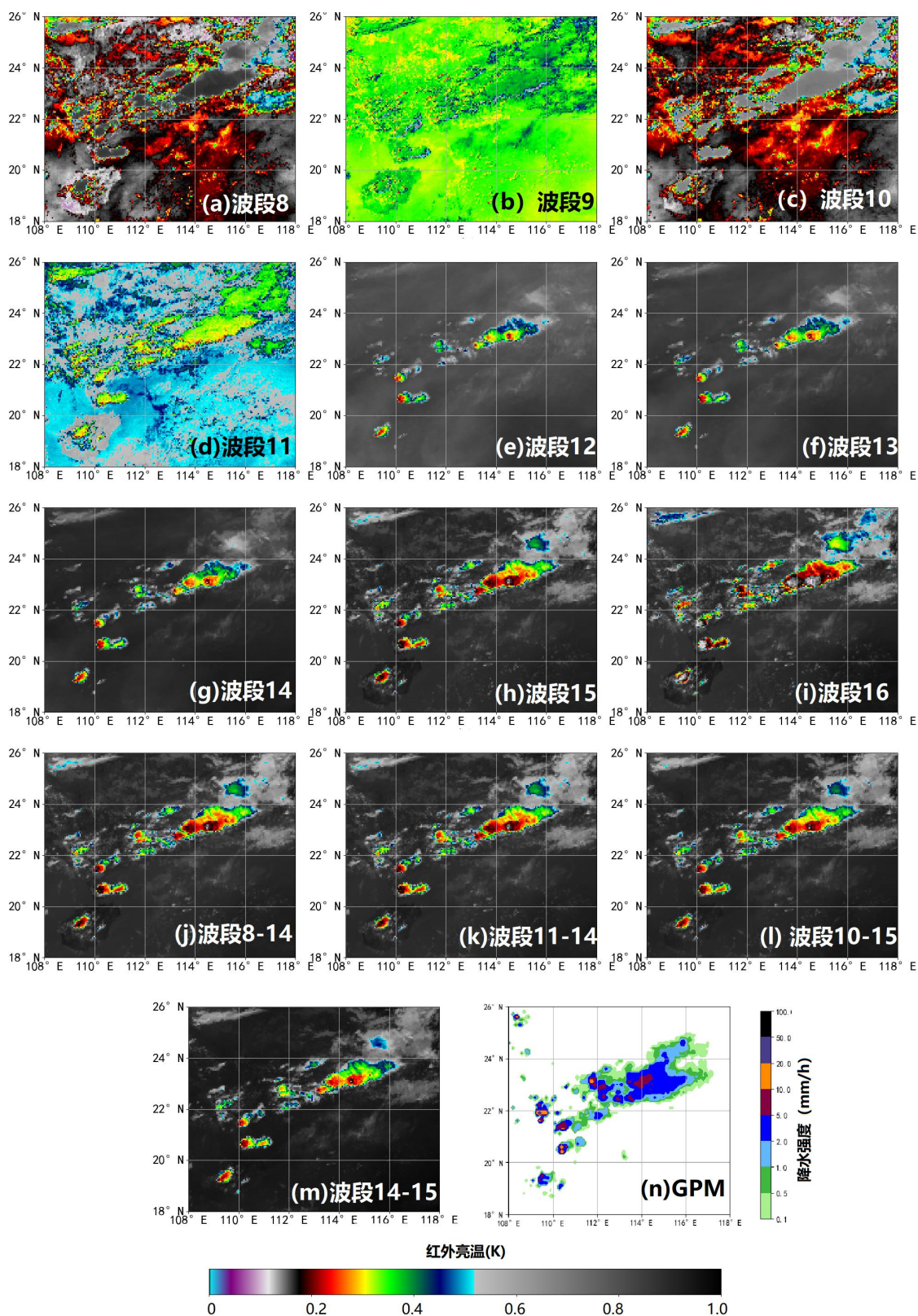


图 2-1 2016 年 5 月 3 日 6 时华南一次降水过程的葵花 8 和 GPM 数据结果:(a)~(i)葵花 8,单通道(8~16)的亮温; (j)~(m)葵花 8, 不同通道的亮温差; (n)GPM 降水数据。

研究发现通过计算适当的亮温差组合可以推出云的物理性质，从而更准确地识别深

对流层。一般而言，水汽通道与长波红外通道之间的亮温差（波段 8-波段 14 和波段 10-波段 15）对云顶高度的变化较敏感。对于云层上部的冰云或水云等云相态类型信息则可以通过 8-11 μm 和 11-12 μm 之间的差异（波段 11-波段 14 和波段 14-波段 15）来获取^[45]。研究已经表明这些亮温差的组合可以为识别对流降水主导的降水区域提供有利信息^[36]。图 2-1 展示了这几个亮温差，可以看出其与 GPM IMERG 降水的分布相比具有较高的一致性，尤其在强降水区域。然而，并非所有红外通道的特征均与降水一致，如通道 8、9，且部分红外通道所表示的云体结构信息并不完全对应着降水，而这些无关信息对应着杂波，无法被深度学习模型拟合。因此，分析不同通道对模型反演结果的影响，并选择最优通道组合是本研究的目标之一。

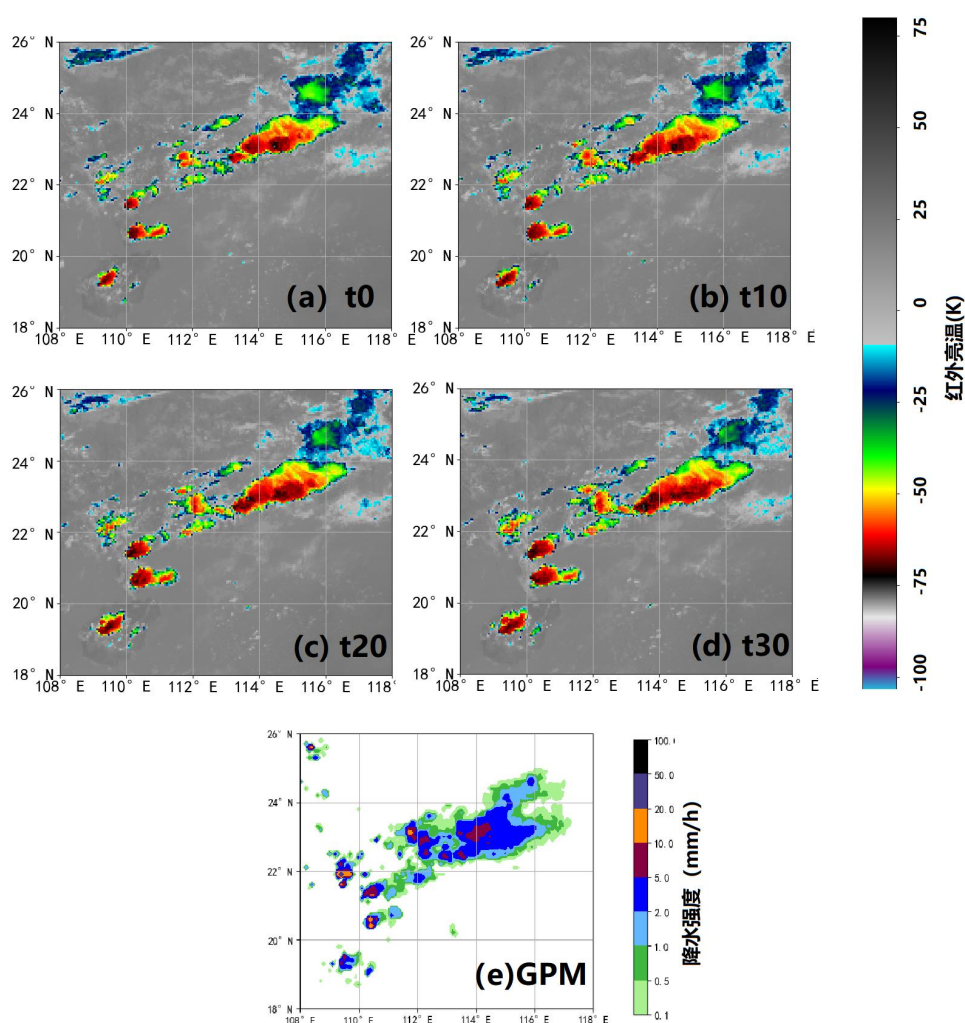


图2-2 2016年5月3日6时华南一次降水过程的葵花8和GPM数据结果:(a)~(d)葵花8,多时刻(t0~t30)的亮温;(e)GPM观测数据。

最后，图 2-2 分别是起止时间是 2016 年 5 月 3 日 6 时至 2016 年 5 月 3 日 6 时 30

分、2016年5月3日6时30分至2016年5月3日7时、2016年5月3日7时至2016年5月3日7时30分的葵花8卫星红外通道13和GPM降水的比较。通过对葵花8多个时刻的观测结果和GPM IMERG降水数据的进行比较,我们发现GPM IMERG和葵花8卫星不同时刻数据存在一定的差异。其中葵花8的观测为单一时刻的观测结果,而GPM IMERG降水数据则是30分钟时段内的平均值。从红外亮温和GPM IMERG降水的对比(图2-2)可知,在30分钟的时间间隔下,不仅降水区域具有明显的变化,而且表现在大雨区域的不断扩大。单个时刻的红外亮温数据是无法捕捉到降水30分钟内的变化情况。那么多时刻降水反演的作用就体现的尤为重要。且从红外亮温图来看,亮温图的具体表现为低温区域由低温中心逐渐向外围扩散,而对应的降水区域由低温中心对应的小雨区逐渐转为大雨区,而周围区域则逐渐转为雨区。这些比较说明了葵花8和GPM IMERG数据的时间对应关系并不明确,因此有必要对不同时刻以及多时刻输入进行反演试验和比较,选出最优时刻。

2.1.4 数据集构建

本研究聚焦华南地区,将研究区域定义为东经 109° 至东经 117° 和北纬 18° 至北纬 26° 区域,主要包含广东省、海南省及其近海一部分海域。该区域属于东亚季风区,是中国年降水量最多的区域之一。由于GPM IMERG数据的时间分辨率为30分钟,而葵花8卫星资料为10分钟,因此基于葵花8反演GPM IMERG数据不仅可以采用多通道,也可以使用不同时刻的卫星资料或其组合。我们将模型输入维数为定义为 $160 \times 160 \times m \times n$,其中160代表经度方向的格点数,160代表纬度方向的格点数, m 为输入通道个数, n 为输入卫星图像个数。为了更好的分析和比较葵花8卫星反演的输出结果和GPM基准值,我们将最终模型标签以及模型输出取相同的维数定义,即 $80 \times 80 \times 1$,其中80代表经度东西方向的格点数,80代表纬度方向的格点数,1代表单个时刻的降水场。

为了筛选有效的降水样本,本研究将数据样本集中降水数值大于 2mm/h 的格点数大于320个(总格点数的 $1/20$)的样本作为有效样本从而保证每个样本训练都包含有雨的格点对应,剔除降水强度过小的样本。此外,由于葵花8资料存在通道数据缺失的情况,我们将对应样本删除。最终,我们共获得8920个有效样本。将数据样本按照时间顺序存储,按6:2:2的比例划分训练集、验证集和测试集。

2.2 深度学习模型简介

深度学习作为一种新兴技术，不断有新模型被提出，但这些模型大多被应用于计算机视觉（CV）和自然语言处理（NLP）领域，与本研究的降水反演目标并不一致。因此，本研究将聚焦三种不同类型的深度学习模型，探究各模型在降水反演效果和适配性。

2.2.1 U-Net 模型介绍

如图 2-3，U-Net 模型是一种改进的全卷积神经网络（FCN），由于其结构形似字母 U 而得名。U-Net 最初被应用于图像语义分割，但由于在反演问题上的出色表现，U-Net 已逐渐成为在大气科学研究中应用最为广泛的深度学习模型。U-Net 是典型的编解码模型，由编码器（Encoder）和解码器（Decoder）组成。其中，编码器是典型的卷积神经网络结构，它重复采用 2 个卷积层和 1 个最大池化层的结构，每进行一次池化操作后特征图的通道就增加 1 倍，而特征图的长、宽也减小一倍。在解码器中，每进行 1 次反卷积操作，特征图的通道数就减半，而长、宽翻倍。U-Net 模型通过长连接（skip-connection）的方法，将编码器中各级高维抽象特征与解码器中通过反卷积放大的特征拼接，重新组成一个 2 倍大小的特征图，再采用 2 个卷积层进行特征提取。在最后的输出层，使用 K 个 1×1 的卷积层将 64 维的特征图映射成最终反演结果， K 取决于反演量的维数。

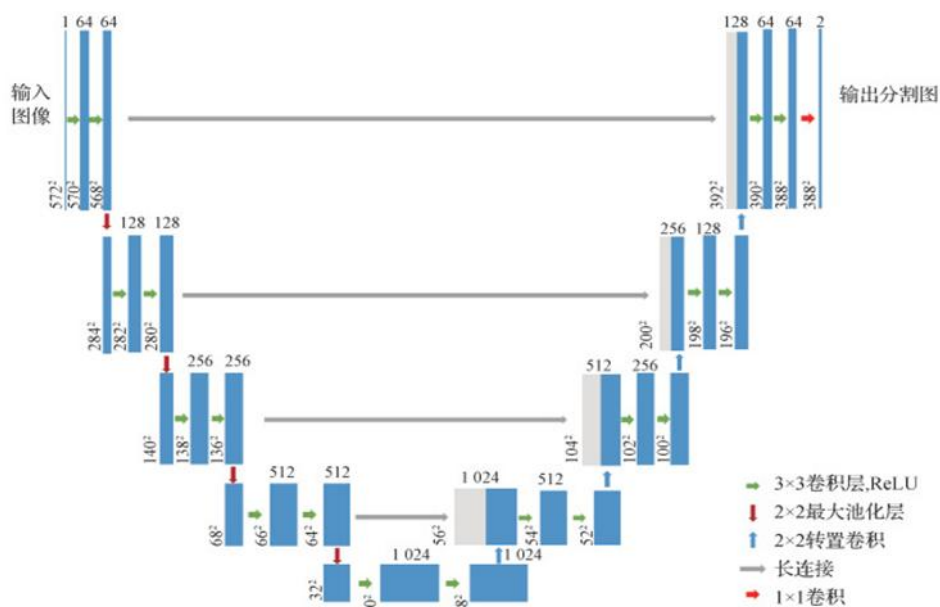


图 2-3 U-Net 模型结构图

U-Net 模型是 FCN 的改进和延伸，它沿用了 FCN 中编解码的思想，并融合了跳跃网络的特点多次进行长连接操作，在模型结构上更加优雅且巧妙，主要体现在以下两点：

(1) 采用编解码结构。编码器逐级提取输入的高维抽象特征，解码器逐步将抽象的特征图放大到输入大小，并还原影像的位置信息，且 U-Net 模型的每一个隐藏层都有较多的特征维数，这有利于模型学习更加多样、全面的特征。

(2) 多次使用长连接结构。一般而言，编解码模型中，编码器提取的特征越抽象，解码器的解码效果在宏观上越好，但是更容易丢失细节特征。而长连接结构能够实现编码器中各级抽象特征在输出前的融合，使得模型能够从上下文信息和细节信息由组合得到更加精确的输出特征图。

2.2.2 ConvMixer 模型介绍

ConvMixer 模型是 Google Brain 团队在 2021 年提出的全卷积神经网络模型，主要用于图像分类任务（图 2-4）。ConvMixer 模型受 Vision Transformer (ViT) 启发，将输入的图片划分为多个块，并对每个块应用卷积操作，然后利用 1×1 卷积实现块间的信息交互，从而利用各个块的特征信息。此外，模型还使用类似于 Transformer 的 MLP 模块来增强模型的非线性表示能力。该模型由于采用了分块的方法，在处理较大输出时表现良好。

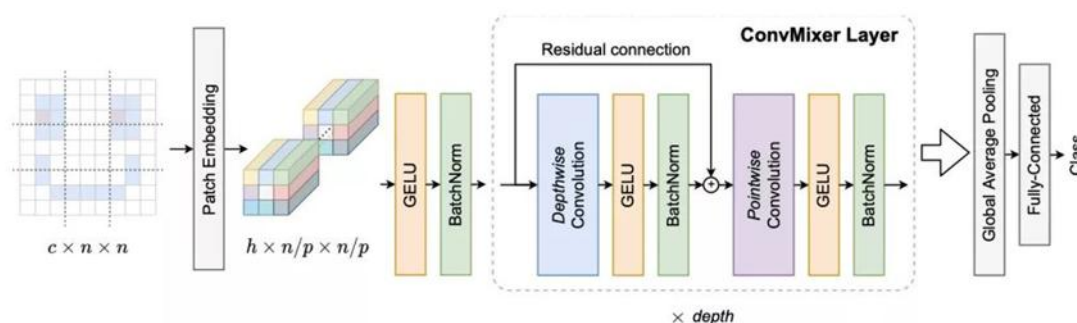


图 2-4 ConvMixer 模型的结构图

与经典卷积神经网络模型相比，ConvMixer 模型具有以下优点：

(1) 参数更少。相对于 ResNet 等传统卷积神经网络，ConvMixer 模型参数更少，因此可以更方便地进行部署和推理。

(2) 模型结构简洁但高效。ConvMixer 模型的结构较为简单，易于理解和修改，使得该模型可以更适应不同的任务需求，具有较好的性能表现：实验结果表明，ConvMixer 模型在 ImageNet、CIFAR-10 和 CIFAR-100 数据集上的分类性能与其他先进的卷积神经网络相当，并且在某些情况下表现更好。

总的来说, Convmixer 模型将卷积神经网络和 Transformer 网络的优势相结合, 通过较少的参数和简洁的网络结构, 在图像分类任务上取得了很好的性能表现, 并有着广泛的应用前景, 是当前卷积神经网络中较为先进的模型。由于本研究是对降水二维场进行预测, 因此将最后一层卷积的输出作为反演结果, 并未使用全连接层。

2.2.3 pix2pixGAN 模型介绍

pix2pixGAN 是一种基于生成对抗网络 (GAN) 的图像翻译模型 (图 2-5), 不同于以上介绍的两种模型, pix2pixGAN 是一种条件生成模型, 可以根据给定的条件生成具有所需特征的图像或数据, 被广泛应用于图像翻译、图像超分辨率、图像风格迁移等领域, 为各项任务提供了高效的解决方案。此外, pix2pixGAN 具有强大的风格迁移能力, 与反演问题具有高度适配性。

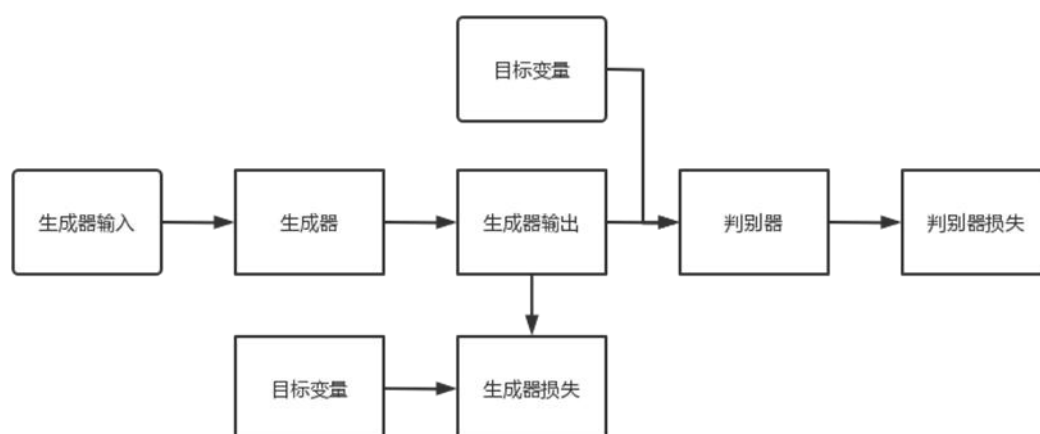


图 2-5 pix2pixGAN 模型的结构图

pix2pixGAN 的核心思想是使用两个深度神经网络, 一个生成器和一个判别器, 互相竞争来完成图像翻译任务。生成器的作用是将输入的图像通过一系列的卷积、反卷积、残差连接等操作, 生成目标图像。而判别器则尝试区分生成器生成的图像与真实的目标图像之间的差异, 从而指导生成器不断进行改进。通过生成器和判别器互相博弈, 不断提高模型的精度和鲁棒性, 从而达到更高质量的生成效果。同时, 为了增强模型的表达能力, pix2pixGAN 还采用了多尺度判别器, 用来检测不同分辨率下图像的细节特征。这一模型具有较强的实用价值, 广泛应用于图像翻译、图像超分辨率、图像风格迁移等领域, 为各项任务提供了高效的解决方案。

然而，与其它 GAN 模型相同，pix2pixGAN 也存在几个问题。首先，生成器的训练需要大量的数据集和时间，否则容易出现过拟合的情况，导致模型的泛化能力不足。同时，模型的训练过程中也容易出现梯度消失或梯度爆炸等问题，需要采取一系列的技巧进行调整。

2.3 反演结果的评估方法

本研究除了对于典型的几个案例进行可视化的对比和分析外，还需要对大量的模型反演结果和 GPM IMERG 数据进行定量对比和检验。因此这里我们引入了多个气象学常用的验证指标，其中主要包含采用命中率 (POD)、虚警率 (FAR)、临界成功指数 (CSI)、HSS 评分、ETS 评分和偏差 Bias。

通过对于引用上述评价指标，我们在模型反演和 GPM IMERG 降水资料的对比和分析部分中，提供精确和客观的降水反演效果的评估和比较。本文所采用的指标计算方法如表 2-2 所示。

表 2-2 降水指标评分具体含义

名称	公式	范围	最优值
命中率 (Probability of detection)	$POD = \frac{a}{a+c}$	[0, 1]	1
虚警率 (False-alarm ratio)	$FAR = \frac{b}{a+b}$	[0, 1]	0
临界成功指数 (Critical success index)	$CSI = \frac{a}{a+b+c}$	[0, 1]	1
HSS 评分 (Heidke skill score)	$HSS = \frac{2(ad-bc)}{(a+c)(c+d) + (a+b)(b+d)}$	$[-\infty, 1]$	1
ETS 评分 (Equitable threat score)	$ETS = \frac{a-dr}{a+b+c-dr}$ where $dr = \frac{(a+c)(a+b)}{a+b+c+d}$	$[-1/3, 1]$	1
偏差 (Bias)	$Bias = \frac{a+b}{a+c}$	$[0, \infty]$	1

其中， a 表示模型反演有雨，观测有雨； b 表示模型反演晴空，观测有雨； c 为模型反演有雨，观测晴空； d 为模型反演晴空，观测晴空。

其中，POD 表示在实际观测到有降水时模型对降水像素点进行正确识别的可能性。FAR 显示了在模拟到的降水像素点中模型识别降水的错误率。CSI 为当移出被正确分类的非降水点时，正确分类的像素点中的降水像素点的比例。HSS 计算了在排除那些完全由于随机性而正确模拟的像素点后，被正确分类的像素点的比例。ETS 表示了考虑偶然性之后被正确分类的像素点占比，得分相对更公正。此外，本研究采用偏差 Bias 描述与实际观测情况相比，模型正确分类降水的频次以及对降水情况低估（Bias<1）或高估（Bias>1）的趋势。

表 2-3 中国短时降水等级表

雨级	5 分钟	10 分钟	30 分钟	1 小时
短时小雨	<0.2	<0.3	<0.7	<2
短时中雨	0.2-0.5	0.3-0.8	0.7-1.8	2.0-4.0
短时大雨	0.5-1.2	0.8-2.0	1.8-4.0	4.0-8.0
短时暴雨	1.2-4.0	2.0-6.0	4.0-12.0	8.0-20.0

为了评估模型对不同强度降水的反演能力，并考虑到本文的研究区域主要以降雨为主。因此我们基于中国短时降雨等级（表 2-3），对于反演结果的降雨等级进行分类，并根据各类评估指标（表 2-2）对个等级降水分别进行检验和分析。其中，由于 GPM IMERG 的降水资料时间分辨率是 30 分钟，因此我们也主要采用表 2-3 中 30 分钟内雨量等级划分进行反演结果的分级检验。需要注意的是，如果该方法应用于其他地区，则可能需要更具当地的纬度、海拔和气候类型对各类细分降水的进行分别讨论，如降雨、降雪、冰雹、霰等过程。

第三章 不同卫星观测通道对降水反演的影响

本章节中,主要利用不同通道以及多通道组合的方式,对葵花 8 卫星进行降水反演。首先指出,葵花 8 卫星的各红外亮温通道对应着不同大气成分的观测信息(表 2-1),如通道 8、通道 9、通道 10 是观测水汽的通道,这三个通道分别观测对流层上层水汽、对流层中层水汽和对流层中低层水汽,通道 12 主要用于观测臭氧,而通道 16 得到的主要是云顶高度的信息。然而并非所有通道信息都对降水反演有益,部分通道由于含有有益信息较少,在深度学习模型中会成为“噪声”,从而干扰模型反演效果。因此,本章将分别测试单通道与多通道对模型反演结果的影响,并选出最优输入通道,试验均采用 U-Net 模型作为基准模型。虽然不同模型之间会存在一定的差异,但以往的研究表明由于不同通道的观测目标信息存在很大差异,因此不同通道的反演差异会显著大于不同模型之间的差异。本文 U-Net 模型的超参数配置如下: learning rate: 0.001; epoch: 100; batchsize: 128。为了保持在训练过程中每个实验都获得良好的结果,使用了利用 ETS 之和(阈值 0.1/1.4/3.6)保存了在训练 100 轮次的过程中,每 2epoch 进行一次验证集的验证,验证集在模型表现上最优被保存为测试模型。本章使用命中率 POD,虚警率 FAR,临界成功指数 CSI, HSS 评分, ETS 评分和偏差 Bias 来评估模型精度。

此外,不同降水等级的降水过程需要进行一定的分类检验,来反应不同通道针对不同等级降水的效果和适配性。因此,本研究在降水强度不同阈值下(0.01mm/h 和 1.4mm/h)进行不同通道的检验。其中当降雨强度大于 0.1mm/h 的格点判断为有降水,小于 0.1mm/h 的格点判断为无降水。

3.1 不同卫星观测单通道的降水反演

在训练过程中,随着 epoch 的增加,模型会更加拟合数据集本身,根据多个单通道(通道 11、12、13)实验发现,在训练 10epoch 之后,验证集 RMSE 不会随着 epoch 变化,说明模型的泛化性能较好,不会出现过拟合现象。如图 3-1,为了更好的评价模型性能,本文采用三个阈值无雨(<0.1)、小雨($<1.4 \> 0.1$)、中雨($<3.6 \> 1.4$)之和判断模型性能。从图 3-1 得出,中雨、大雨的指标评分先趋于稳定,有雨的指标评分逐渐升高后平稳,故采用三者之和作为判断模型的评价标准。

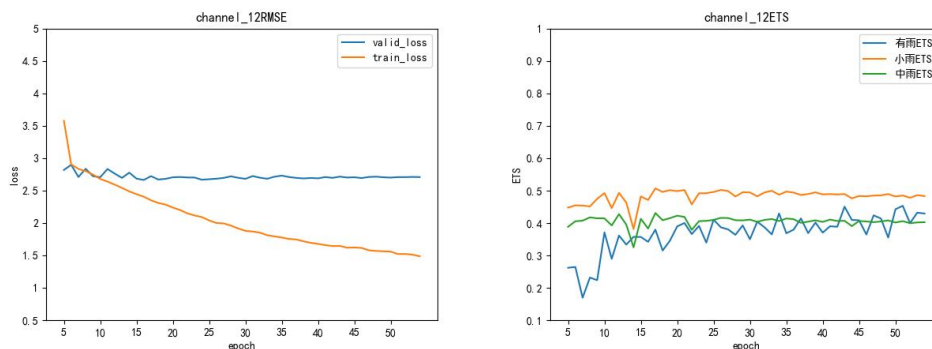


图 3-1 通道 12 作为模型输入验证集随着 epoch 的变化曲线(左图为 RMSE,左图蓝线为验证集 RMSE, 橙线为训练损失 (RMSE), 右图为 ETS, 蓝线为有雨, 橙线为小雨, 绿线为中雨)

本章对各通道反演结果均在测试集上的表现进行统计分析, 由于国际标准规定大于 0.1mm/h 的代表有降水发生, 将阈值设置为 0.1mm/h 来评判各个波段反演结果的降水区域的反演情况, 并计算各项指标, 结果如下表 3-1 所示。

表 3-1 葵花 8 单通道和多通道的反演降水指标结果 (阈值为 0.1mm/h, 加粗表现最优)

	通道 8	通道 9	通道 10	通道 11	通道 12	通道 13	通道 14	通道 15	通道 16	全通道	三通道
POD	0.73	0.78	0.8	0.82	0.81	0.83	0.9	0.85	0.82	0.92	0.92
CSI	0.49	0.49	0.5	0.52	0.53	0.52	0.49	0.5	0.51	0.52	0.48
HSS	0.45	0.44	0.46	0.48	0.5	0.48	0.41	0.44	0.46	0.46	0.39
ETS	0.29	0.28	0.3	0.31	0.33	0.32	0.26	0.29	0.3	0.29	0.24
BIAS	1.25	1.37	1.4	1.4	1.35	1.42	1.74	1.55	1.45	1.68	1.82
FAR	0.59	0.57	0.57	0.58	0.6	0.59	0.52	0.55	0.57	0.54	0.5

表 3-2 葵花 8 单通道的反演降水指标结果 (阈值为 1.4mm/h)

	通道 8	通道 9	通道 10	通道 11	通道 12	通道 13	通道 14	通道 15	通道 16	全通道	三通道
POD	0.54	0.6	0.59	0.62	0.63	0.61	0.64	0.62	0.61	0.82	0.82
CSI	0.4	0.42	0.42	0.44	0.44	0.44	0.44	0.43	0.44	0.46	0.42
HSS	0.52	0.53	0.53	0.55	0.55	0.55	0.55	0.54	0.53	0.56	0.51
ETS	0.35	0.36	0.36	0.38	0.38	0.38	0.38	0.37	0.36	0.39	0.34
BIAS	0.88	1.04	0.99	1.03	1.05	1.02	1.11	1.09	1.07	1.6	1.78
FAR	0.61	0.58	0.6	0.6	0.6	0.6	0.58	0.57	0.57	0.51	0.46

从表 3-1 的单通道反演的指标情况来看,表中加粗代表单通道最优值,降雨区域存在着明显的小雨区的高估,这与图 3-1、图 3-2、图 3-3 案例的结论基本一致。其中各个通道对于大多数降雨评估指标(除命中率外)评分较低,各个单通道命中率 POD 较高(0.73~0.9),其中通道 14 的命中率最高。在其他较严格指标中,通道 12 在各个单通道 4 项评分中,均取得了指标评分最优值。只有偏差 Bias 中,通道 12 不是最优。综合比较单通道反演结果,通道 12 的降水区域的估计指标评价最优。

依照中国短时降水标准(表 2-3),半小时累积降水量达到 0.7 毫米为中雨,即降水强度为 1.4mm/h 为中雨。为了评判模型的中雨区及以上降水强度的反演情况,本文以 1.4mm/h 作为阈值来定义各个波段反演结果中雨及以上的降水,并计算各个单通道反演结果的检验指标,其结果如表 3.2 所示。首先,各单通道的命中率(POD)均介于 0.54-0.64 之间,低于全通道和三通道(通道 11、通道 12、通道 13)作为模型输入的结果。这可能是因为全通道/三通道的偏差(BIAS)较大,容易高估降水。事实上,各通道反演结果均有高估倾向(BIAS>1),这个现象在降雨区中也有体现,这可能是因为模型采用均方差作为损失函数,具有关注高值区的倾向。通道 13 反演结果的 BIAS 为所有通道最小(1.02),但其反演效果相比其他通道较差。结合从公平性评分(ETS)、关键成功指数(CSI)、HSS 等指标我们都可以看出通道 11、通道 12、通道 13、通道 14 的反演效果是并列最优值,总体优于三通道,低于 9 个通道当做模型输入。因此对于做模型反演降水来说,根据阈值为 0.1mm/h 与 1.4mm/h 的指标总体评价,选择通道 12 作为模型反演的输入数据是最优选择。

首先,针对葵花 8 卫星 9 个红外通道逐一进行单通道试验。本研究将选取 3 个典型案例,逐一分析各个通道对反演结果的影响。

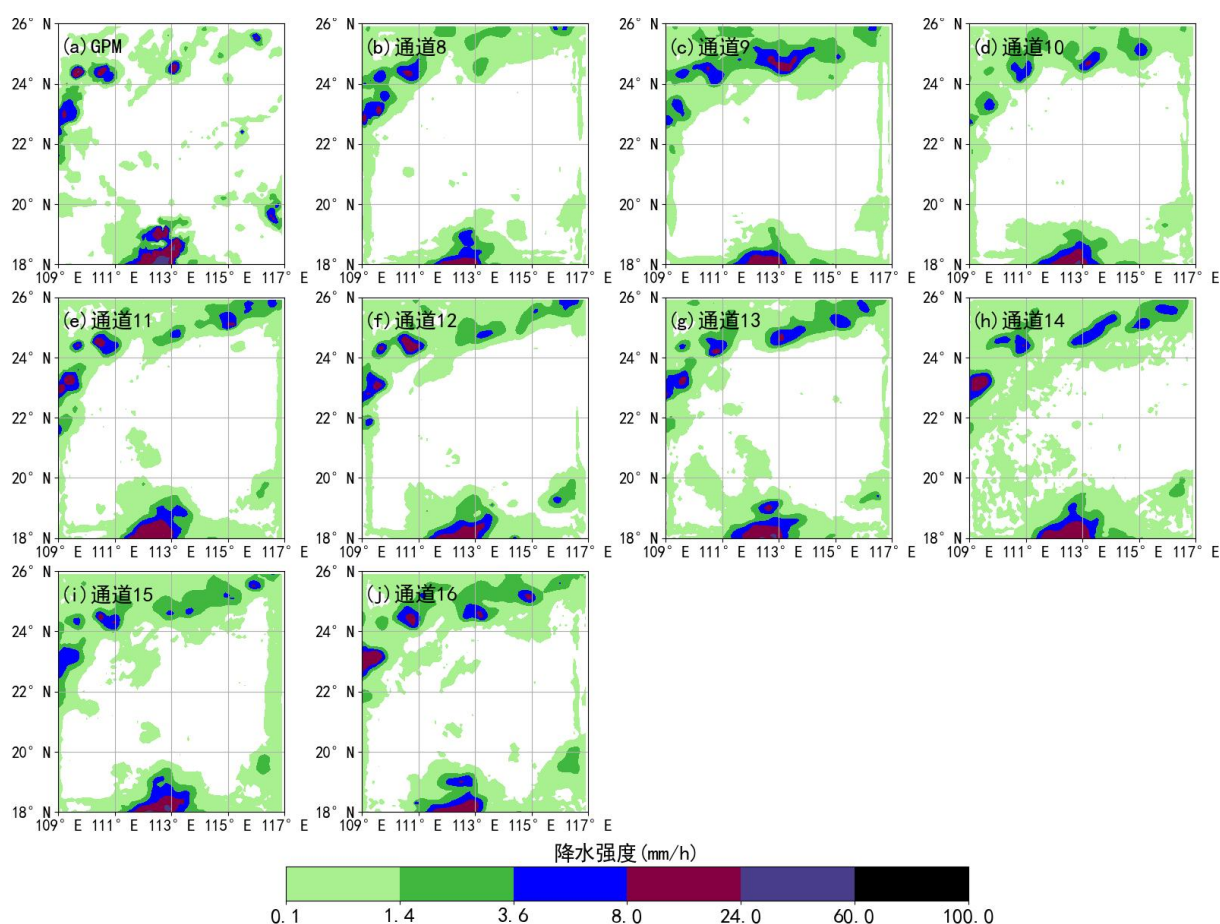


图 3-2 2019 年 6 月 30 日 16 时华南地区一次降雨过程的不同通道反演结果：(a) GPM IMERG 降水数据；(b) - (j) 波段 8~波段 16 各段反演结果。

图 3-2 展示了 2019 年 6 月 30 日 16 时各通道反演的降水强度。图 3-2 (a) 中显示了 GPM IMERG 降水数据结果，从图中可以看出，降水中心位于本研究区域的南方，且都是短时暴雨和短时大雨区，中间没有短时中雨区过度，降雨强度由中心区域向外逐渐降低，在研究区域中心分布着一块暴雨区（褐色区域）。图 3-2 (b) - (j) 分别代表通道 8 至通道 16 的模型反演结果，各反演结果的降雨中心的位置均在研究区域的南方，虽然研究区域中心的暴雨区在基于各同道的反演结果上都有不同程度的体现，但是降水中心的范围和反演强度与 GPM 数据有一定的差异。整体上看，通道 13 的降水中心区域反演效果从降水强度和位置来说均较好，且通道 13 的小雨反演区域误报也较少。东经 117°、北纬 20° 的区域大雨各个单通道反演结果都没有体现。综上，该案例的反演结果中各个单通道反演结果都是基本反演出降雨范围和降雨强度，细节表现方面，通道 8、通道 9 在研究区域的北方区域较差，误报较多，通道 11、通道 12、通道 13 在小雨区的误报率和大雨区预报准确度均较高。

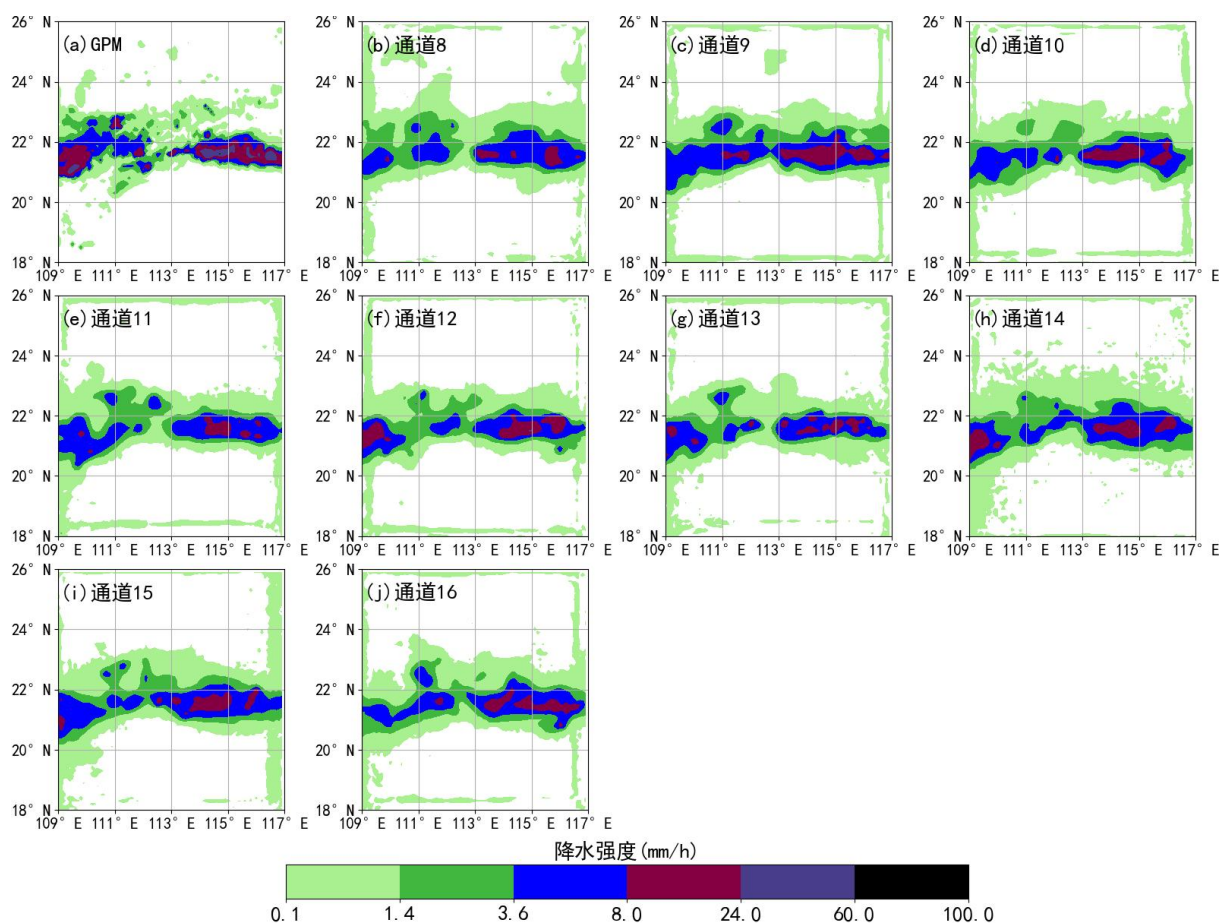


图 3-3 2019 年 8 月 17 日 3 时华南地区一次降雨过程的不同通道反演结果：(a) GPM IMERG 降水数据；(b) - (j) 通道 8~通道 16 各段反演结果。

图 3-3 展示的是 2019 年 8 月 17 日 3 时各通道反演的降水强度，以及 GPM IMERG 降水数据。该案例降水中心分布在研究区域的正东方和正西方，降水区域整体上比较集中，连贯性较强。但是从各个波段的总体反演结果来看，GPM IMERG 的两个降水中心在各波段的反演的结果上均有体现，对于小雨区域的误报均存在不同程度的问题。且不少波段的反演结果在研究区域西南方向存在的小雨区的误报。此外，各波段的反演结果对研究区域的中心的暴雨区均存在明显的低估，这可能是由于降水区域在亮温上的特征已经逐渐消散，由于降水发展时间滞后于亮温的低温时间，才会出现明显的低估。

而从各波段反演的结果看，通道 11、12、13 的小雨区的误报相对较小，且将雨区的分布结构反演结果更为接近 GPM 数据结果。对于通道 12、通道 14，对于研究区域正西方的降水中心的暴雨区比其他通道反演更好，但是通道 14 小雨误报较多。而对于大雨、暴雨区域和强度反演角度而言，除通道 12、14 外均未呈现出研究区域西方的暴雨

强度，尤其通道 11 的结果中研究区域东方的暴雨区也没有体现。

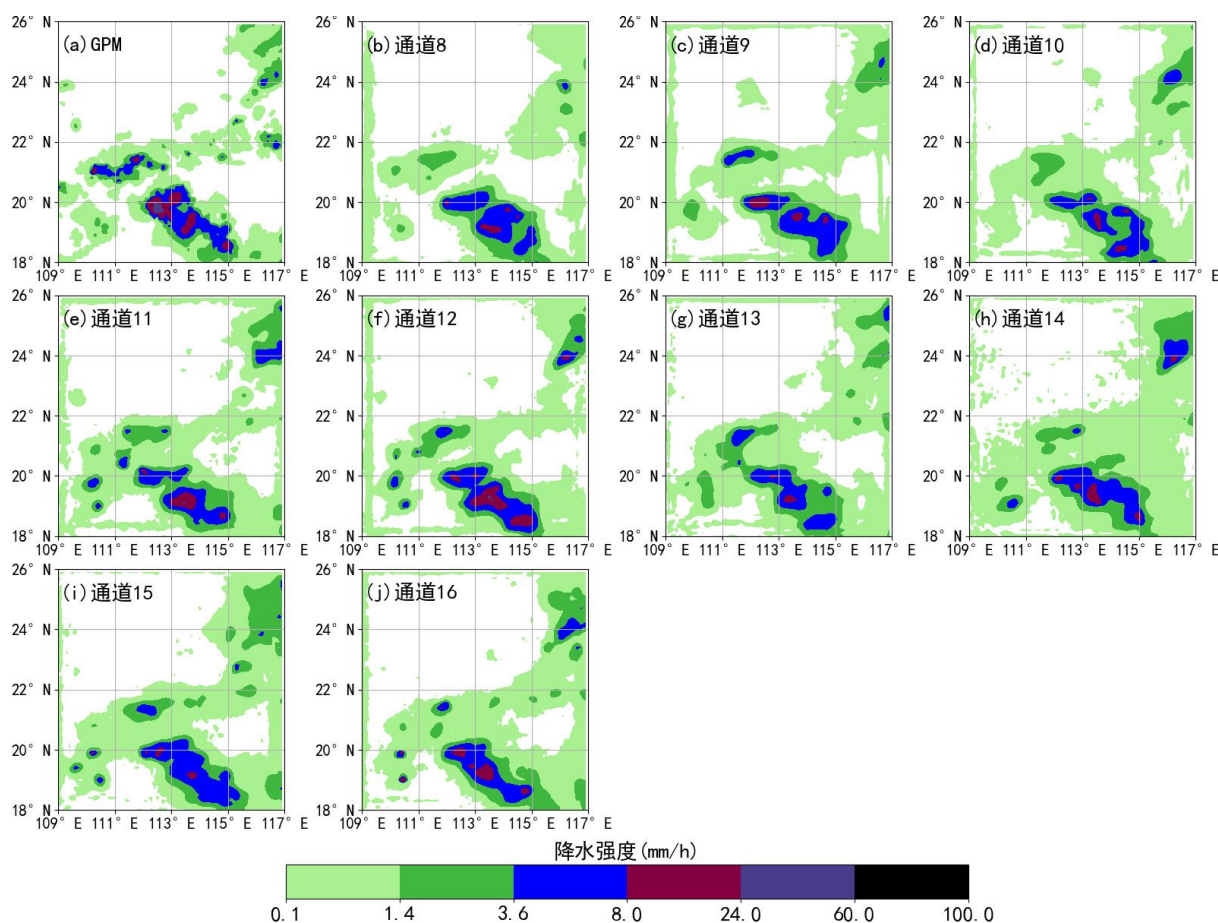


图 3-4 2019 年 8 月 30 日 16 时华南地区一次降雨过程的不同通道反演结果：(a) GPM IMERG 降水数据；(b)-(j) 通道 8~通道 16 各波段反演结果。

图 3-4 展示的是 2019 年 8 月 30 日 16 时各通道反演和 GPM 的降水量，其中图 3-4 (a) 为 GPM IMERG 的降水数据结果。此次案例的降水区域为带状，斜分布在研究区域南方，其他区域存在零星分布的小雨，是一次局部强降水事件。对于研究区域的暴雨中心，波不同通道反演结果的区域大小、位置而言，与 GPM 数据均有比较好对应关系。而对于降水强度而言，各反演结果均未体现该强度，均偏低（暴雨区→大雨区）；尤其通道 15，它对于整体南部降水结构未能较好呈现。除此降水中心之外，对于周边小雨的反演，通道 11、12、13 误报相对较小，而其他通道在其他区域误报较多，这可能是由于红外亮温反演降水强度的对应关系是红外亮温越低，降水强度越高。

根据单通道的反演结果，葵花 8 各个单通道对大量降水案例反演效果存在明显的差异，从统计指标分析上，通道 11、12、13 作为模型输入时的反演结果表现为前三，其中通道 12 最优（指标评分第一个数最多）。不过从单通道几个典型的降雨过程的分析

和大量统计评估指标中，我们发现即使作为最优的单通道（通道 12），其反演出的大雨区仍然存在一定的低估和小雨区的误报。

3.2 不同卫星观测多通道的降水反演

由 3.1 节的单通道反演的统计指标表现和案例分析，选取通道 11、通道 12、通道 13 三个指标前三的通道分别作为模型输入变量、再增加 9 通道（9 个通道）和 3 通道输入（通道 11、通道 12、通道 13）分别作为模型输入变量，多通道的降水反演可以融合各单通道侧重的信息（如水汽信息、云高信息、云相信息），是否能实现对于反演降水精度的提升。本节选取三个典型案例，分析比较多通道与单通道输入（通道 11、12、13）的反演结果，并且对单通道和多通道反演结果进行统计检验对比（表 3-1 和表 3-2），这里我们主要对比的通道 11、通道 12、通道 13 和多通道（3 通道和 9 通道）反演结果。

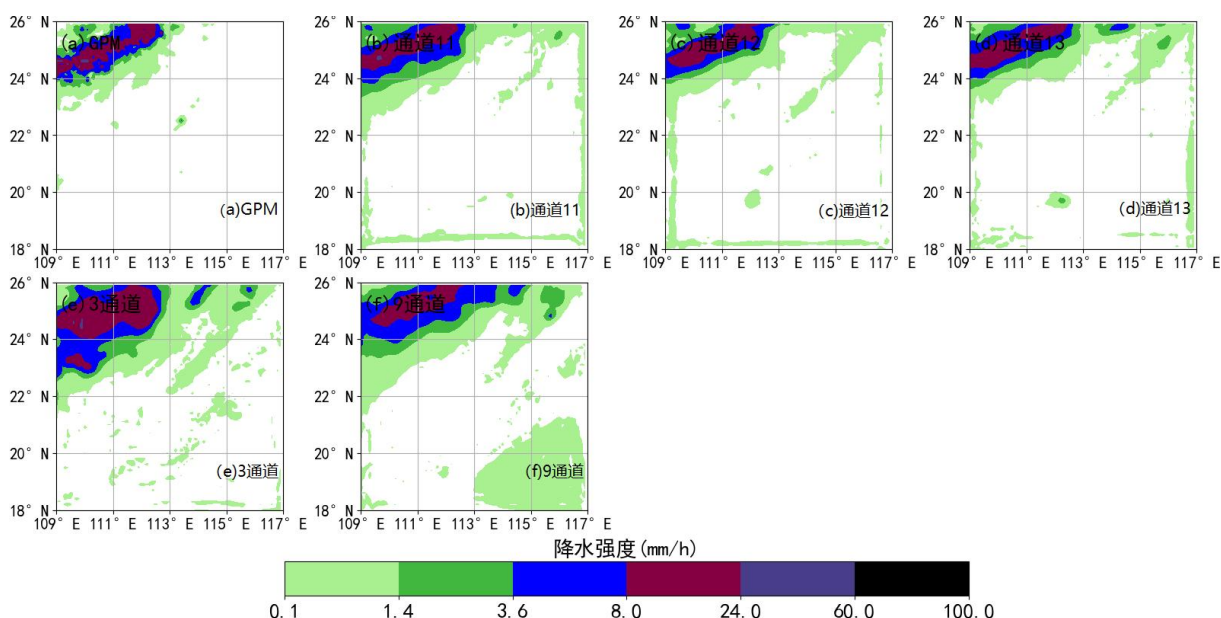


图 3-5 2019 年 7 月 12 日 21 时华南地区一次降雨过程的单通道和多通道的反演结果:(a) GPM IMERG 降水数据; (b)- (d) 单通道反演结果(通道 11、通道 12、通道 13); (e) 3 通道反演结果(8-16 多通道模型输入)。

图 3-5 展示的是 2019 年 7 月 12 日 21 时单通道和多通道反演的降水反演结果对比，其中图 3-5 (a) 表示 GPM IMERG 降水数据的真实情况，图 3-5 (b) - (e) 展示的是通道 11、12、13 的降水反演结果，图 3-5 (e) 展示的 3 通道作为模型输入的反演结果，图 3-5 (f) 展示的 9 通道作为模型输入的反演结果。GPM IMERG 降水数据显示，本次案例的降雨大值区主要位于研究区域的西北部分，降雨的中心出现较大区域的暴雨。从

单通道和多通道的反演结果来看，两者均将东南方的暴雨中心强度呈现出来。不过相对而言，通道 11、12、13 的反演结果在降雨区域上要优于 3 通道和 9 通道输入，并且单通道结果左上部分降水区域的中、小雨范围反演更为贴切；而 9 通道结果在研究区域东南方出现大片小雨区的误报，3 通道结果的暴雨区的反演降水形状与 GPM 资料的降水形状不一致，说明通道资料越多，对于反演降水结果不一定更好。

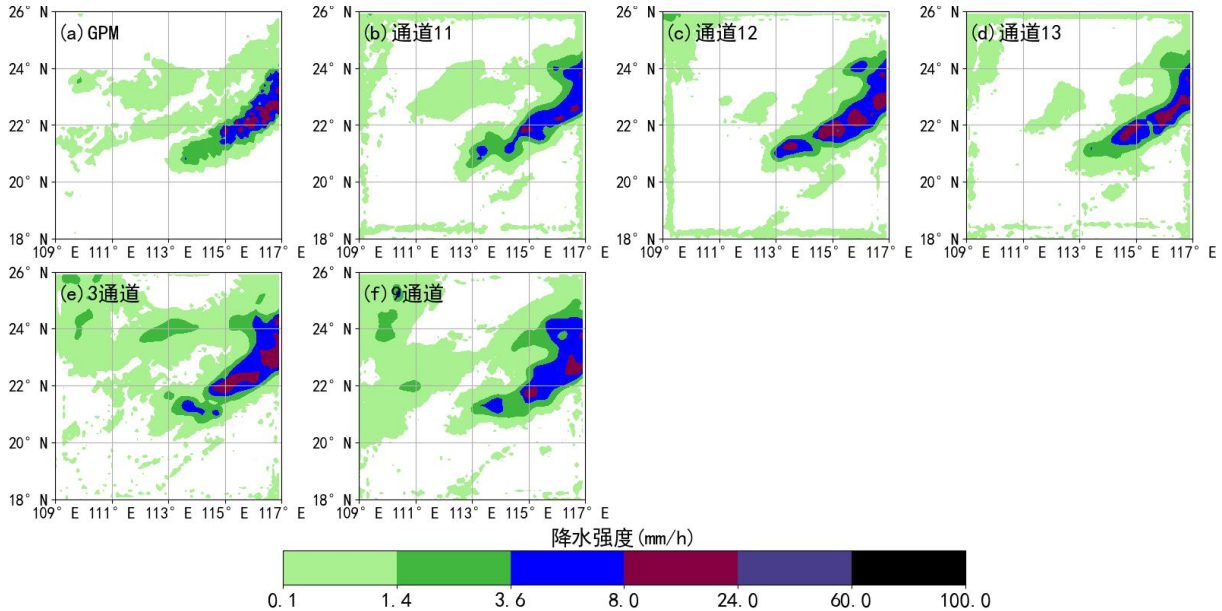


图 3-6 2022 年 7 月 10 日 22 时 30 分华南地区一次降雨过程的单通道和多通道的反演结果：(a) GPM IMERG 降水数据；(b)-(d) 单通道反演结果(通道 11、通道 12、通道 13)；(e) 3 通道反演结果(8-16 多通道模型输入)。

图 3-6 展示的是 2019 年 7 月 10 日 22 时 30 分单通道和多通道反演的降水对比，其中图 3-6 (a) 是 GPM IMERG 降水数据，图 3-6 (b) - (e) 展示的是通道 11、12、13 的降水反演结果，图 3-6 (e) 展示的 3 通道作为模型输入的反演结果，图 3-6 (f) 展示的 9 通道作为模型输入的反演结果。从判别有/无雨看，单通道（通道 11、12、13）的反演效果明显优于多通道（3 通道、9 通道）输入的反演结果，对于大雨带的分布区域和强度大小均有较好的体现；然而多通道输入的反演结果呈现情况为雨区误报、实际为小雨、反演为大雨区，对于 GPM IMERG 实际降雨对比来看，各个反演结果均对于强降水区域进行了有效的反演，不足之处在于无雨区误报为有雨区，小雨区误报为大雨区（多通道），分析原因是模型输入通道数增加，也伴随着噪声增加，导致了雨区的误报，。

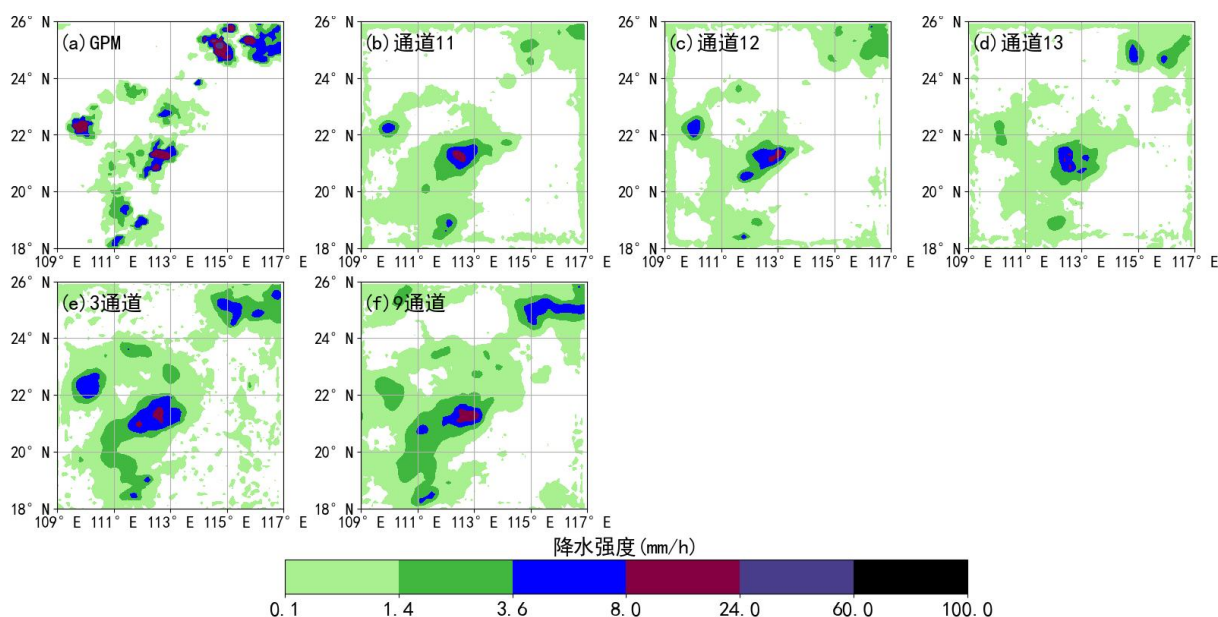


图 3-7 2019 年 5 月 31 日 13 时 30 分华南地区一次降雨过程的单通道和多通道的反演结果：(a) GPM IMERG 降水数据；(b)- (d) 单通道反演结果(通道 11、通道 12、通道 13)；(e) 3 通道反演结果(8-16 多通道模型输入)。

图 3-7 展示的是 2019 年 5 月 31 日 13 时 30 分单通道和多通道反演的降水对比，其中图 3-7 (b) - (e) 展示的是通道 11、12、13 的降水反演结果，图 3-7 (e) 展示的 3 通道作为模型输入的反演结果，图 3-7 (f) 展示的 9 通道作为模型输入的反演结果。研究区域的降水区域分散，主要对比区域东北方和中心区域的反演效果。东北方的降水区在单通道（通道 11、12、13）上没有体现，但是在 3 通道和 9 通道反演结果的大雨区对应 GPM。多通道反演降水有利有弊，增加误报的可能性，但是也大大减小了漏报率。

对比表 3-1 和表 3-2 中各个单通道和多通道的反演结果的评价指标，我们可以发现一致的结论，多通道在阈值为 1.4mm/h（小、中雨分界线）的各项指标均比单通道的反演结果更优，阈值为 0.1mm/h 的情况则相反。从案例可以看出，多通道的反演结果没有优于单通道的反演结果，还增加了降水区误报，但是对于中大雨区的反演精度没有单通道反演效果好。

综合本节中多个典型案例和大量反演实验的评估指标来看，多通道的模型反演效果总体而言不如单通道（通道 11、12、13）。并且，多通道反演的结果可以看出该方法仍然存在诸多问题，比如多通道输入并没有很好的融合多通道的信息，从而达到取长补短的效果，更多时候起到的作用是综合各个通道的主要信息，过滤掉差异化较大的信息，因此检验结果中多通道表现处于中等偏下水平。因此如果要使用多通道作为输入，那么

如何保留有用的差异化信息，以及如何滤掉无用的噪音信息，则是影响多通道模型反演结果的重要因素。但就目前试验结果而言，多通道反演结果并非最合适的模型输入。

3.3 本章小结

本章以葵花 8 卫星的 9 个红外通道及其多通道作为模型输入，进行了大量的降水反演试验。我们分析了华南地区 3 个典型降雨案例不同通道及其全通道的反演结果，利用多个常用降水指标对降水范围以及中雨以上的降水区间进行定量指标评价，重点分析和讨论了葵花 8 卫星最优通道（各单通道和多通道）。

结果表明，基于各个通道的模型反演结果存在明显差异，模型反演的降水区域整体存在高估，主要表现为小雨区的误报。通道 13 在通道实验的整体表现要优于其他单通道和多通道作为模型输入，在降水区域的结果分析以及降水强度的指标表现上均表现出色。基于多通道的反演结果整体上劣于最优单通道输入，但由于其输入信息更加全面，具备单通道难以表征的云层的结构信息，在降水区域分布及细节上明显优于单一通道。基于本章结论，本文认为葵花 8 通道 13 作为降水反演的最优模型输入。

第四章 不同卫星观测时刻对降水反演的影响

GPM IMERG 降水资料的时间分辨率为 30 分钟, 而葵花 8 资料的时间分辨率为 10 分钟, 这意味着每个 GPM IMERG 降水样本对应的时间段内存在 4 个时刻的葵花 8 数据, 本文以 t_0 、 t_{10} 、 t_{20} 、 t_{30} 表示这 4 个时刻。其中 t_0 代表 GPM IMERG 数据中 30 分钟时间中的初始时刻, t_{10} 和 t_{20} 为初始时刻后的 10 分钟和 20 分钟, t_{30} 为最终时刻(30 分钟)。鉴于这四个时刻的亮温数据均与对应的 GPM IMERG 降水资料均存在较好的对应关系(图 2.2), 本章将评估基于不同单时刻及多时刻的亮温的反演结果, 并选择模型最优时刻。本章试验采用与第三章相同的模型及模型超参数, 并选择通道 12/13 作为模型输入的亮温通道, 防止出现单通道的偶然性。为了使用降水指标更好的评价各个实验的反演降水性能, 本文利用三个阈值 0.1mm/h(有雨), 1.4mm/h(中雨), 3.6mm/h(大雨)的降水指标值的平均值评估模型性能。本章对各通道反演结果均在测试集上的表现进行统计分析。

4.1 不同卫星观测单时刻的降水反演

表 4-1 单时刻和多时刻(通道 12)输入的降水反演的指标均值

	t_0	t_{10}	t_{20}	t_{30}	$t_0\sim t_{30}$
POD	0.63	0.63	0.63	0.62	0.65
CSI	0.43	0.42	0.43	0.42	0.42
HSS	0.47	0.46	0.47	0.46	0.46
ETS	0.33	0.32	0.33	0.32	0.33
BIAS	1.11	1.13	1.12	1.1	1.23

表 4-2 单时刻和多时刻（通道 13）输入的降水反演的指标均值

	t0	t10	t20	t30	t0~t30
POD	0.64	0.63	0.63	0.62	0.66
CSI	0.43	0.42	0.43	0.42	0.42
HSS	0.48	0.47	0.47	0.46	0.46
ETS	0.33	0.32	0.33	0.32	0.33
BIAS	1.1	1.13	1.12	1.09	1.23

表 4-1、表 4-2 表示的葵花 8 卫星数据单时刻单通道反演的降水指标评价结果。从指标结果来看，各个时刻的反演降水效果差异并不大，基于 t0 时刻反演的降水场的命中率、关键成功指数、HSS 和公平性评分均为最优值。这可能是因为本研究反演的是累计降水，相比于表达云的亮温信息存在一定的滞后性。表 4-1 也展示了将 t0 至 t30 的四个时刻同时作为模型输入的反演降水结果，在大部分指标上均不如基于 t0 输入的降水反演结果。从目前的结果来看，最优输入反演结果的偏差（BIAS）指标普遍偏高，这是因为模型总体反演结果还是存在一定的高估。

本节针对葵花 8 卫星通道 12/13 进行单通道的多个单时刻反演试验，并在所有的模型反演结果中，选取 3 个典型个例（通道 13），逐一分析各个时刻对反演结果的影响。

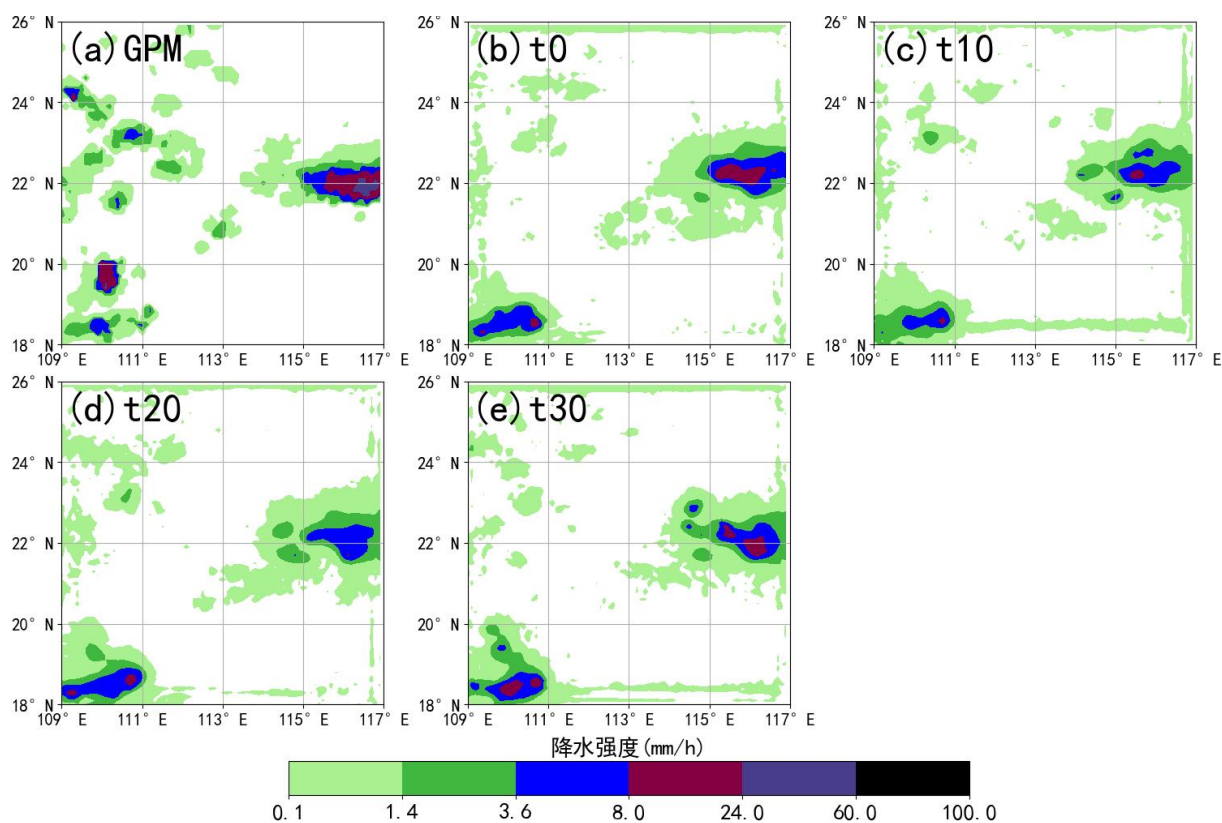


图 4-1 2019 年 8 月 16 日 4 时华南地区一次降雨过程不同时刻(通道 13)的反演结果:(a) GPM IMERG 降雨资料; (b)-(c)不同时刻的降雨资料,其中 t_0 代表初始时刻数据, t_n 代表初始时刻后 n 分钟的数据。

图 4-1 展示了 2019 年 5 月 6 日 5 时 30 分基于不同时刻输入的反演结果。图 4-1 (a) 表示 GPM IMERG 降水数据的真实情况,该案例降水中心位于本研究区域的正东方,此外在研究区域中心分布着两块暴雨区,且小雨区覆盖范围较小。图 4-1 (b) - (e) 分别代表不同时刻亮温输入的模型反演结果,对于降水中心的位置反演都是在研究区域的正东方向,但是降水中心的范围是反演强度是不一致的,从图中看出对于降水中心的反演结果上,每个时刻的反演结果上都是很好的反映。从时间的不断推移上,可以发现雨区的误报率是越来越大的,在中、大雨区的反演结果上,也可以看出基于 t_0 时刻输入的反演结果在研究区域正东区域有很好的反演效果,但其是在研究区域的西南方向的高强度降水区域反演效果并不理想。基于其他时刻的反演的结果也存在相同的问题,且整体效果不如 t_0 。这说明选择合适的输入时刻的必要性,在 10 分钟的时间间隔内,亮温与降水之间的对应关系已经发生了变化。此外,除了 t_0 时刻的反演结果可以有效地区分降水区域和非降水区域之外,其他单一时刻的反演结果并不能明显地区分降水区域,只能区分高强度降水的降水区域。随着时间的推移,中、大雨区的反演结果会消失,原因在于

降水和亮温之间对应关系随着时间变化。

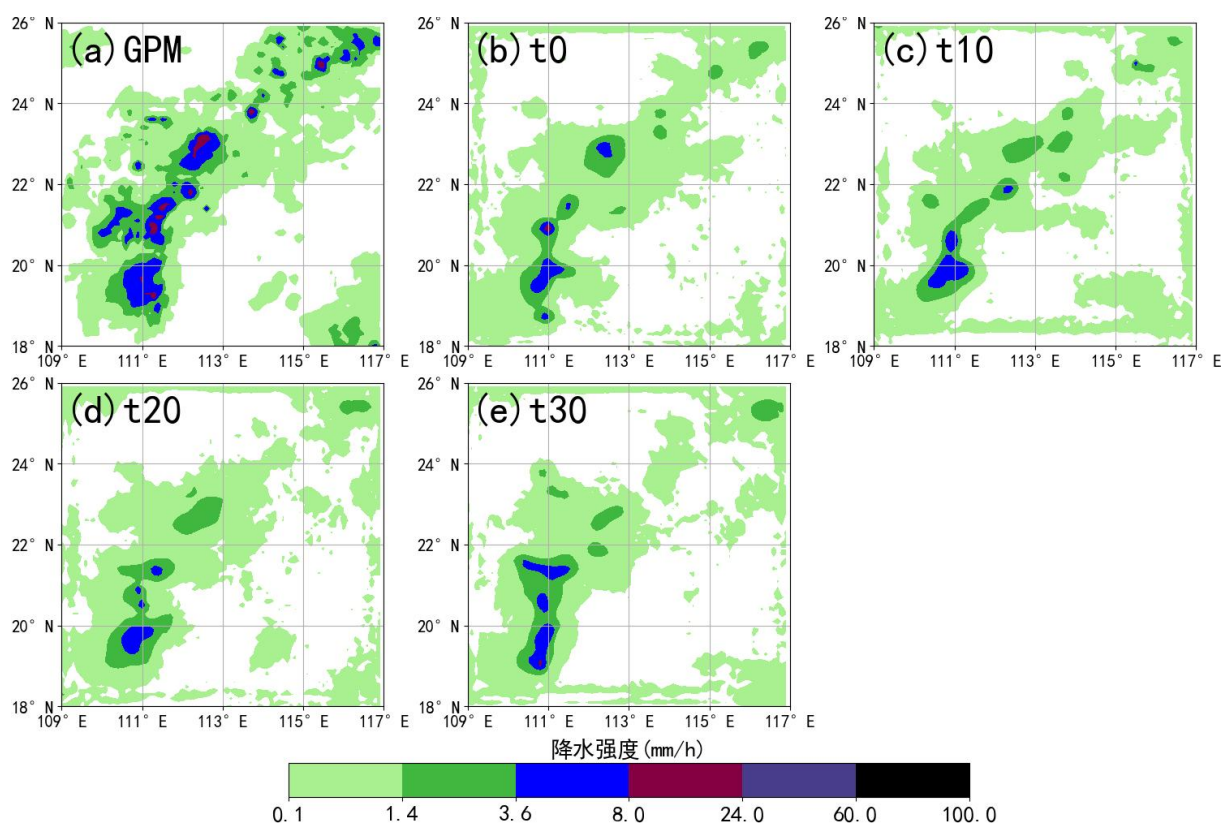


图 4-2 2019 年 5 月 31 日 10 时华南地区一次降雨过程不同时刻(通道 13)的反演结果:(a) GPM IMERG 降雨资料; (b)-(c) 不同时刻的降雨资料, 其中 t_0 代表初始时刻数据, t_n 代表初始时刻后 n 分钟的数据。

图 4-2 展示的是 2019 年 5 月 31 日 10 时基于不同时刻输入的反演结果, 其中图 4-2 (a) 是 GPM IMERG 降水数据。该时刻的大雨区分布在研究区域的中心偏左, 降水区比较分散, 连贯性较差。图 4-2 (b) - (e) 分别代表基于 t_0 至 t_{30} 各时刻亮温输入的模型反演结果, 可以看到随着时间降水位置的判断越来越差, 但强度上存在一定误差。此外, t_0 时刻的模型反演效果明显优于其他单时刻的反演效果, 该时刻的反演结果存在一定的漏报, 但明显优于其他时刻。此外, 基于 t_0 时刻亮温输入的反演结果中强降水区域以及强降水强度均较为准确。基于 t_{10} 时刻亮温的反演结果明显偏低且范围更小, 且与基于 t_{20} 、 t_{30} 时刻的亮温反演结果一样, GPM 存在的降水区域在其反演结果上完全消失。

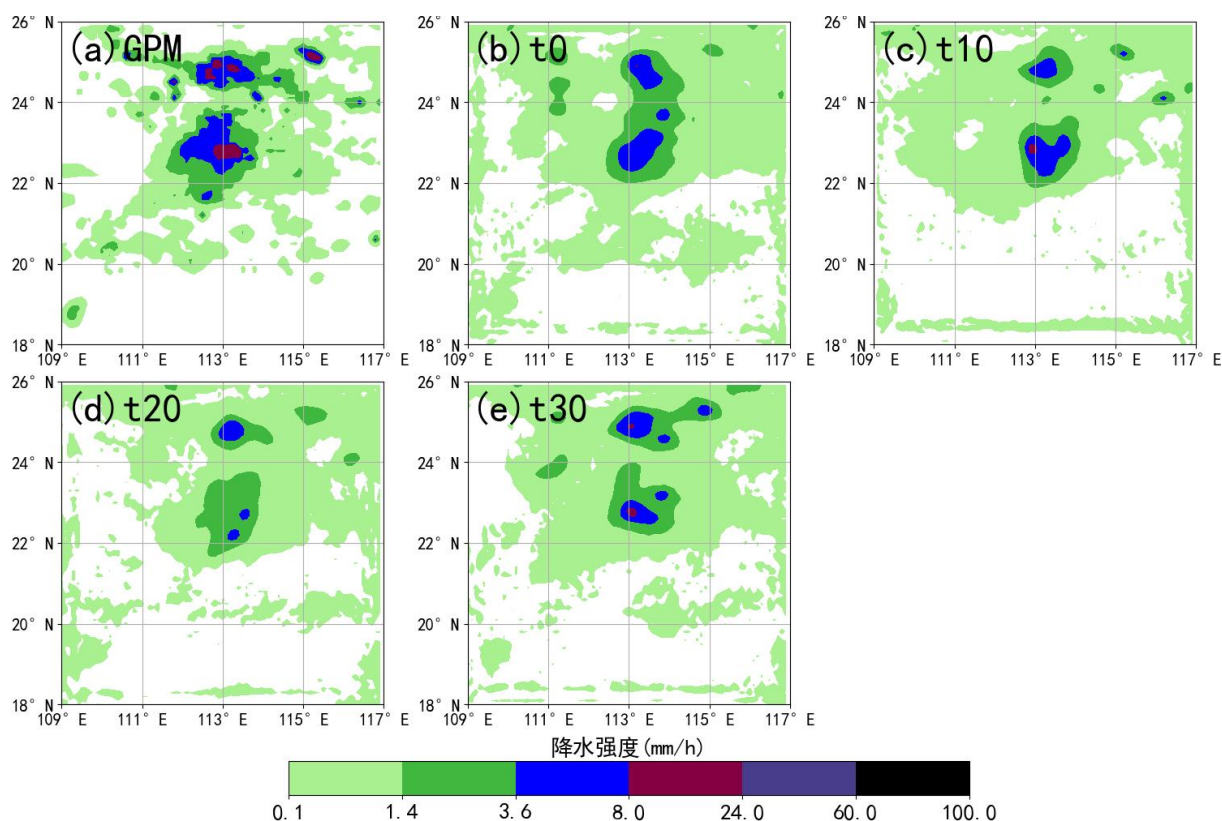


图 4-3 2019 年 6 月 11 日 13 时 30 分华南地区一次降雨过程不同时刻(通道 13)的反演结果:(a) GPM IMERG 降雨资料; (b)-(c) 不同时刻的降雨资料, 其中 t_0 代表初始时刻数据, t_n 代表初始时刻后 n 分钟的数据。

图 4-3 展示的是 2019 年 6 月 11 日 13 时 30 分一次降雨过程基于不同时刻输入的反演结果。图 4-3 (a) 为 GPM IMERG 的半小时降水强度, 该案例降水区域主要在研究区域的北方, 分为两块降水区。图 4-3 (b) - (e) 分别代表 t_0 至 t_{30} 的模型反演结果, 各反演结果的降雨中心与 GPM IMERG 数据有着很好的对应, 较为准确。各反演结果随着时间偏移, 与 GPM 的降水区域对应性越来越差。本案例中基于 t_0 时刻亮温输入的反演结果效果较好, 而基于其他时刻输入的反演结果较差, 与前两个案例的结果一致。

4.2 不同卫星观测多时刻的降水反演

将多时刻的亮温信息同时进行输入理论上可以表征红外亮温的时空变化特征, 本节将对比多时刻输入与最优单时刻输入 (t_0) 的反演结果。为了不引入更多的模型结构, 本研究通过增加卷积通道的方式引入多时刻信息, 即将输入形状调整至为 $200 \times 160 \times 4$ 。

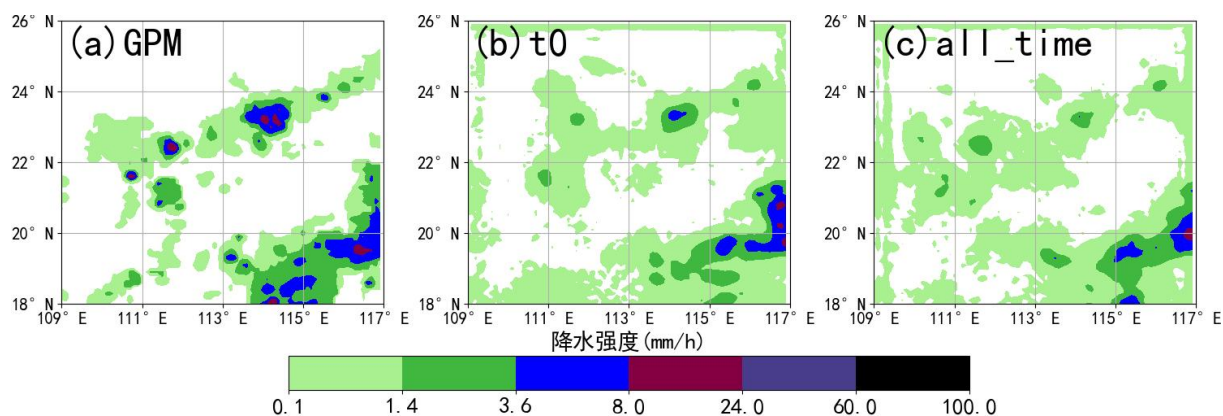


图 4-4 2019 年 5 月 28 日 16 时华南地区一次降雨过程单时刻和多时刻(通道 13)反演的比较:(a) GPM IMERG 降雨资料; (b) t0 (最优单时刻); (c) all_time (t0 至 t30 作为模型输入)。

图 4-4 展示了 2019 年 5 月 28 日 16 时的华南地区一次降雨过程单时刻和多时刻反演的比较。图 4-4 (a) 表示 GPM IMERG 降水数据的真实情况,可以看出,降水中心位于本研究区域的东南方,降雨强度由中心区域向外逐渐降低,在研究区域中心分布着两块大雨区。在北纬 23°、东经 114° 的分布一块大雨区。图 4-4 (b) 表示 t0 时刻的降水反演结果,图 4-4 (c) 表示多时刻作为模型输入的降水反演结果。基于 t0 时刻输入的反演结果很好地对应了 GPM 中的短时大雨区的位置,但是反演的降水强度不足,和多时刻反演降水都存在雨区误报的问题。

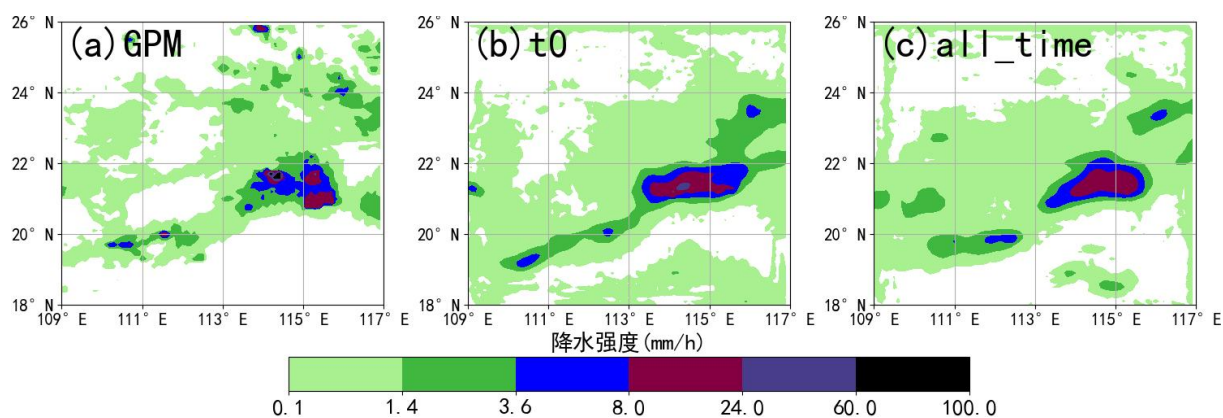


图 4-5 2019 年 8 月 17 日 11 时华南地区一次降雨过程单时刻和多时刻反演的比较:(a) GPM IMERG 降雨资料; (b) t0 (最优单时刻); (c) all_time (t0 至 t30 作为模型输入)。

图 4-5 展示了 2019 年 8 月 17 日 11 时华南地区一次降雨过程单时刻和多时刻反演的比较。此案例不论是单时刻反演降水,还是多时刻反演降水,在研究区域上的降水区域都有很好的对应。在东经 115° 北纬 23° 的强降水区域的轮廓反演模糊,暴雨区也存在高估,在这块的降水强度都是大雨强度以上,所以对于大雨区强度的反演比较好。多时

刻和单时刻的反演结果在暴雨区域还是存在偏差，但都涵盖了 GPM 的强降水区域，多时刻反演在研究区域南方没有出现误报。

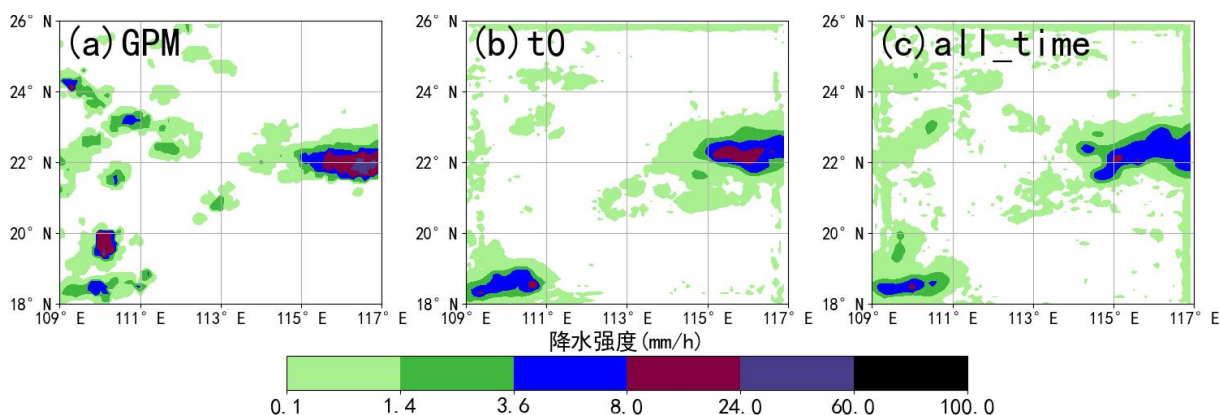


图 4-6 2019 年 8 月 16 日 4 时华南地区一次降雨过程单时刻和多时刻反演的比较：(a) GPM IMERG 降雨资料；(b) t_0 （最优单时刻）；(c) all_time（ t_0 至 t_{30} 作为模型输入）。

图 4-6 展示了 2019 年 8 月 16 日 4 时的华南地区一次降雨过程单时刻和多时刻反演的比较。图 (a) 显示为 GPM IMERG 的半小时降水强度，模型反演结果的中、大雨区基本对应。基于多时刻输入的反演结果具有较为准确的降水中心，但是强度明显偏弱，不如基于最优单时刻输入的反演结果，但是其中大雨区的位置要亚于单时刻反演的中大雨区的位置。

由表 4-1 可以看出，基于多时刻输入的反演不如基于 t_0 时刻的反演结果，在相对由于其他单时刻的反演结果。这说明模型基于多时刻的输入能够融合了不同单时刻的信息，但并没有充分利用各个时刻的有用信息实现取长补短的效果，这可能是因为模型自身辨别干扰信息的能力不足。总体而言，葵花 8 卫星红外亮温数据的时刻反演从案例和统计数据结果看， t_0 时刻的反演效果优于其他单时刻。

4.3 本章小结

本章节讨论了基于葵花 8 不同时刻亮温输入的降水反演的效果并进行评估。实验分为单时刻反演降水 and 多时刻反演降水，时刻的选取对应 GPM IMERG 数据对应的起始时间段。通过多个案例分析，发现基于起始时刻 (t_0) 的反演结果在降水强度以及降水范围上优于多时刻输入的反演结果以及基于其他单一时刻的反演降水结果。通过分析，本研究认为可能的原因是代表云体信息的亮温低值区在时间上先于降雨发生的时间。

第五章 不同反演模型对降水反演的影响及其改进

本章节主要分析和讨论不同深度学习模型对于降水过程的反演效果。主要针对三个模型结构分别进行模型反演实验并且对相应的结果进行检验，使用模型包括 U-Net、ConvMixer、pix2pixGAN。pix2pixGAN 的生成器是 U-Net 模型结构，所以本论文是事先比较 U-Net 模型和 ConvMixer 模型，两者比较模型性能，选择优者当做模型生成器。为了保证实验的一般性，将葵花 8 卫星红外资料的 9 个单通道中通道 11/12/13（单通道）作为模型输入，其对应时刻为 GPM IMERG 降水数据的起始时间（ t_0 ）。本章的测试数据集仍然选择 2019 年 5 月 1 日至 2019 年 9 月 31 日中被筛选为有效样本的部分。为了保持实验环境的一致性，三模型选用的参数均一致：设置的模型学习率（lr）：0.001；batch_size（训练批次大小）：128；训练轮次（epoch）：100；早停法耐心值：20；耐心值指标：ETS 之和（阈值 0.1/1.4/3.6），唯一不同的是 pix2pixGAN 的判别器学习率设置为 0.002，且 pix2pixGAN 的生成器损失为 L1Loss+BinaryLoss，其余模型均为 L2Loss。以下选取三个典型案例进行具体分析。本文分别以通道 11/12/13 三组不同输入，用来比较 U-Net 和 ConvMixer 的模型性能，由于三通道各个实验表现结果一致，故以通道 12 为例。

5.1 不同反演模型对降水反演的影响比较

表 5-1 不同模型的反演指标评价均值

	U-Net	ConvMixer	pix2pixGAN
POD	0.63	0.6	0.66
CSI	0.43	0.41	0.46
HSS	0.47	0.36	0.5
ETS	0.33	0.32	0.35
BIAS	1.11	1.31	1.14

综合表 5-1 中的各项评价指标，总体而言，可以发现 pix2pixGAN 模型对于研究区域的降水过程反演的效果相比其他两个模型更好，其 POD 达到了 0.66，Bias 达到了 1.14。而 ConvMixer 模型反演的指标评价最低，主要是由于 ConvMixer 是作为图像语义分割模型，和 U-Net 是同一种模型且适用于小数据集，但是 ConvMixer 的整体结果表现较差。

U-Net 模型的表现优势在于 U-Net 更好的保存了特征信息，虽然 ConvMixer 也有残差连接来保存特征信息，但是不如 U-Net 的特征保存信息完整。pix2pixGAN 的生成器是 U-Net 模型结构，整体上评价优于 U-Net 模型也是合理的，对于小雨区的误报处理上 pix2pixGAN 的判别器能够很好地甄别误报的小雨区，虽然也同时漏报了一些小雨区。pix2pixGAN 在 BIAS 指标的反应上是高估情况降低，接近真实的降雨状况，命中率 (POD)、关键成功指数 (CSI)、公平性评分 (ETS) 等都是比 U-Net、ConvMixer 模型要好。

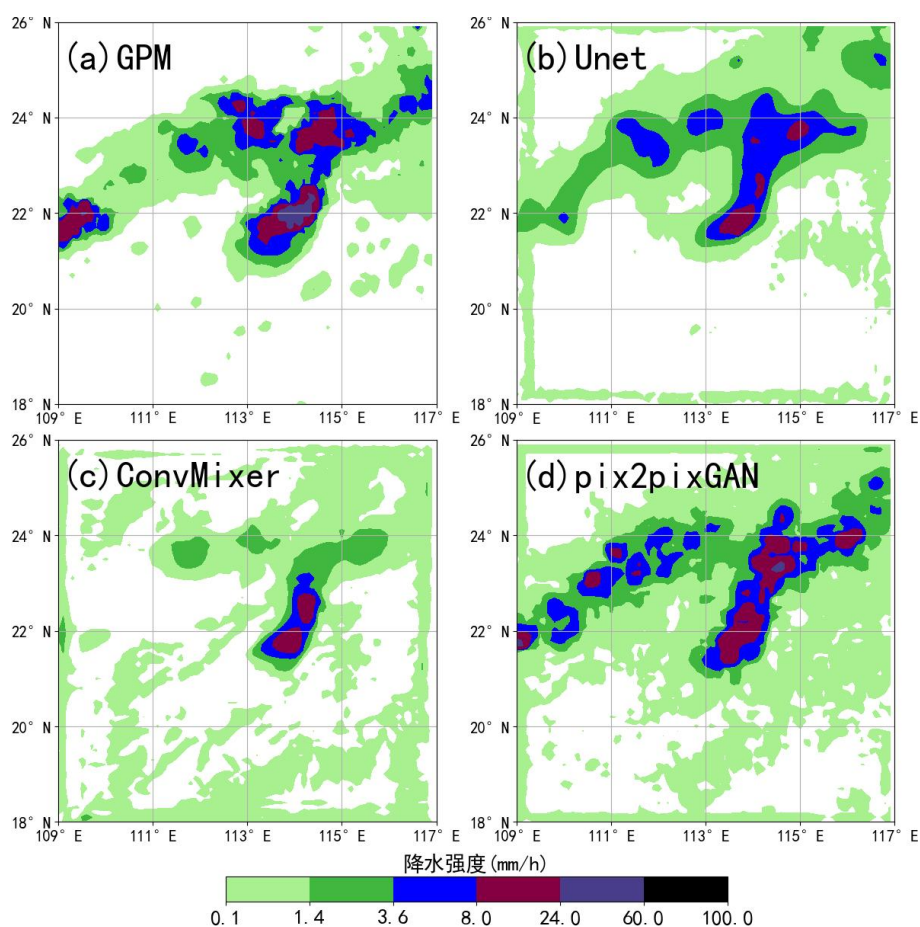


图 5-1 2019 年 6 月 13 日 13 时华南地区一次降雨过程的不同模型反演结果：(a)GPM IMERG 降雨资料；(b)-(d)三个深度学习模型的反演结果。

图 5-1 为 2019 年 6 月 13 日 13 时的华南地区一次降雨过程的三种模型反演结果以及 GPM IMERG 数据。其中图 (a) 为 GPM IMERG 的半小时的降水强度分布，图 (b) - (d) 分别表示各模型降水反演的结果。此次降水过程中，东南方出现大到暴雨，中部地区出现南北走向的带状降水区。总体而言，对于大雨区和暴雨区的位置，U-Net 和 pix2pixGAN 均能够反演达到中雨以上的降水强度。其中，pix2pixGAN 反演的降水强度

更为接近 GPM IMERG 数据，尤其对于东经 114° 附近的大雨、暴雨；对于东经 112° 的两块大雨区，pix2pixGAN 有较好的反演结果，降水强度和降水区域均相对更为合理，且降雨轮廓更契合 GPM IMERG 资料；U-Net 在这个区域将部分暴雨区低估成大雨区，且 ConvMixer 存在大面积的小雨区的误报。因此，总体情况来看 pix2pixGAN 的模型表现效果最好。

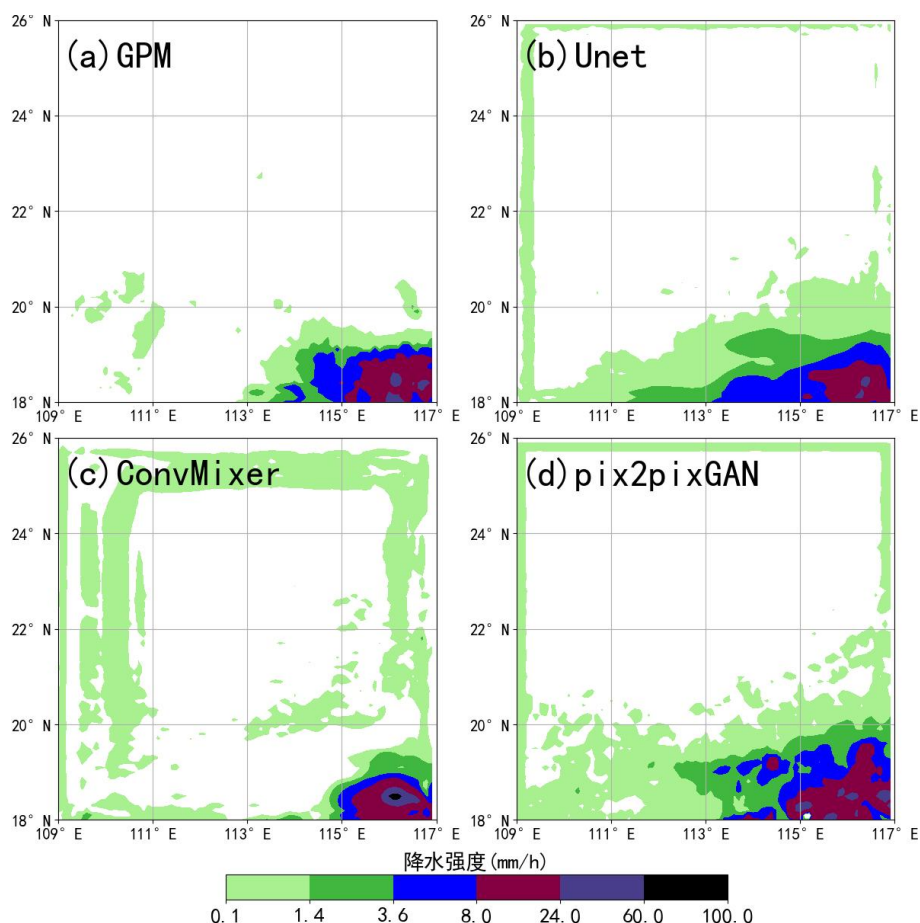


图 5-2 2019 年 6 月 14 日 12 时 30 分华南地区一次降雨过程的不同模型反演结果: (a)GPM IMERG 降雨资料; (b)-(d)三个深度学习模型的反演结果。

图 5-2 展示了 2019 年 6 月 14 日 12 时 30 分华南地区一次降雨过程的三种模型反演结果及 GPM 降水。图 5-2 (a) 为 GPM IMERG 的半小时的降水强度，而图 5-2 (b) - (d) 表示三种模型降水反演的结果。此次降水过程中，大到暴雨分布在研究区域的南部，以及中部存在狭长的东西走向的雨带。对于小雨区，ConvMixer 同样存在小雨区的误报，U-Net 相对较好但也存在一定的小雨区误报，而 pix2pixGAN 的小雨误报率相对最小。对于靠近研究区域中心的暴雨区，pix2pixGAN 的反演结果从大雨/暴雨分布范围上有较好对应，然而研究区域的降水中心在降水强度上存在高估，没有呈现出暴雨；

U-Net 和 ConvMixer 对应的降水强度与 GPM IMERG 对应更不尽人意,尤其是 ConvMixer。在反演的降水区域上, pix2pixGAN 的暴雨区反演结果更接近 GPM IMERG 数据,与上一案例一致, pix2pixGAN 模型性能更好。

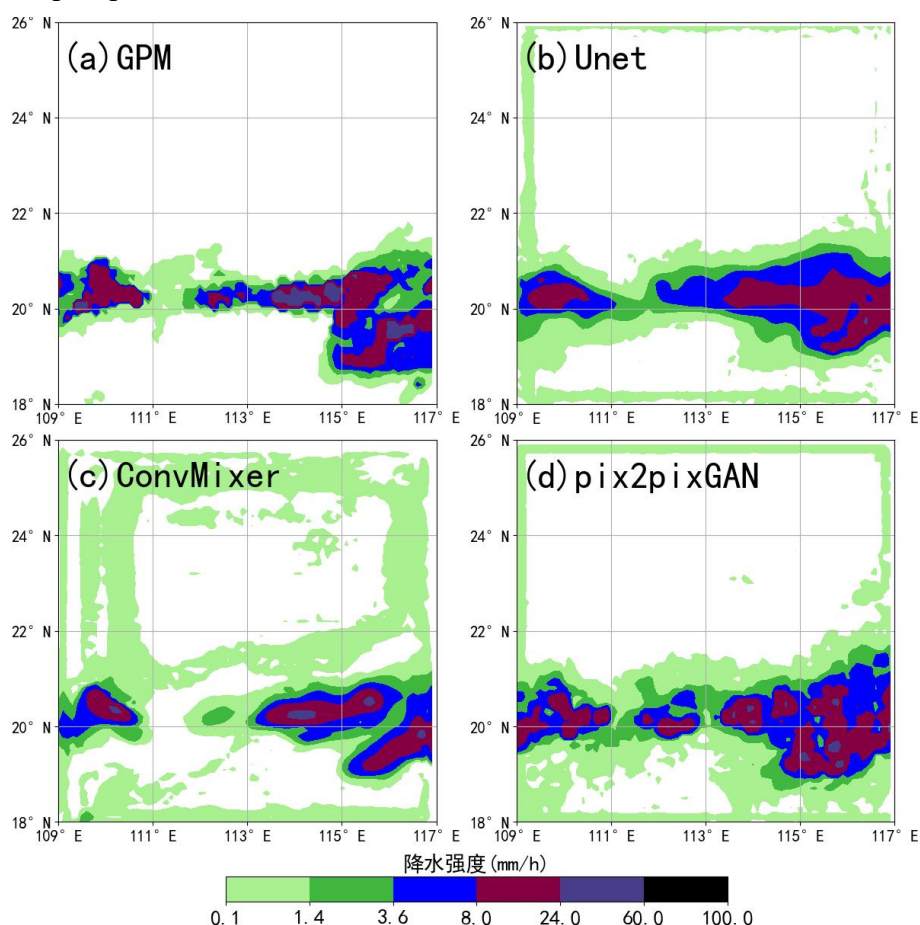


图 5-3 2019 年 6 月 14 日 23 时华南地区一次降雨过程的不同模型反演结果: (a)GPM IMERG 降雨资料; (b)-(d)三个深度学习模型的反演结果。

图 5-3 展示了 2019 年 6 月 14 日 23 时某次降水过程中三模型反演及 GPM IMERG 数据的半小时的降水强度分布。图 (a) 为 GPM IMERG 的半小时降水强度, 图 (b) - (d) 表示三种模型降水反演的结果。此次降水过程的大到暴雨区域较为集中, 主要分布在中部, 周边仅零散分布小雨区域。总体而言, GPM 数据显示的暴雨区, 在模型反演结果中都得到了较好的反演, 反演的降水强度和降水范围都符合预期。其中, U-Net 和 ConvMixer 的反演结果整体上和 GPM IMERG 的降水区域相差较多, 且雨区结构、降水强度上模型处理的效果也不到位: 特别是雨区轮廓非常粗糙, 存在严重的误报和漏报。同时, ConvMixer 还对研究区域的北部区域存在严重的误报。pix2pixGAN 的效果相对更好, 暴雨的降水强度上是三个模型中对应最好的, 然而周边的雨区结构细节未能得以

呈现，且周边的小雨区域也存在误报漏报。pix2pixGAN 对比 GPM IMERG 的降水在东经 112° 北纬 20° 反演的降雨区域在其他两个模型都没有体现出来，然而对于研究区域整体降水反演效果相对更好。

5.2 基于两步预报法的模型改进

由于多个模型的反演对于降水区域仍然存在小雨区误报和漏报的问题，本研究基于两步预报法的思想，对 5.1 节中最优模型 pix2pixGAN 进行改进、测试和检验。其中，模型输入输出和 5.1 节保持一致，对于模型的输出分两步处理。首先，该方法对降雨是否发生进行判断，并输出降雨发生判别标识；其次，对降雨的强度和分布进行同时反演；最终，在得到降雨发生判别标识和降雨的强度分布后，该方法对于降雨发生的区域和强度进行综合输出。

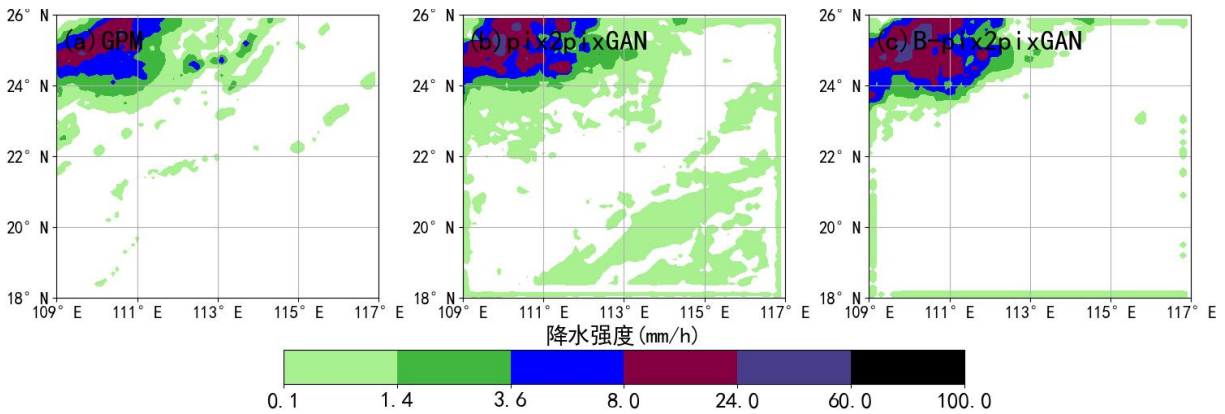


图 5-4 2019 年 7 月 12 日 16 时 30 分华南地区的一次降雨过程 pix2pixGAN 和其改进模型 B-pix2pixGAN 的反演效果比较(a) GPM IMERG 降雨资料；(b) pix2pixGAN 模型的反演结果；(c) B-pix2pixGAN 模型的反演结果。

图 5-4 展示了 2019 年 7 月 12 日 16 时 30 分的半小时内的降水强度，其中图 5-4 (a) 为 GPM IMERG 的半小时降水强度，图 5-4 (b) 是 pix2pixGAN 反演的结果，图 5-4 (c) 是基于两步预报法改进模型之后反演的结果。通过对比 GPM IMERG 降雨数据，可以发现虽然 pix2pixGAN 和 B-pix2pixGAN 均能较好的反演出降雨中心的暴雨区域，但是 pix2pixGAN 的小雨误报率很明显。但也不难看出，经过两部预报法的改进后，B-pix2pixGAN 对于小雨区的误报有着较好的抑制作用，大雨和中雨区反演的结果和 pix2pixGAN 相比更加接近 GPM IMERG 降雨数据。总体来看，B-pix2pixGAN 的降水区域的处理要优于 pix2pixGAN，整体上对降水区域的把握更准确，在部分的场景下，可

以减少了原 pix2pixGAN 模型中小雨区的误报。

不过需要注意的是，B-pix2pixGAN 模型的使用仍然存在着一定的局限性。虽然在降雨区域较小的时候，B-pix2pixGAN 可以更精确地判断小区域的降水是否发生；但是在降水区域较大的时候，该模型反演结果相比 pix2pixGAN 模型并没有明显的改善。B-pix2pixGAN 是在此基础上把区域的有无降水精细到单个格点级别，因此该方法对于小区域降水可以提供更加精确的判断。而降雨区域较大的时候，pix2pixGAN 模型的反演结果已经较为优秀，主要由于该模型使用局部判断生成器输出的真伪，所以基于两步预报法的 B-pix2pixGAN 的改进空间就较为有限。

5.3 本章小结

本章节基于 U-Net、ConvMixer、pix2pixGAN 三种深度学习模型反演降水。本研究选取的模型是常用于图像分割的研究领域的模型。这些深度学习模型本质上输入输出都是二维网格数据，在深度学习领域表现为模型输入输出均为图片。

U-Net 模型的优势在于模型是经典的编码解码结构，模型从低维度特征提取成高维度特征，再把高维度特征还原成低维度特征，从结果看出现一定的低估和误判，原因是在于降水区域在葵花 8 瞬时亮温上没有完全体现。

ConvMixer 模型的优势在于参数量相对较少，从效果来看不如 U-Net 模型。从测试指标的结果来看，U-Net 模型优于 ConvMixer 模型。

pix2pixGAN 模型分为生成器和判别器两部分，生成器的结构基于 U-Net 的结构，符合本研究的数据集，判别器提高了整体模型的反演精度，并利用 patch-D 的思想对网格数据的局部区域分别判断模型反演降水结果的正确性，使得模型反演的小区域处理的更到位。基于两步预报法的 B-pix2pixGAN 模型的模型精度比 pix2pixGAN 更加精确，针对格点的反演精度更高。

第六章 结论与展望

6.1 结论

降水在全球和区域水循环中起到关键作用。准确的降水观测或估计在气候、水文和天气预报模式中至关重要。及时准确的观测降水需要部署大量的雨量计，但雨量计的部署和维护尤其是在陆地复杂地形区域或海洋非常困难，很难大范围全域部署雨量计观测降水。因此，通过天气雷达或卫星观测间接估计降水是非常必要的，尤其随着新一代高时空分辨率静止卫星的发展以及数据驱动的深度学习方法的应用，为高时空分辨率的降水估计带来了新的机遇。

本研究基于 Himawari-8 静止卫星多通道红外亮温观测和 NASA GPM IMERG Final 降水产品，结合深度学习方法研究降水反演技术。考虑到不同通道红外亮温观测间的差异、不同时间卫星亮温观测，以及深度学习模型结构的差异，本文详细研究了不同红外通道亮温观测对降水反演的影响、不同时间卫星亮温观测对降水反演的影响以及深度学习模型结构对降水反演的影响。针对上述研究内容本文得到的主要结论如下：

- 1) 基于 Himawari-8 静止卫星 9 个红外通道亮温观测作为输入，结合编码-解码结构的 U-Net 模型分析了不同红外通道亮温观测对降水反演的影响，结果表明，当采用 Himawari-8 静止卫星通道 13 红外亮温作为模型输入反演降水时效果最优。而当采用 Himawari-8 静止卫星的 9 个红外通道亮温观测作为模型输入反演降水时比仅采用通道 13 红外亮温观测反演降水的精度要低。
- 2) 考虑到降水发展演变的复杂性，仅依赖单个时刻卫星红外亮温观测可能无法准确估计降水。因此，本文基于 Himawari-8 不同时刻的多通道红外亮温观测，结合 U-Net 模型进一步研究了不同时刻红外亮温观测对降水反演的影响，结果表明：GPM IMERG 降水资料的起始时刻对应的 Himawari-8 红外亮温观测作为模型输入要优于其他单时刻和多时刻作为模型输入的降水反演精度，具体表现在起始时刻作为模型输入在反演降水强度和降水的区域轮廓细节上最优。
- 3) 考虑到深度学习模型结构的影响，本文在编码-解码结构 U-Net 模型的基础上，选择了 pix2pixGAN(生成对抗网络模型结构)和 ConvMixer(VIT 模型结构)进一步改进降水反演效果，并对不同模型结构的表现进行了对比，结果如下：U-Net 模型、ConvMixer 模型和 pix2pixGAN 模型都能对研究区域的降水区域有较好的反演，但

是存在对无雨区误报为小雨区和对中、大雨区存在低估的问题。最后，本文构建了基于两步预报法和 pix2pixGAN 模型改进预报模型（即 B-pix2pixGAN），并加入雨强掩膜控制，有效地改进了无雨区误报为小雨区的问题，并提高了中、大雨区的降水反演精度。

6.2 创新点

本文基于多通道红外亮温观测，结合深度学习方法研究了不同红外通道亮温观测和不同时间观测对降水估计的影响。针对小雨区误报和大雨区低估的问题，发展了一个基于两步预报法和 pix2pixGAN 模型改进的预报模型（即 B-pix2pixGAN），发挥生成对抗网络高分辨率映射的优势和雨强掩膜训练控制，有效地改进了降水反演的准确率。

6.3 不足与展望

本研究基于 Himawair-8 多通道红外亮温观测资料，通过建立了双层叠连 B-pix2pixGAN 模型改进了降水反演，尤其是小雨区的误报和大雨区估计准确率。但本文仍然存在一些不足之处：

1) 本研究对于通道实验的仅研究了 Himawari-8 存在 9 个红外通道作为模型输入时对降水反演的影响。并未测试 Himawari-8 可见光通道和短波红外通道对降水反演的影响。未来应加入天气雷达和更多卫星通道观测作为输入，进一步改进降水反演精度。

2) 考虑到 NASA GPM IMERG 降水是 30 分钟平均降水量，本研究仅针对 NASA GPM IMERG 降水分辨率（30 分钟）所对应的观测时间内 Himawari-8 红外亮温观测对降水反演的影响，未进一步研究在外围相近时刻的卫星观测对降水反演的影响。未来应试验研究前后时刻卫星观测对降水反演的影响，探讨是否能进一步改进降水反演精度。

3) 本研究仅考虑了基于静止卫星亮温观测，结合深度学习方法的降水反演问题。这一技术具有支持的降水临近预报的潜力，未来，应基于深度学习方法，应用高时空分辨率的静止卫星观测资料，实现降水的临近预报。

参考文献

- [1]. Cheng H, Wu T, Dong W. Thermal contrast between the middle-latitude Asian continent and adjacent ocean and its connection to the East Asian summer precipitation[J]. *Journal of climate*, 2008, 21(19): 4992-5007.
- [2]. Yao C, Yang S, Qian W, et al. Regional summer precipitation events in Asia and their changes in the past decades[J]. *Journal of Geophysical Research: Atmospheres*, 2008, 113(D17).
- [3]. 王祝. 广东省降水的趋势变化和时空分布特性分析[J]. *人民珠江*, 2006, 27(02): 37-39+50 .
- [4]. 刘永林, 延军平, 岑敏仪. 广东省降水非均匀性与气候变化的响应关系[J]. *中山大学学报: 自然科学版*, 2015, 54(5): 138-146.
- [5]. 廖一帆, 林炳章, 丁辉. 广东省暴雨高风险区划[J]. *水资源保护*, 2022, 38(2): 7-16.
- [6]. Anagnostou E N, Maggioni V, Nikolopoulos E I, et al. Benchmarking high-resolution global satellite rainfall products to radar and rain-gauge rainfall estimates[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2009, 48(4): 1667-1683.
- [7]. Kidd C, Levizzani V. Status of satellite precipitation retrievals[J]. *Hydrology and Earth System Sciences*, 2011, 15(4): 1109-1116.
- [8]. Stampoulis D, Anagnostou E N. Evaluation of global satellite rainfall products over continental Europe[J]. *Journal of Hydrometeorology*, 2012, 13(2): 588-603.
- [9]. Adler R F, Mack R A. Thunderstorm cloud height–rainfall rate relations for use with satellite rainfall estimation techniques[J]. *Journal of climate and applied meteorology*, 1984: 280-296.
- [10]. Wardah T, Bakar S H A, Bardossy A, et al. Use of geostationary meteorological satellite images in convective rain estimation for flash-flood forecasting[J]. *Journal of Hydrology*, 2008, 356(3-4): 283-298.
- [11]. 唐国强, 万玮, 曾子悦, 等. 全球降水测量 (GPM) 计划及其最新进展综述[J]. *遥感技术与应用*, 2015, 30(4): 607-615.

- [12].Min M, Bai C, Guo J, et al. Estimating summertime precipitation from Himawari-8 and global forecast system based on machine learning[J]. IEEE transactions on geoscience and remote sensing, 2018, 57(5): 2557-2570.
- [13].薛鹏飞, 余钟波, 谷黄河. 雅鲁藏布江流域 GPM 和 TRMM 遥感降水产品精度评估[J]. 水电能源科学, 2020, 38(11):13-16.
- [14].张天宇, 桂术, 杨若文, 等. TRMM 和 CMORPH 卫星资料对三峡库区降水的评估分析[J]. 气象, 2020, 46(8): 1098-1112.
- [15].杨娜, 卢莹, 杨茜然,等. PERSIANN-CDR 产品对淮河流域干旱特征辨识性能评估[J]. 水电能源科学, 2020, 38(3): 1-4+12.
- [16].Skofronick-Jackson G, Petersen W A, Berg W, et al. The Global Precipitation Measurement (GPM) mission for science and society[J]. Bulletin of the American Meteorological Society, 2017, 98(8): 1679-1695.
- [17].任英杰, 雍斌, 鹿德凯, et al. 全球降水计划多卫星降水联合反演 IMERG 卫星降水产品在中国大陆地区的多尺度精度评估[J]. 湖泊科学, 2019(2):13.
- [18].孔宇. 中国大陆 GPM/IMERG 产品的精度评估[D]. 南京信息工程大学, 2017.
- [19].孙桂凯, 魏义熊, 王国帅,等. GPM IMERG 卫星降水数据在广西地区适用性分析[J]. 人民长江, 2022(007):053.
- [20].Wang Z, Zhong R, Lai C, et al. Evaluation of the GPM IMERG satellite-based precipitation products and the hydrological utility[J]. Atmospheric Research, 2017, 196: 151-163.
- [21].Foelsche U, Kirchengast G, Fuchsberger J, et al. Evaluation of GPM IMERG Early, Late, and Final rainfall estimates using WegenerNet gauge data in southeastern Austria[J]. Hydrology and Earth System Sciences, 2017, 21(12): 6559-6572.
- [22].Wang G, Wang K, Han W, et al. Typhoon Maria precipitation retrieval and evolution based on the infrared brightness temperature of the Feng-Yun 4A/advanced geosynchronous radiation imager[J]. Advances in Meteorology, 2020, 2020: 1-12.
- [23].Arkin P A. The relationship between fractional coverage of high cloud and rainfall accumulations during GATE over the B-scale array[J]. Monthly weather review, 1979, 107(10): 1382-1387.

- [24]. Arkin P A, Joyce R, Janowiak J E. The estimation of global monthly mean rainfall using infrared satellite data: The GOES Precipitation Index (GPI)[J]. *Remote Sensing Reviews*, 1994, 11(1-4): 107-124.
- [25]. Mishra A K, Gairola R M, Varma A K, et al. Improved rainfall estimation over the Indian region using satellite infrared technique[J]. *Advances in space research*, 2011, 48(1): 49-55.
- [26]. Todd M C, Kidd C, Kniveton D, et al. A combined satellite infrared and passive microwave technique for estimation of small-scale rainfall[J]. *Journal of Atmospheric and Oceanic Technology*, 2001, 18(5): 742-755.
- [27]. Adler R F, Negri A J. A satellite infrared technique to estimate tropical convective and stratiform rainfall[J]. *Journal of Applied Meteorology and Climatology*, 1988, 27(1): 30-51.
- [28]. Petersen W A, Rutledge S A. On the relationship between cloud-to-ground lightning and convective rainfall[J]. *Journal of Geophysical Research: Atmospheres*, 1998, 103(D12): 14025-14040.
- [29]. Grecu M, Anagnostou E N, Adler R F. Assessment of the use of lightning information in satellite infrared rainfall estimation[J]. *Journal of Hydrometeorology*, 2000, 1(3): 211-221.
- [30]. Reudenbach C, Nauss T, Bendix J. Retrieving Precipitation with GOES, Meteosat, and Terra/MSG at the Tropics and Mid-latitudes[J]. *Measuring Precipitation From Space: EURAINSAT and the Future*, 2007: 509-519.
- [31]. Ebert E E, Janowiak J E, Kidd C. Comparison of near-real-time precipitation estimates from satellite observations and numerical models[J]. *Bulletin of the American Meteorological Society*, 2007, 88(1): 47-64.
- [32]. Früh B, Bendix J, Nauss T, et al. Verification of precipitation from regional climate simulations and remote-sensing observations with respect to ground-based observations in the upper Danube catchment[J]. *Meteorologische Zeitschrift*, 2007, 16(3): 275-293.
- [33]. Ba M B, Gruber A. GOES multispectral rainfall algorithm (GMSRA)[J]. *Journal of Applied Meteorology and Climatology*, 2001, 40(8): 1500-1514.

- [34].Nauss T, Kokhanovsky A A. Discriminating raining from non-raining clouds at mid-latitudes using multispectral satellite data[J]. Atmospheric Chemistry and Physics, 2006, 6(12): 5031-5036.
- [35].Ebert E E, Manton M J. Performance of satellite rainfall estimation algorithms during TOGA COARE[J]. Journal of the Atmospheric Sciences, 1998, 55(9): 1537-1557.
- [36].Torricella F, Cattani E, Levizzani V. Rain area delineation by means of multispectral cloud characterization from satellite[J]. Advances in Geosciences, 2008, 17: 43-47.
- [37].Cattani E, Torricella F, Laviola S, et al. On the statistical relationship between cloud optical and microphysical characteristics and rainfall intensity for convective storms over the Mediterranean[J]. Natural Hazards and Earth System Sciences, 2009, 9(6): 2135-2142.
- [38].Thies B, Nauß T, Bendix J. Precipitation process and rainfall intensity differentiation using Meteosat second generation spinning enhanced visible and infrared imager data[J]. Journal of Geophysical Research: Atmospheres, 2008, 113(D23).
- [39].Feidas H, Giannakos A. Classifying convective and stratiform rain using multispectral infrared Meteosat Second Generation satellite data[J]. Theoretical and applied climatology, 2012, 108: 613-630.
- [40].Lazri M, Ameur Z, Ameur S, et al. Rainfall estimation over a Mediterranean region using a method based on various spectral parameters of SEVIRI-MSG[J]. Advances in space research, 2013, 52(8): 1450-1466.
- [41].Bellerby T, Todd M, Kniveton D, et al. Rainfall estimation from a combination of TRMM precipitation radar and GOES multispectral satellite imagery through the use of an artificial neural network[J]. Journal of applied Meteorology, 2000, 39(12): 2115-2128.
- [42].师春香, 卢乃锰, 张文建. 卫星面降水估计人工神经网络方法[J]. 气候与环境研究, 2001, 6(4): 467-472.
- [43].熊秋芬, 胡江林, 夏军. 神经网络方法在静止卫星多通道资料估算降水中的应用[J]. 气象, 2002, 28(9): 17-21.

- [44].Hong Y, Hsu K L, Sorooshian S, et al. Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system[J]. *Journal of Applied Meteorology*, 2004, 43(12): 1834-1853.
- [45].Hamidi O, Poorolajal J, Sadeghifar M, et al. A comparative study of support vector machines and artificial neural networks for predicting precipitation in Iran[J]. *Theoretical and applied climatology*, 2015, 119: 723-731.
- [46].Ma L, Zhang G, Lu E. Using the gradient boosting decision tree to improve the delineation of hourly rain areas during the summer from advanced Himawari imager data[J]. *Journal of Hydrometeorology*, 2018, 19(5): 761-776.
- [47].Kühnlein M, Appelhans T, Thies B, et al. Precipitation estimates from MSG SEVIRI daytime, nighttime, and twilight data with random forests[J]. *Journal of Applied Meteorology and Climatology*, 2014, 53(11): 2457-2480.
- [48].Das S, Chakraborty R, Maitra A. A random forest algorithm for nowcasting of intense precipitation events[J]. *Advances in Space Research*, 2017, 60(6): 1271-1282.
- [49].Turini N, Thies B, Bendix J. Estimating high spatio-temporal resolution rainfall from MSG1 and GPM IMERG based on machine learning: Case study of Iran[J]. *Remote Sensing*, 2019, 11(19): 2307.
- [50].Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks[J]. *science*, 2006, 313(5786): 504-507.
- [51].Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[J]. *Advances in neural information processing systems*, 2012, 60, 84-90.
- [52].Wang C, Xu J, Tang G, et al. Infrared precipitation estimation using convolutional neural network[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2020, 58(12): 8612-8625.
- [53].Wang Y, Zhang J, Zhu H, et al. Memory in Memory: A Predictive Neural Network for Learning Higher-Order Non-Stationarity From Spatiotemporal Dynamics[C]// 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2019.

- [54].Venkatesh R, Balasubramanian C, Kaliappan M. Rainfall prediction using generative adversarial networks with convolution neural network[J]. Soft Computing, 2021, 25: 4725-4738.
- [55].Sadeghi M, Nguyen P, Hsu K, et al. Improving near real-time precipitation estimation using a U-Net convolutional neural network and geographical information[J]. Environmental Modelling & Software, 2020, 134: 104856.
- [56].Trebing K, Stańczyk T, Mehrkanoon S. SmaAt-U-Net: Precipitation nowcasting using a small attention-U-Net architecture[J]. Pattern Recognition Letters, 2021, 145: 178-186.