

基于神经网络的太赫兹冰云探测反演算法研究

陈柯¹, 张兰¹, 张幼明², 董彬彬¹, 刘艳¹, 吴琼³, 商建³

1. 华中科技大学 电子信息与通信学院, 武汉 430074;

2. 武汉船舶通信研究所, 武汉 430079;

3. 许健民气象卫星创新中心, 中国气象局中国遥感卫星辐射测量和定标重点开放实验室/国家卫星气象中心
(国家空间天气监测预警中心), 北京 100081

摘要: 太赫兹频段在冰云探测上具有独特优势, 但是目前的太赫兹冰云反演算法将不同种类冰相粒子(主要是冰和霰)视为冰粒子统一计算。本文根据冰云太赫兹辐射特性实现了一种预分类的神经网络算法, 能够从太赫兹亮温中分别反演得到冰、霰两种粒子的统计参数和廓线分布。首先, 基于WRF数值模式和ATMS载荷真实观测的冰云霰数据构建了包含冰、霰粒子密度廓线的混合冰云数据库, 然后, 使用DOTLRT辐射传输模式模拟183 GHz、243 GHz、325 GHz、448 GHz、664 GHz和874 GHz这6个频段的星载太赫兹冰云探测亮温, 最后, 开展冰云参数探测仿真试验, 验证反演算法性能。仿真试验中反演得到的冰和霰的路径总量均方根误差分别为 8.97 g/m^2 和 10.90 g/m^2 , 等效粒径均方根误差分别为 $7.54 \mu\text{m}$ 和 $25.38 \mu\text{m}$, 反演的冰、霰密度廓线也具有较高的精度。研究结果表明本文算法能够以较好的精度从多频太赫兹冰云探测亮温数据中分别反演得到冰、霰两种粒子的路径总量、等效粒径、等效云高和密度廓线, 突破现有研究仅仅计算单一冰粒子的局限, 更加符合冰云真实情况。

关键词: 太赫兹, 冰云探测, 神经网络, 冰和霰粒子, 冰云参数反演

引用格式: 陈柯, 张兰, 张幼明, 董彬彬, 刘艳, 吴琼, 商建. 2022. 基于神经网络的太赫兹冰云探测反演算法研究. 遥感学报, 26(10): 2043-2059

Chen K, Zhang L, Zhang Y M, Dong S B, Liu Y, WU Q and Shang J. 2022. Research on retrieval algorithm of terahertz ice cloud sounding based on neural network. National Remote Sensing Bulletin, 26(10): 2043-2059 [DOI: 10.11834/jrs.20210110]

1 引言

冰云位于大气对流层上层6—15 km高度范围(王虎等, 2017), 完全或者部分由冰相粒子组成(张欣等, 2018)。由于冰云对能量收支、水循环等地球物理现象会产生重要影响, 因此需要在大气数值模式中对冰云进行合理的表示(王虎等, 2017)。然而, 目前全球大气模式之间的冰云分布存在显著差异, 全球和区域平均云冰量相差一个数量级(Mendrok等, 2009), 造成这种模式间差异的主要原因是当前还缺乏足够的全球冰云探测数据。因此, 获取全球冰云参数(高度、冰水路径、等效粒径)的真实观测对于改善大气数值模式中的对冰云的表征性能, 减少气候预测和气象

预报的不确定性具有重要研究意义(Liou, 1986)。

冰云中冰晶粒子尺寸主要集中在20—600 μm (Garrett等, 2005; Donovan, 2003), 目前一些现有的卫星遥感技术能够实现对部分冰云参数的探测, 但是都存在各自的局限(刘磊等, 2020)。微波辐射计探测仪只对500 μm 以上的大尺寸冰晶敏感, 而红外和可见光频段的探测仪只对50 μm 以下的小尺寸冰晶敏感(Waliser等, 2009); 而处于以上两者之间的太赫兹频段, 其波长范围与冰云粒子尺寸分布接近, 探测范围能够覆盖冰云的全部冰晶尺寸, 兼顾了穿透能力与敏感性, 因此星载太赫兹辐射探测仪对于冰云探测具有独特的优势(王虎等, 2017), 全球在该领域已经开展多年研究。美国研制出机载太赫兹冰云探测仪CoSSIR

收稿日期: 2020-05-09; 预印本: 2020-09-29

基金项目: 国家自然科学基金(编号: 42275141); 国防科工局民用航天十三五预先研究基金(编号: D040110); 上海航天科技创新基金(编号: SAST2019-102)

第一作者简介: 陈柯, 研究方向为微波大气遥感、微波辐射计研制。E-mail: chenke@hust.edu.cn

通信作者简介: 商建, 研究方向为微波遥感探测及卫星图像定位。E-mail: shangjian@cma.gov.cn

(Compact Scanning Submillimeterwave Imaging Radiometer) 并开展了一系列机载试验, 使用蒙特卡罗贝叶斯算法反演冰云廓线、冰水路径 IWP (Ice Water Path) 和等效粒径 D_{mc} (the median mass equivalent sphere diameter), 验证了被动太赫兹具备探测冰云的能力 (Evans 等, 2005 和 2012)。Jiménez 等 (2007) 利用辐射传输模式对太赫兹频段星载冰云探测亮温进行模拟, 并利用神经网络算法完成了对冰水路径 IWP、等效粒子尺寸 D_{mc} 和等效云高 Z_{mc} (the median IWP cloud altitude) 的反演。目前, 欧空局已经在欧洲第二代气象卫星 MetOp (Meteorological Operational Satellite Program) 上规划研制了星载太赫兹冰云成像仪 ICI (Ice Cloud Imager), 探测频段包括 5 个, 即 183 GHz、243 GHz、325 GHz、448 GHz 和 664 GHz, 计划 2022 年发射 (Kangas 等, 2014)。

冰云的冰相粒子除了固态的冰粒子 (Ice) 外, 还存在有空气、液态水和固态冰混合构成的霰粒子 (Graupel)。两者介电常数、粒径分布都有一定差异, 但是目前的太赫兹冰云探测研究通常将冰云视为仅由冰粒子构成, 仅仅反演冰粒子参数, 与冰云真实情况存在差异 (Piyush 等, 2017; Brath 等, 2018)。本文利用不同太赫兹频段亮温对冰、霰两种粒子的辐射特性差异, 提出一种预分类的神经网络算法从 183—874 GHz 多频段太赫兹亮温数据中反演得到冰、霰两种粒子参数。为了验证该算法, 本文使用 WRF (Weather Research and Forecasting Model) 模式预报场的冰密度廓线 IWC (Ice Water Content) 和 ATMS (Advanced Technology Microwave Sounder) 微波载荷实际探测的霰密度廓线 GWC (Graupel Water Content) 数据构建冰云数据库, 再使用 DOTLRT (Discrete-Ordinate Tangent Linear Radiative Transfer) 辐射传输模式模拟 183 GHz、243 GHz、325 GHz、448 GHz、664 GHz 和本文 874 GHz 这 6 个频段 (表 1) 的星载太赫兹冰云探测亮温, 然后使用神经网络算法从太赫兹亮温数据中反演得到冰、霰两种粒子的参数和密度廓线, 并对冰云探测频率通道选择、反演算法优化、探测精度分析等太赫兹冰云探测的关键问题开展研究。本文研究成果可为中国将来发展太赫兹冰云探测载荷提供技术支撑和反演算法。

表 1 太赫兹冰云探测仿真频率通道

Table 1 Frequency channels of Terahertz ice cloud sounding simulation experiments

编号	中心频率/GHz	灵敏度/K
1	183.31±7.0	0.6
2	183.31±3.4	0.7
3	183.31±2.0	0.7
4	243.2±2.5	0.6
5	325.15±9.5	1.1
6	325.15±3.5	1.2
7	325.15±1.5	1.4
8	448.0±7.2	1.3
9	448.0±3.0	1.5
10	448.0±1.4	1.9
11	664.0±4.2	1.5
12	874.38	2.5

2 太赫兹冰云探测亮温模拟

2.1 辐射传输模式

目前全球还没有在轨运行的太赫兹频段大气探测载荷, 因此本文使用大气辐射传输模式计算模拟的太赫兹冰云探测亮温数据。辐射传输模式通过对微分辐射传输方程求解实现从大气物理参数到大气辐射亮温的映射计算。

DOTLRT 是美国 NOAA (The National Oceanic and Atmospheric Administration) 的科学家 Voronovich 等 (2004) 使用离散纵标法对差分辐射传输方程求解而实现的离散正切线性辐射传输模式。DOTLRT 输入参数除了常规的表面辐射率、气压、高度、大气温度和水汽廓线外, 还支持五中水凝物, 包括云液态水、雨、冰、雪、霰, 仿真频率可高至 1000 GHz, 适用于本论文的太赫兹频段仿真。本文研究的太赫兹冰云探测频率通道采用了与 ICI 相同的频率再外加更高 874.38 GHz 频率, 以提升对小尺寸粒子探测的灵敏度。仿真频率列表和各个频率通道的灵敏度如表 1 所示。

2.2 冰云数据库的构建

试验案例选择了 2016 年经过中国南海的台风“尼伯特”和台风“莫拉蒂”。台风“尼伯特”分析时刻为 2016 年 7 月 7 日 UTC: 04:45, 台风“莫拉蒂”为 2016 年 9 月 13 日 UTC: 05:00。大气和冰云参

数以美国 NCEP (National Centers for Environmental Prediction) 的 GFS(Global Forecast System for Medium Range) 再分析资料作为边界场和初始场, 使用 WRF数值预报模式 6 h 预报输出。

为了验证冰云大气数据库和模拟冰云观测亮温数据的可信度, 首先将时间、空间、频率匹配的 DOTLRT 模拟亮温数据与美国 NPOESS (The National Polar-orbiting Operational Environment Satellite System) 气象卫星上的 ATMS 微波大气探测载荷的真实观测亮温数据进行对比。ATMS 载荷工作在 23.8—183.31 GHz 频段的 22 个频率通道 (Weng 等, 2012), 本文只选择了对冰云敏感的

183 GHz 频段 5 个频率通道进行对比。图 1 和图 2 分别展示了 ATMS 18 通道 183.31 ± 7 GHz 和 22 通道 183.31 ± 1 GHz 两个频率的台风“尼伯特”场景亮温对比, 其中图 1 (a) 和图 2 (a) 是 ATMS 真实观测亮温图像, 图 1 (b) 和图 2 (b) 是使用 WRF 大气参数输入 DOTLRT 模式模拟的亮温图像, 可以发现台风对流区域的冰云模拟亮温与 ATMS 真实观测值差异较大, 仅能在台风眼外围区域观察到冰云散射造成的低温值, 且周围的螺旋云带分布与观测亮温也不太一致, 冰云区域的低温值与观测值相比也不太一致。

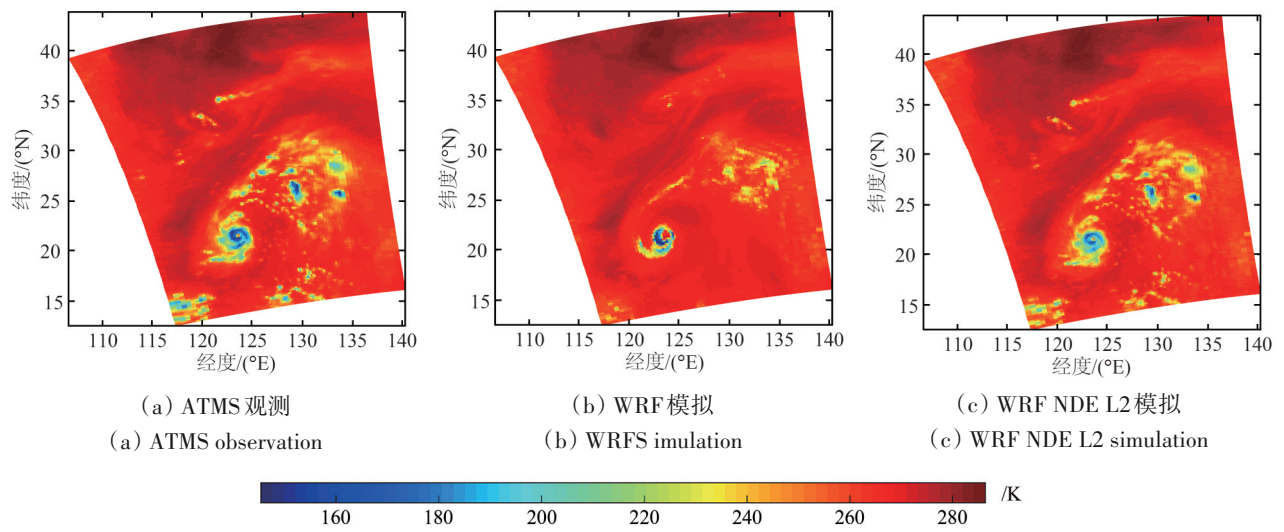


图 1 ATMS 18 通道 183.31 ± 7 GHz 亮温对比

Fig. 1 Comparison of ATMS channel 18 183.31 ± 7 GHz brightness temperature

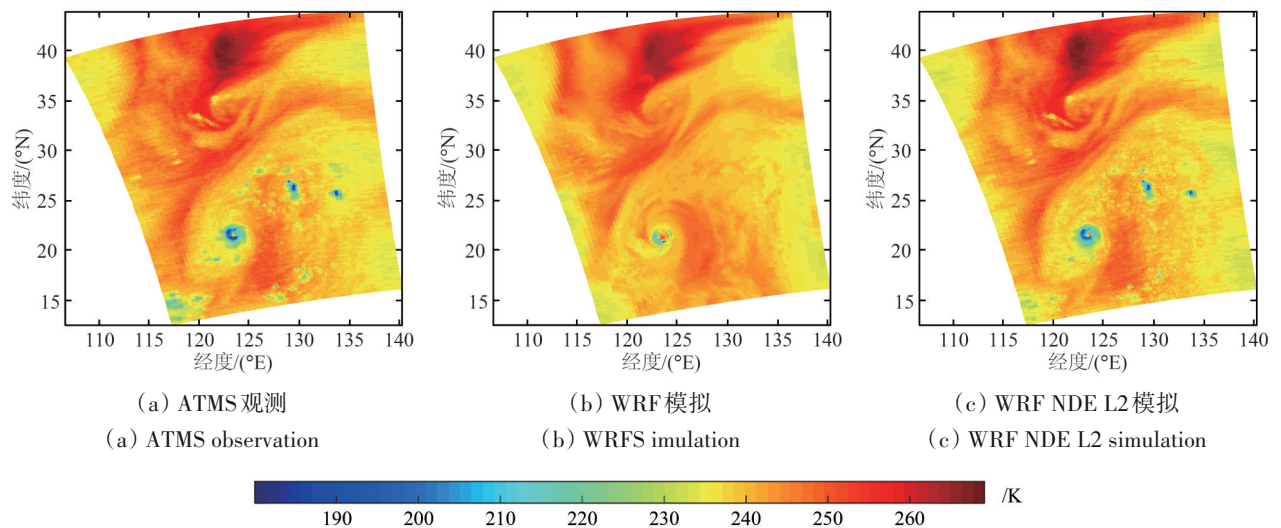


图 2 ATMS 22 通道 183.31 ± 1 GHz 亮温对比

Fig. 2 Comparison of ATMS channel 22 183.31 ± 1 GHz brightness temperature

为了提高冰云数据库的准确性，本文将美国 NOAA 发布的 ATMS 载荷大气参数产品 NDE L2 级数据也引入冰云数据库的构建。NDE L2 数据产品是由 NOAA 开发的微波综合反演系统 MIRS (Microwave Integrated Retrieval System) 从 ATMS SDR (Sensor Data Records) 观测亮温数据中反演生成的 (Boukabara 等, 2011)。由于 ATMS 最高探测频段只到 183 GHz, 因此对于冰云相关的冰密度廓线 IWC 和霰密度廓线 GWC, NDE L2 数据只提供了 GWC 数据, IWC 数据是无效的。图 3 展示了“尼伯特”场景的 WRF 和

NDE L2 数据的霰粒子路径 GWP 分布对比。从图 3 中可以看出, WRF 数据的霰粒子仅仅分布在台风中心外围, 而 NDE L2 数据的霰粒子不仅分布在台风中心区域, 在外围螺旋云带区域也存在, 与 ATMS 观测亮温图像中的低温冰云区域分布更加一致。将来如果 ATMS NDE L2 GWC 数据被用于实际冰云探测卫星的反演中, 与仿真试验不同的一点是必须考虑 ATMS GWC 数据与实际卫星亮温数据的时空匹配问题, 这会引起反演精度的下降, 具体将取决于两个卫星数据时空匹配的接近程度。

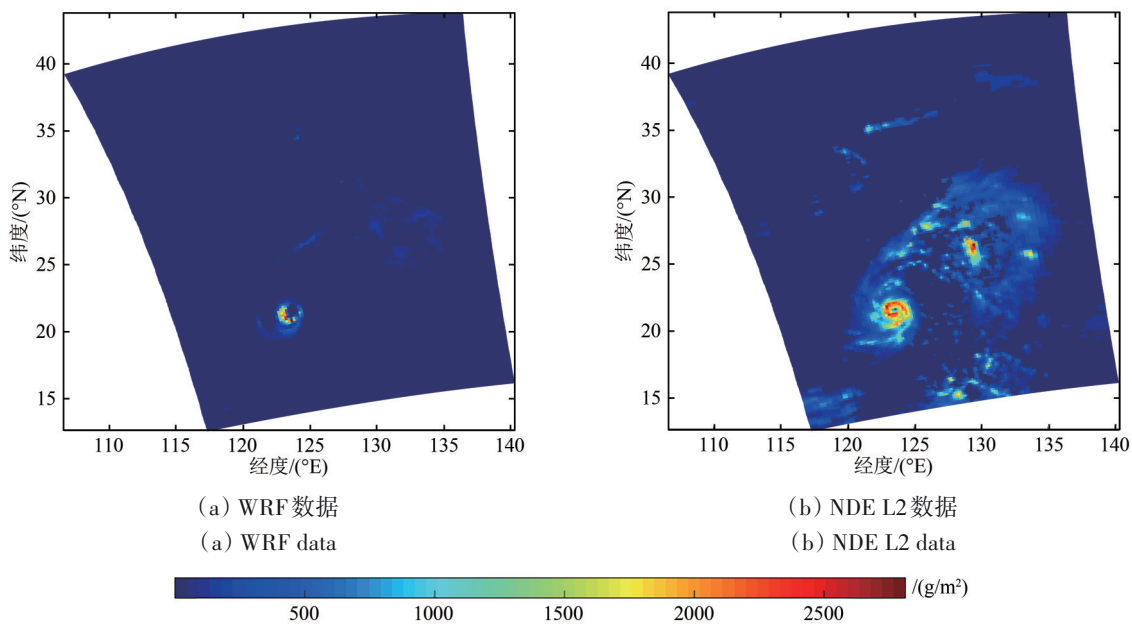


图3 WRF和NDE L2数据的GWP分布

Fig. 3 GWP distribution of WRF data and NDE L2 data

因此本文将 NDE L2 级数据中的 GWC 数据替换掉冰云数据库中的 WRF 数据, WRF 和 NDE L2 混合数据模拟的冰云亮温如图 1 (c) 和图 2 (c) 所示。可以看出用 NDE L2 霰数据替换 WRF 霰数据之后 DOTLRT 模型计算的亮温明显与 ATMS 观测亮温数据更加接近, 台风中心区域和外围螺旋云带区域的冰云散射效应也很明显, 与观测亮温一致性较好。图 4 列出了 ATMS 183 GHz 频段 5 个频率通道的两种模拟亮温与 ATMS 真实观测亮温数据的均方根误差 RMSE, 可以看出数据分析结果与视觉观感是一致的, 加入真实观测霰数据的混合冰云数据库输入 DOTLRT 计算的模拟亮温具有更小的误差, 在 183 GHz 最远翼频率通道 183.31±7 GHz, RMSE 从 WRF 模拟亮温的 11.9 K 下降到 3.9 K, 下降幅度最大, 在最中心的 183.31±1 GHz 频率误差也

从 4.1 K 下降到了 2.7 K。

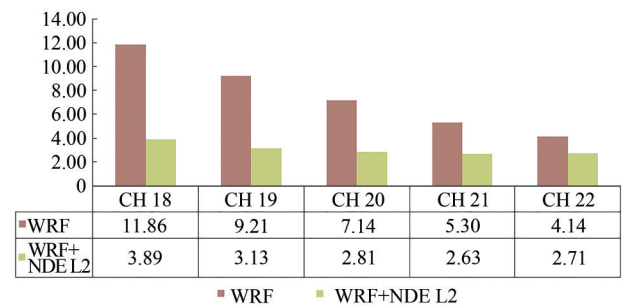


图4 ATMS 183 GHz 观测亮温与模拟亮温的 RMSE

Fig. 4 Root mean square errors of the simulated brightness temperature with respect to ATMS observation at 183GHz band

ATMS 观测亮温对比验证试验表明: (1) WRF 模式预报的台风对流区域的冰云霰密度廓线 GWC 与真实数值差异较大, 而从对比验证结果来

看, ATMS探测的NDE L2的GWC数据相比WRF模式预报的GWC数据更加准确; (2) 在183 GHz频段冰云的霰粒子对亮温的影响是最明显的; (3) DOTLRT模拟亮温可以较好的反映台风对流区域的冰云特征。因此本文采用WRF数据+NDE L2霰数据混合构建大气冰云数据库。

2.3 冰云参数

不同大气数值模式GCM (Global Climate Model) 中的冰云参数分布的差异可以高达一个数量级, 是目前大气数值模式不确定性的最大来源, 这些冰云参数差异产生的原因主要是由于不同的数值模式对冰粒子特性和将云冰转化为降水的阈值做出了不同的假设 (Heymsfield 和 Iaquinta, 2000)。为了提高气候和气象预报水平, 有必要对全球数值模式中冰云的关键分布参数进行观测验证并解决其差异 (Buehler等, 2007), 这些关键参数包括冰云含量廓线 (XWC)、总路径 (XWP)、有效粒子尺寸 (D_{me}) 和等效云高 (Z_{me}), 这些也是本论文要研究的冰云参数 (X可以代表不同种类粒子, 例如冰含量IWC、霰含量GWC)。下面给出这些冰云参数的具体定义。

XWC是冰云粒子含量, 代表冰云粒子在垂直高度上的密度分布, 单位为 g/m^3 ; XWP是冰云总路径, 是XWC在高度上的积分 (Buehler等, 2007), 单位: g/m^2 , 定义为

$$XWP = \int XWC dz \quad (1)$$

D_{me} 是粒子的质量加权中位尺寸, 为粒子等效直径 (Buehler等, 2007), 单位: μm , 定义为 (Jiménez等, 2007):

$$\frac{XWP}{2} = \int_0^{D_{me}} m(D) N(D) dD \quad (2)$$

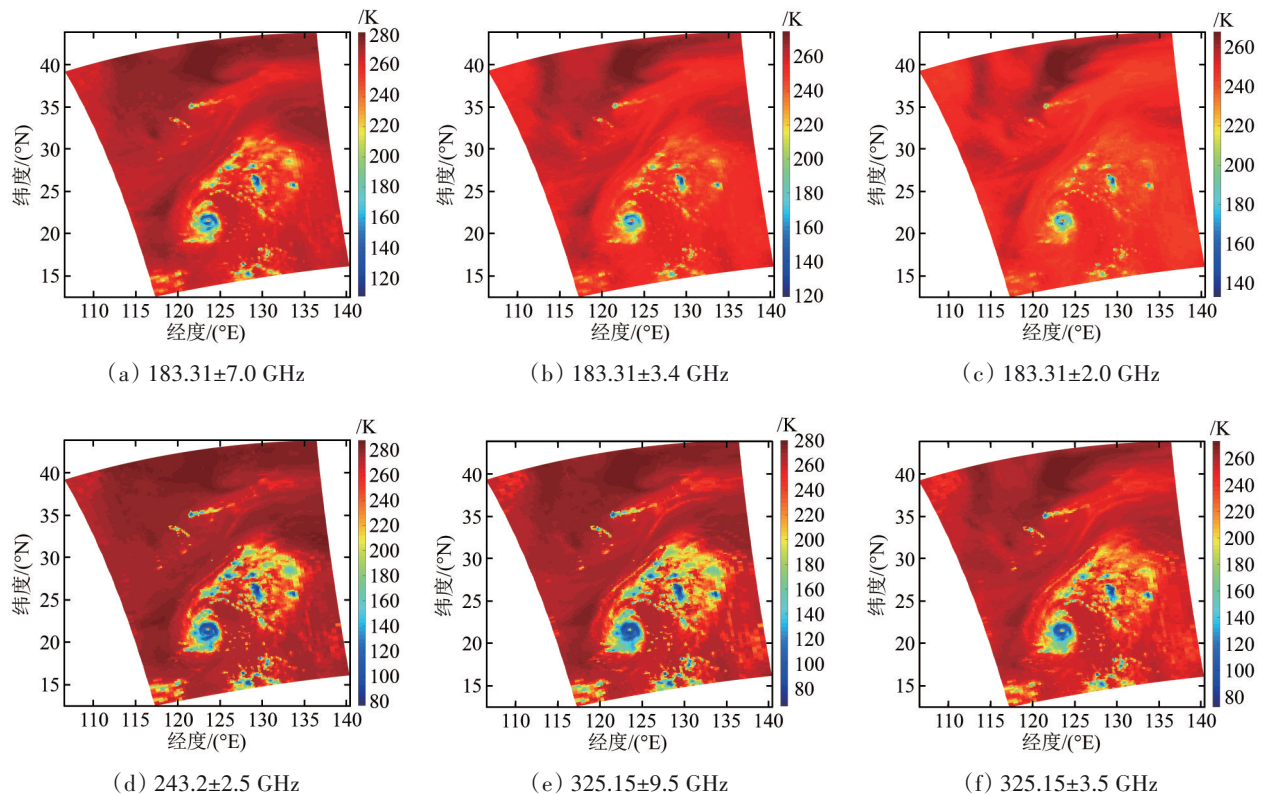
式中, $N(D)$ 为粒子尺寸分布函数, 表示等效球形半径 D 和 $D+dD$ 之间的粒子数。

Z_{me} 等效云高定义为XWP一半时对应的高度 (Buehler等, 2007), 单位: km, 定义为

$$\frac{XWP}{2} = \int_0^{Z_{me}} XWC dz \quad (3)$$

2.4 模拟太赫兹冰云亮温

图5展示了WRF数据+NDE L2霰数据混合大气冰云数据库输入DOTLRT模式计算得到的台风“尼伯特”场景的太赫兹模拟亮温图像, 频率如表1所示。



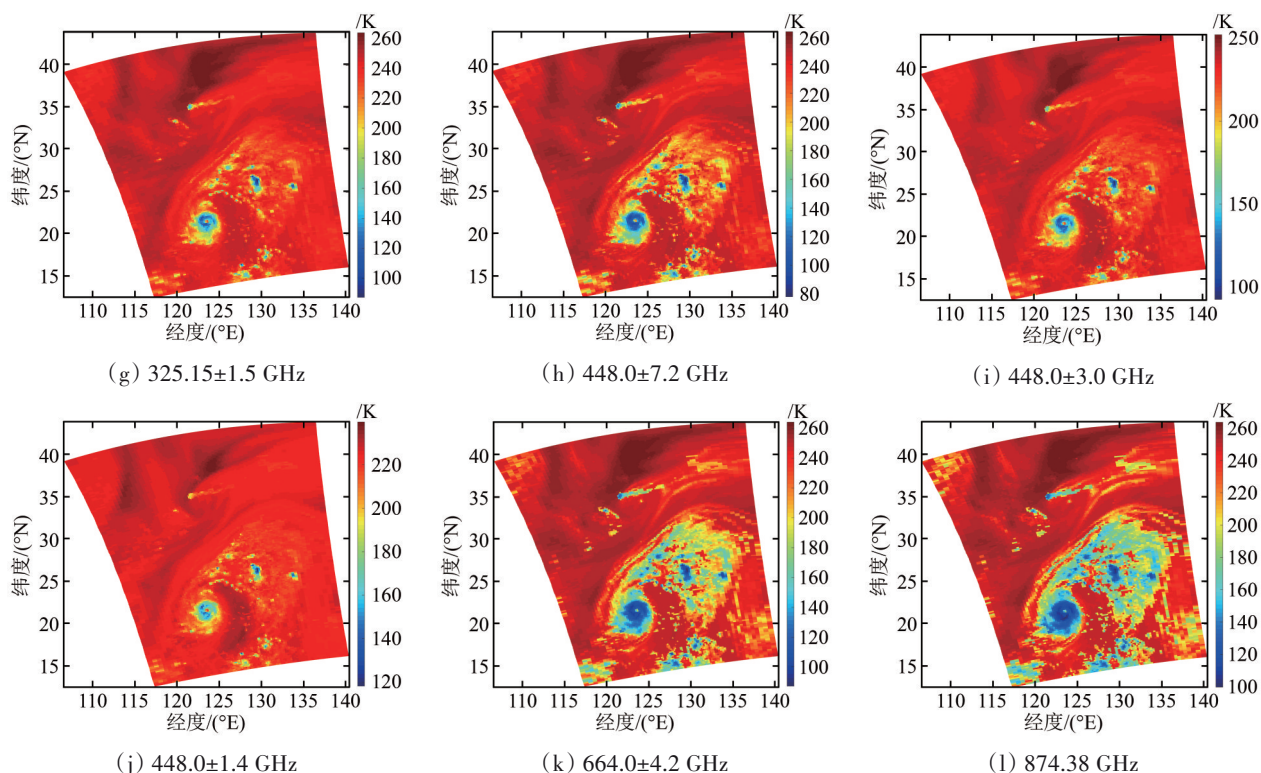


图5 台风“尼伯特”场景的太赫兹模拟亮温图像

Fig. 5 Simulated brightness temperature image of typhoon Nepartak at Terahertz band

从图5中183—874 GHz亮温图像来看, DOTLRT计算的亮温图像能够反映台风对流区域的冰云粒子散射造成的低温值, 与ATMS真实观测的冰云区域一致, 而且在太赫兹频段最高的664 GHz和874 GHz, 低温区域明显扩大, 说明在太赫兹高频更小的冰云粒子散射效应变强, 能够对大气上视辐射亮温造成下降, 这与辐射传输理论也是吻合的。

3 霰粒子反演仿真试验

3.1 实验方案和数据

将冰云分为冰、霰两种粒子构成时, 太赫兹频段的冰云亮温下降是由两种粒子共同散射作用引起的, 由于冰、霰两种粒子尺寸分布、介电常数不同, 具有不同的散射谱特性 (Wang等, 2016), 因此在使用神经网络反演冰云参数时需要将两种粒子分开。图6展示了仅考虑冰粒子对亮温散射作用时不同冰水路径IWP下的冰云辐射亮温特性谱, 可以看出183 GHz频段几乎不受冰粒子散射影响。而本文2.2节ATMS观测亮温与模拟亮温的对比验证试验表明, 183 GHz频段亮温对冰云的霰粒子有足够的灵敏度。利用183 GHz频段冰、霰两种粒子的辐射特性差异, 可以使用神经网络算法从183 GHz

频段亮温先单独反演冰云的霰粒子参数。

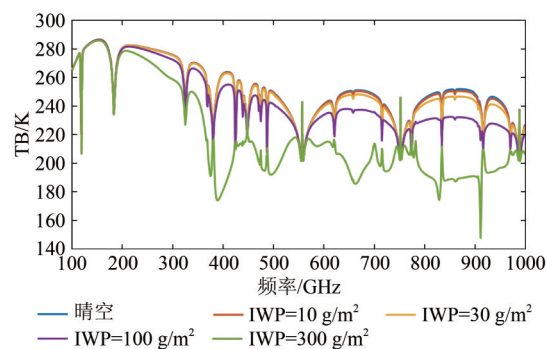


图6 冰粒子亮温特性谱

Fig. 6 Spectrum of brightness temperature of ice particles

神经网络算法的训练数据库需要两部分数据: 一是亮温数据作为输入, 由DOTLRT模拟亮温实现; 二是冰云数据作为输出, 由输入到DOTLRT模式的冰云数据库实现。将两部分数据组成数据对, 选择以台风“莫拉蒂”的亮温—冰云数据对作为训练数据集, 台风“尼伯特”的数据对作为测试数据集。本文用神经网络算法分别实现冰云霰粒子的统计参数和垂直密度廓线GWC的反演, 反演的霰参数包括路径总量 (GWP)、等效粒径 (G_D_{me})、等效云高 (G_Z_{me}) 这3个参数。

3.2 霰参数反演试验及结果分析

冰云霰粒子的统计参数反演试验流程如下：

(1) 对数据库中的亮温数据进行筛选，筛选依据是霰粒子散射在 183.31+7 GHz 频率引起的亮温差超过该频率通道灵敏度 0.6 K 时，则认为对应的霰粒子总量 GWP 能够被准确探测到，从 DOTLRT 亮温中计算得到对应 GWP 门限阈值为 12 g/m²。

(2) 使用作为训练集的台风“莫拉蒂”数据训练神经网络，网络隐藏层采用 32 个神经元，训练函数为 *trainscg*，网络收敛判别指标采用 MSE（均方误差）。神经元个数 N_h 通过以下经验公式得到：

$$N_h = \frac{N_s}{(\alpha \times (N_i + N_o))} \quad (4)$$

式中， N_i 是输入层神经元个数； N_o 是输出层神经元个数； N_s 是训练集样本数； α 是可以自取的任意值变量，通常范围可取 2—10。

(3) 将训练好的网络对测试数据集中的台风“尼伯特”模拟观测亮温进行反演，得到霰参数。

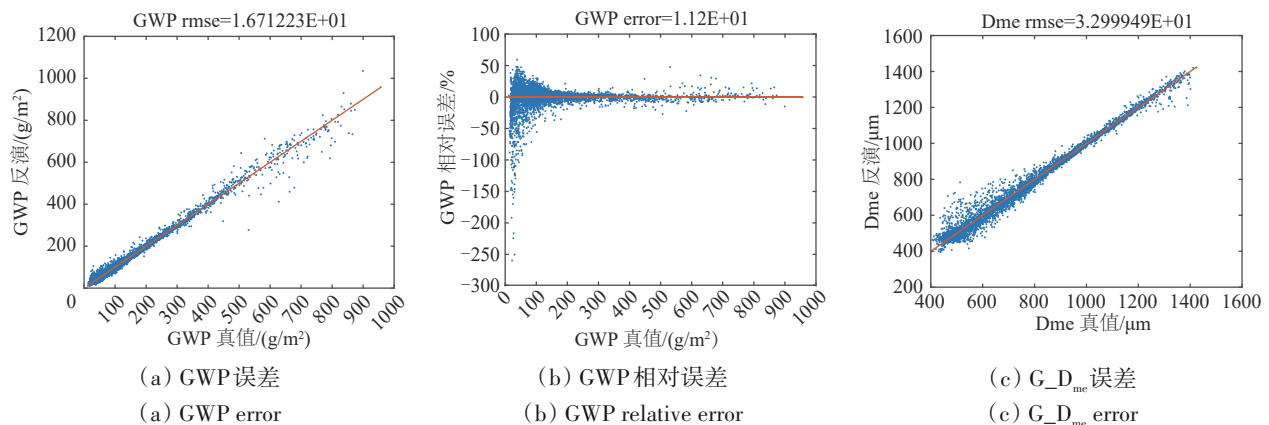
(4) 将反演得到的霰参数估值与测试数据集中对应的真实霰参数进行对比分析，评估反演精度。

图 7 展示了反演的 3 种霰参数的散点图，横轴代表真值，纵轴为反演参数值。分别使用了 4500 个样本训练，3000 个样本测试。图 7 (a) 是反演的 GWP 散点图，总体上大部分数据点都分布在 $x=y$ 的对角线附近，说明反演 GWP 值与真值比较接近，均方根误差 RMSE 为 16.7 g/m²。从数据分布来看，绝大部分冰云样本点位于 $GWP < 500$ g/m² 的范围，而在 $GWP > 500$ g/m² 的范围里绝对误差变得更大，出现更多明显偏离对角线的散点；而从图 7 (b) 展示的 GWP 相对误差散点图来看，反而是 GWP 越小相对误差越大，总体来说 95% 以上测试数据的 GWP 反演相对误差能够控制在 $\pm 50\%$ 以内，只有 $GWP < 100$ g/m²

的部分样本相对误差超过这个范围；图 7 (c) 是反演的 $G_{D_{mc}}$ 散点图，总体上大部分数据点都分布在 $x=y$ 的对角线附近，说明反演 $G_{D_{mc}}$ 值与真值比较接近，均方根误差 RMSE 为 33.0 μm ；从数据分布来看，绝大部分冰云样本点位于 $G_{D_{mc}} < 1400$ μm 的范围，而在 $G_{D_{mc}} < 800$ μm 的范围里绝对误差变得更大，出现更多明显偏离对角线的散点；而从图 7 (d) 展示的 $G_{D_{mc}}$ 相对误差散点图来看，反而是 $G_{D_{mc}}$ 越小相对误差越大，大部分测试数据的 $G_{D_{mc}}$ 反演相对误差能够控制在 $\pm 20\%$ 以内；图 7 (e) 是反演的 $G_{Z_{mc}}$ 散点图，总体上都分布在 $x=y$ 的对角线两侧，均方根误差 RMSE 为 523.1 m；从数据分布来看，绝大部分冰云样本点位于 $7 \text{ km} < G_{Z_{mc}} < 12 \text{ km}$ 的范围，而在 $8 \text{ km} < G_{Z_{mc}} < 9 \text{ km}$ 和 $G_{Z_{mc}} > 10 \text{ km}$ 的范围里绝对误差变得更大，出现更多明显偏离对角线的散点；而从图 7 (f) 展示的 $G_{Z_{mc}}$ 相对误差散点图来看，大部分测试数据的 $G_{Z_{mc}}$ 反演相对误差能够控制在 $\pm 20\%$ 以内。

3.3 霰廓线 GWC 反演试验及结果分析

冰云霰粒子的密度廓线 GWC 的神经网络反演试验流程与 3.2 节的霰参数反演流程类似，只是将霰参数改成 GWC。图 8 展示了部分测试样本的反演结果，其中蓝线代表真实 GWC 廓线，红线代表反演 GWC 廓线。可以看出，不同测试样本的反演精度差异很大，除了第一条样本廓线外，其它测试样本的反演 GWC 廓线与真实廓线差异较大，说明神经网络算法直接反演冰云霰粒子的 GWC 廓线效果并不理想。主要原因在于神经网络算法非常依赖于训练数据的表征性，变化多样、过于复杂的训练数据会造成算法性能下降 (Zhang 等, 2016)。图 8 是随机选取的一组霰廓线反演结果，这些霰廓线就是一个典型的例子。



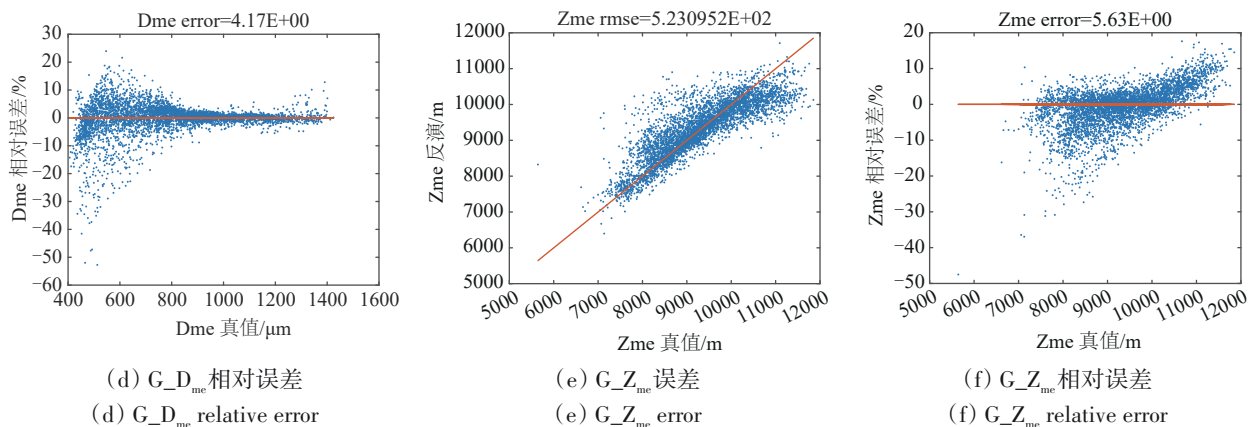


图7 霰参数反演结果散点图

Fig. 7 Scatterplot of the retrieved graupel parameters

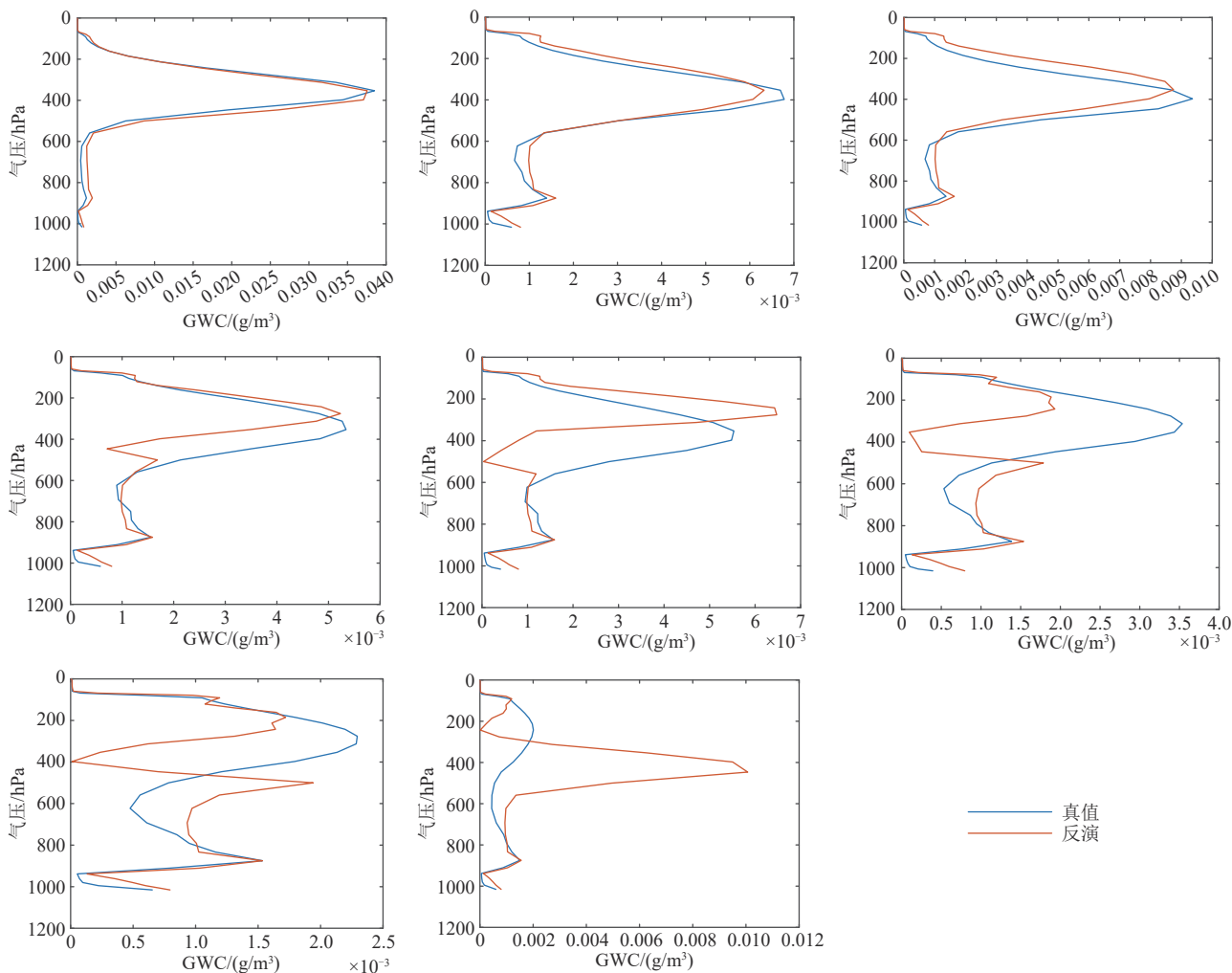


Fig. 8 Retrieval results of GWC using unclassified neural network

解决这种复杂数据问题的一个思路就是对数据进行预分类，将一个复杂数据集的反演问题分成多个相对简单数据集的反演问题。本文试验的分类思路是根据霰廓线的垂直分布特征进行分类，

例如如图8显示的一个明显分类特征是霰廓线两个峰值的大小比例，但是由于实际应用时冰云霰廓线的分布是未知的，因此不能直接根据廓线形状进行分类。而我们研究发现，GWP对具有不同形状

的霰廓线样本具有较好的指示关系, 可以将GWP值作为霰廓线样本分类的指示器, 如图9所示。图9中样本点的不同颜色代表具有不同形状的霰廓线, 蓝色代表低层与高层峰值比 $<1/3$; 橙色为峰值比在 $1/3—2/3$ 之间; 黄色为峰值比在 $2/3—5/3$ 之间; 紫色为峰值比在 $5/3—5$ 之间; 绿色为峰值 ≥ 5 , 可以看出不同类型霰廓线样本集对应的GWP值有明显的差异。从图9可以看出, 大多数样本都属于第一类蓝色霰廓线, 第四类紫色和第五类绿色样本数量很少, 而且已经处于设定的可检测GWP阈值以下, 因此可以将所有霰廓线按照GWP值划分为3类, 对应的GWP取值区间分别为 $GWP>35 \text{ g/m}^2$ 、 $25 \leq GWP < 35 \text{ g/m}^2$ 和 $15 \leq GWP < 25 \text{ g/m}^2$ 这3种阈值。

加入数据预分类的霰廓线神经网络反演试验流程如下: (1) 对训练数据集中的冰云廓线样本根据GWP的3种阈值进行分类; (2) 使用3种类型的训练样本分别训练3个神经网络; (3) 对测试数据集中的所有亮温数据使用3.2节中的神经网络反演得到GWP; (4) 将测试数据集的亮温数据样本根据上面的GWP的3种阈值进行分类; (5) 对分类的亮温数据样本使用对应的网络反演得到霰廓线GWC。

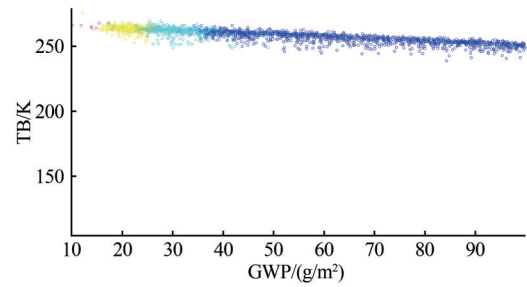


图9 GWC廓线分类与GWP的关系

Fig. 9 Relationship between GWC profile classification and GWP

预分类的GWC反演试验结果如图10所示。分类得到的第一类廓线共有3620个样本, 第二类478个, 第三类281个样本, 可见IWP越小、冰云越薄实际中的样本数越少。图10(a)、10(c)、10(e)中红线是反演GWC的均值, 蓝线是RMSE; 图10(b)、10(d)、10(f)是RMSE与GWC真值比值, 代表相对误差。表2列出了3类样本以及全部样本在不同气压层的反演误差。其中, nn 1表示分类后的第一类样本 ($GWP>35 \text{ g/m}^2$), nn 2表示分类后的第二类样本 ($25 \leq GWP < 35 \text{ g/m}^2$), nn 3表示分类后的第三类样本 ($15 \leq GWP < 25 \text{ g/m}^2$)。

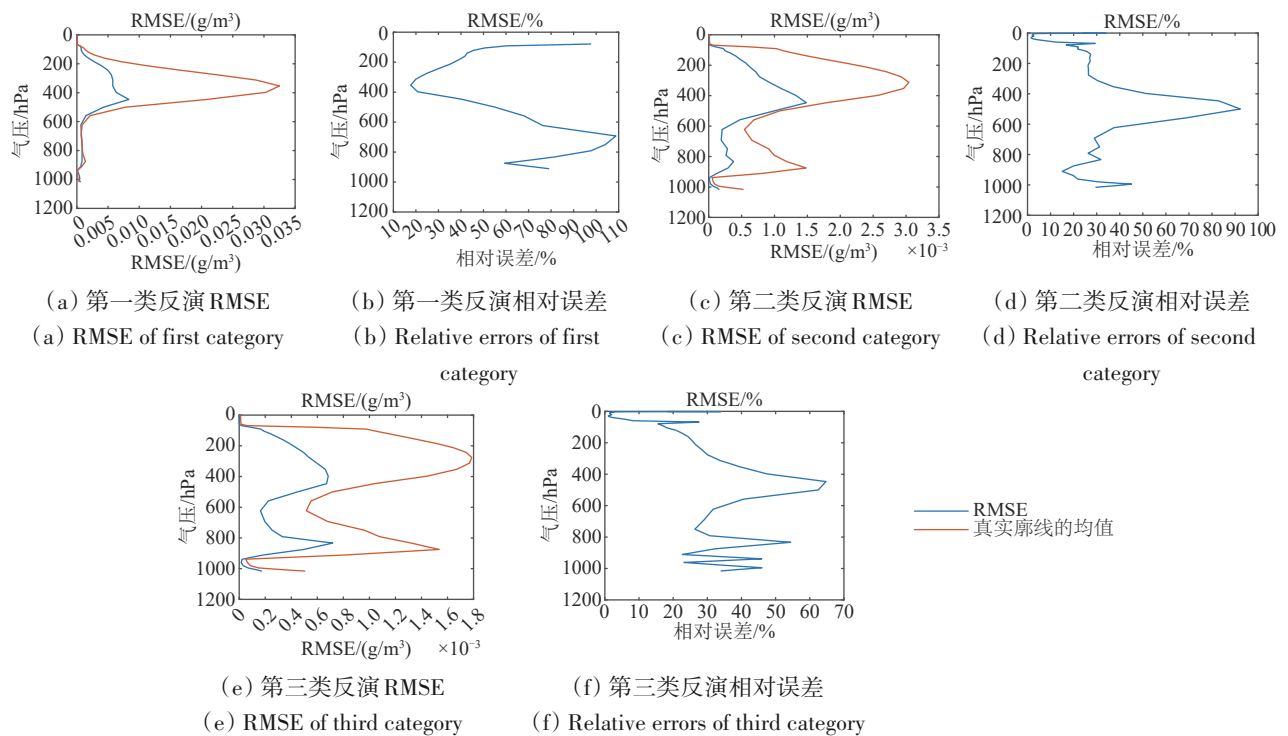


图10 采用预分类神经网络的霰廓线GWC反演结果

Fig. 10 Retrieval results of GWC using pre-classified neural network

从反演的绝对误差来看, 第一类样本误差值最大, RMSE范围为 $0—8.2 \times 10^{-3} \text{ g/m}^3$, 第二类样本RMSE范围为 $0—1.5 \times 10^{-3} \text{ g/m}^3$, 第三类样本误差值

最小, RMSE范围为 $0—0.7 \times 10^{-3} \text{ g/m}^3$ 。这主要是由于这3类样本是按照GWP大小——即冰云厚度来划分的。但从相对误差来看, 3类样本就整体趋于

一致, 误差绝大部分都在60%以下。横向对比3类样本的相对误差, 可以发现, 850—550 hPa范围对应霰廓线第二峰值, 第一类样本误差最大, 是因为第一类样本的峰值集中在550 hPa以上, 第二峰值处霰密度值较小导致相对误差大; 450—350 hPa对应廓线第一峰值, 第一类样本误差最小, 反演精度最高。由于第一类样本数占比达到了82%, 可以看到全部样本的平均误差与第一类样本误差最为接近, 预分类主要改进了冰云较薄的二、三类样本的反演精度。

表2 霰廓线的分层反演误差
Table 2 Layered retrieval errors of GWC

气压/hPa	RMSE/(10^{-3} g/m ³)			
	nn 1	nn 2	nn 3	all
850	0.84	0.34	0.63	0.82
550	1.83	0.54	0.25	1.88
450	8.17	1.44	0.65	7.97
350	5.69	1.10	0.66	5.66
250	5.00	0.71	0.49	4.08
150	1.52	0.44	0.31	1.36
100	0.84	0.24	0.18	0.54

气压/hPa	相对误差/%			
	nn 1	nn 2	nn 3	all
850	73.42	27.02	45.39	68.14
550	64.94	70.43	42.85	69.82
450	41.43	82.92	63.64	44.03
350	17.70	36.99	39.08	20.72
250	28.60	25.75	28.10	28.96
150	42.29	26.96	23.25	45.59
100	54.92	22.04	17.57	45.76

注: nn 1为第一类样本, nn 2为第二类样本, nn 3为第三类样本, all为全部样本。

图11对比了霰廓线在未分类和预分类两种情况下的反演误差, 纵轴是压强 (hPa), 横轴是RMSE (g/m^3)。图11中统计的是二、三类廓线的误差, 它们样本数少, 提升最为明显。蓝色是未分类时二、三类廓线样本的总体RMSE, 红线和黄线分别是分类后二、三类廓线的RMSE; 可以看出分类后的二、三类样本误差值远小于未分类的误差, 验证了预分类对提升这些样本反演精度的有效性。

3.4 霰参数反演精度对比

通过反演的霰密度廓线GWC也可以计算得到GWP、 $G_{D_{me}}$ 、 $G_{Z_{me}}$ 这3种霰参数, 称为间接反演法 (Indirect Inversion Method)。本节将相同冰云样本的间接反演的霰参数精度与3.2节直接反演的结果进行对比。表3展示了两种反演方法的均方根误

差RMSE和相对误差。从数值上来看, 间接反演的GWP, RMSE从直接反演的 11.80 g/m^2 下降到了 10.90 g/m^2 , $G_{D_{me}}$ 的RMSE从 $32.77 \mu\text{m}$ 下降到了 $25.38 \mu\text{m}$, $G_{Z_{me}}$ 的RMSE从 513.62 m 上升到了 571.16 m 。看起来两种方式反演的冰云霰参数精度基本相当。

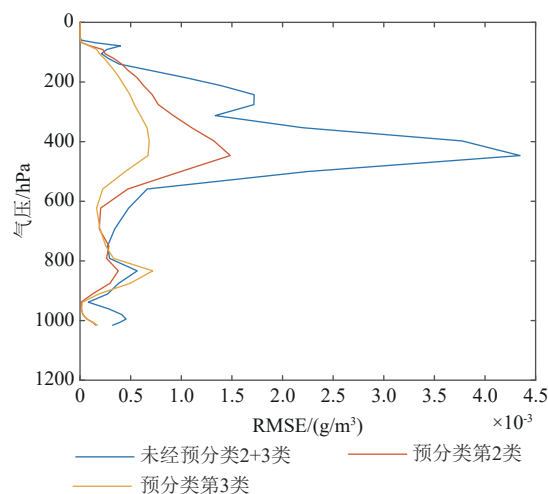


图11 未分类与预分类神经网络的GWC反演误差对比
Fig. 11 Comparison of retrieval errors of GWC using unclassified neural network and pre-classified neural network

表3 霰参数反演误差对比

Table 3 Comparison of the retrieval errors of graupel parameters

变量	直接反演	间接反演
GWP RMSE/(g/m^2)	11.80	10.90
GWP error/%	7.9	7.3
$G_{D_{me}}$ RMSE/ μm	32.77	25.38
$G_{D_{me}}$ error/%	4.1	3.2
$G_{Z_{me}}$ RMSE/m	513.62	571.16
$G_{Z_{me}}$ error/%	5.6	6.2

4 冰粒子反演仿真试验

4.1 试验方案和数据

在183 GHz频段, 亮温的冰云散射可以认为仅仅来自于霰粒子的贡献, 但在更高的太赫兹频段, 散射由冰和霰粒子共同引起 (Wang等, 2016)。因此要从太赫兹亮温中反演冰粒子参数, 需要将之前反演的霰密度廓线作为先验约束条件, 再使用神经网络算法根据冰粒子引起的散射亮温差反演得到冰粒参数。具体实现方法是首先从WRF预报大气参数中去除冰参数, 再加上反演的霰密度廓线GWC输入DOTLRT模式得到不包含冰粒散射的模拟亮温, 然后将无冰亮温和观测亮温相减,

则此亮温差代表了冰粒对大气辐射亮温的贡献。图 13 展示了太赫兹各个频率通道的台风“尼伯特”场景的冰粒子亮温差图像。

用神经网络算法对冰粒子的反演仍然分为统计参数和垂直密度廓线 IWC 的反演, 反演的冰参数也是路径总量 (IWP)、等效粒径 ($I_{D_{me}}$)、等效

云高 ($I_{Z_{me}}$), 定义与霰参数相同。观测数据使用冰粒子在 243 GHz 至 874 GHz 频段 9 个频率通道的亮温差, 如图 12 所示。训练数据集仍然选择台风“莫拉蒂”的亮温差—冰云数据对, 测试数据集使用台风“尼伯特”数据。

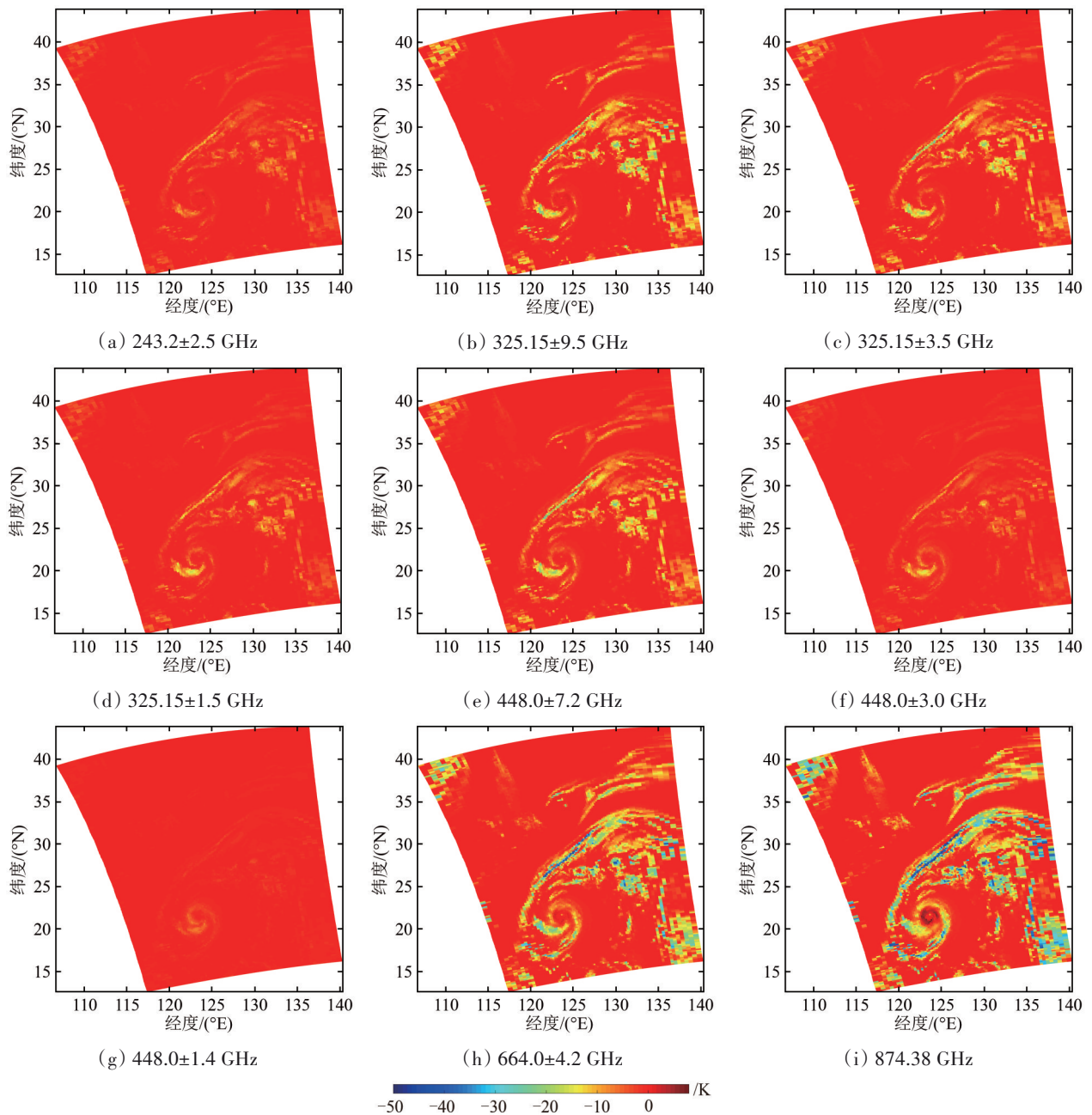


图 12 冰粒子亮温差图像

Fig. 12 Brightness temperature difference caused only by the ice particles at Terahertz band

4.2 冰参数反演仿真试验及结果分析

基于亮温差的冰参数神经网络反演算法训练和反演过程与本文 3.2 节霰参数类似, 训练网络之前

也需要对数据进行筛选。冰参数的筛选条件是当样本 IWP 引起的亮温差值超过 874.38 GHz 频率的系统噪声 2.5 K 时, 则认为对应的冰云样本可以准

确探测到, 计算得到对应IWP门限阈值为 5.5 g/m^2 。

图13展示了神经网络算法反演冰参数的结果散点图, 横轴代表真值, 纵轴是反演值, 分别使用了6000个样本训练, 6000个样本测试。图13(b)、13(d)、13(f)是相对误差分布图。总体上3种参数的大部分数据点都分布在 $x=y$ 的对角线附近, 说明反演的冰参数值总体上精度较好。IWP的RMSE为 11.34 g/m^2 , 在 $\text{IWP} > 120 \text{ g/m}^2$ 时, 绝对误

差变得更大, 出现更多明显偏离对角线的散点。从相对误差来看, 除具有非常小IWP值的样本点相对误差较大外, 总体相对误差都能控制在 $\pm 50\%$ 以内。 $I_{D_{me}}$ 的RMSE为 $10.24 \mu\text{m}$, $I_{D_{me}}$ 在 $100\text{--}140 \mu\text{m}$ 时绝对误差更大, 相对误差基本集中在 $\pm 30\%$ 以内。 $I_{Z_{me}}$ 的RMSE为 575.69 m , 在 $I_{Z_{me}}$ 低于 10 km 和高于 14 km 时绝对误差更大, 相对误差基本集中在 $\pm 20\%$ 以内。

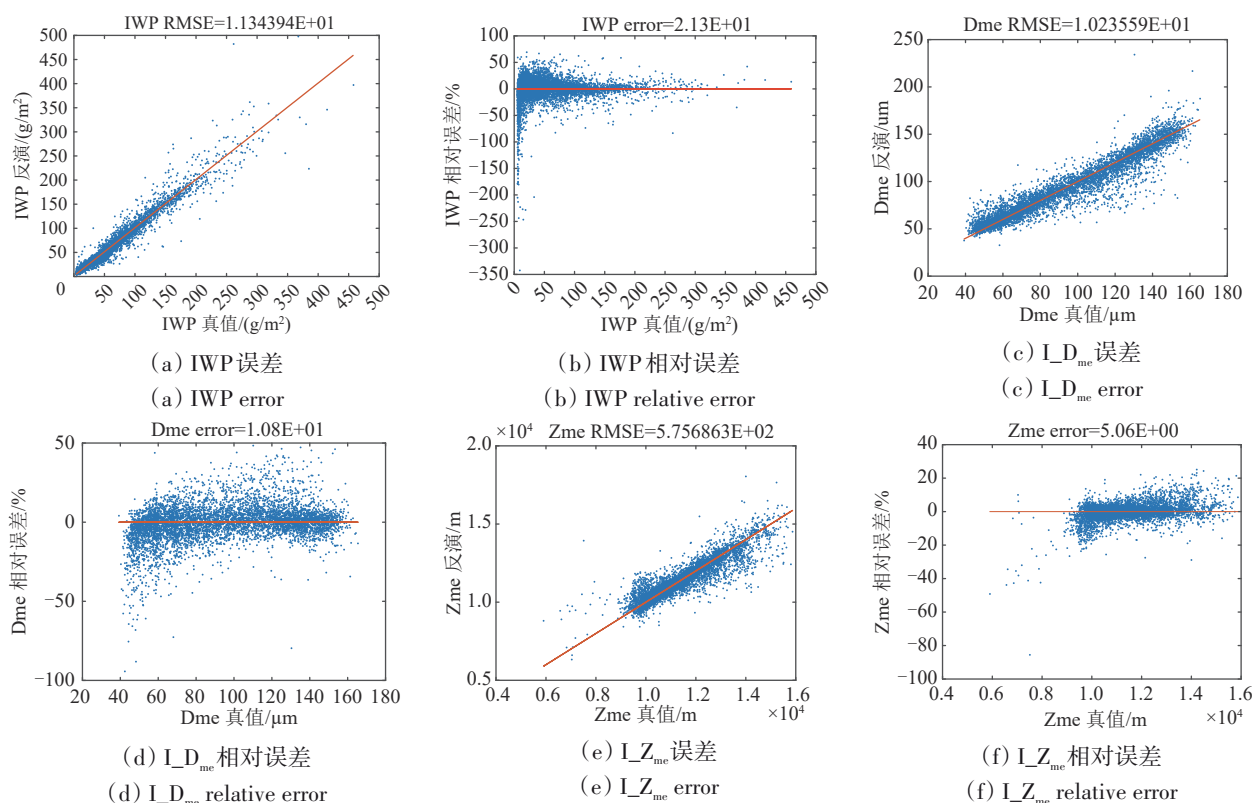


图13 冰参数反演结果散点图

Fig. 13 Scatterplot of the retrieved ice parameters

为了分析太赫兹冰云探测技术中探测频率通道选择对冰参数探测精度的影响, 本文开展了使用不同频率组合的观测数据反演冰参数的对比试验。试验测试了以下5种频率组合的观测亮温数据: (1) $325 \text{ GHz} + 448 \text{ GHz}$ 2个频段的6个频率通道; (2) $325 \text{ GHz} + 448 \text{ GHz} + 664 \text{ GHz}$ 3个频段的7个频率通道; (3) $448 \text{ GHz} + 664 \text{ GHz}$ 2个频段的4个频率通道; (4) $243 \text{ GHz} + 325 \text{ GHz} + 448 \text{ GHz} + 664 \text{ GHz} + 874 \text{ GHz}$ 5个频段的9个频率通道; (5) $664 \text{ GHz} + 874 \text{ GHz}$ 2个频段的2个频率通道。

表4列出了5种频率组合反演3种冰参数的RMSE。可以看出, 仅有最高频率的 $664 \text{ GHz} + 874 \text{ GHz}$ 组合反演的3种冰参数误差都是最大的。对于IWP, $325 \text{ GHz} + 448 \text{ GHz}$ 频段组合反演误差其

次, 而对于 $I_{D_{me}}$ 、 $I_{Z_{me}}$, 其次是 $448 \text{ GHz} + 664 \text{ GHz}$ 频段组合反演误差较大。对3种冰云参数, $325 \text{ GHz} + 448 \text{ GHz} + 664 \text{ GHz}$ 频段组合反演结果比前两种频率组合反演误差都要更小, 但是方案4的 $243 \text{ GHz} + 325 \text{ GHz} + 448 \text{ GHz} + 664 \text{ GHz} + 874 \text{ GHz}$ 频段组合方案反演精度最高。说明更多的太赫兹探测频段能够明显提升冰粒子参数反演精度。

4.3 冰廓线反演仿真试验及结果分析

冰粒子主要分布在 $7\text{--}15 \text{ km}$ 高度范围内, 且不同冰云样本的高度分布差异较大, 例如某些区域冰粒子主要集中在 10 km 以下, 某些区域集中在 10 km 以上。因此冰密度廓线IWC反演问题与GWC类似, 同样具有较高的复杂性, 使用神经网络

络算法反演 IWC 同样需要进行数据预分类, 本文使用了 4.2 节中反演的等效冰云高度 I_Z 对冰云样本进行分类。 $I_Z < 10$ km 为一类, I_Z 在 10—15 km 范围内的则按照 1 km 等高度间距分类。

表 4 不同频率组合的冰参数反演误差对比

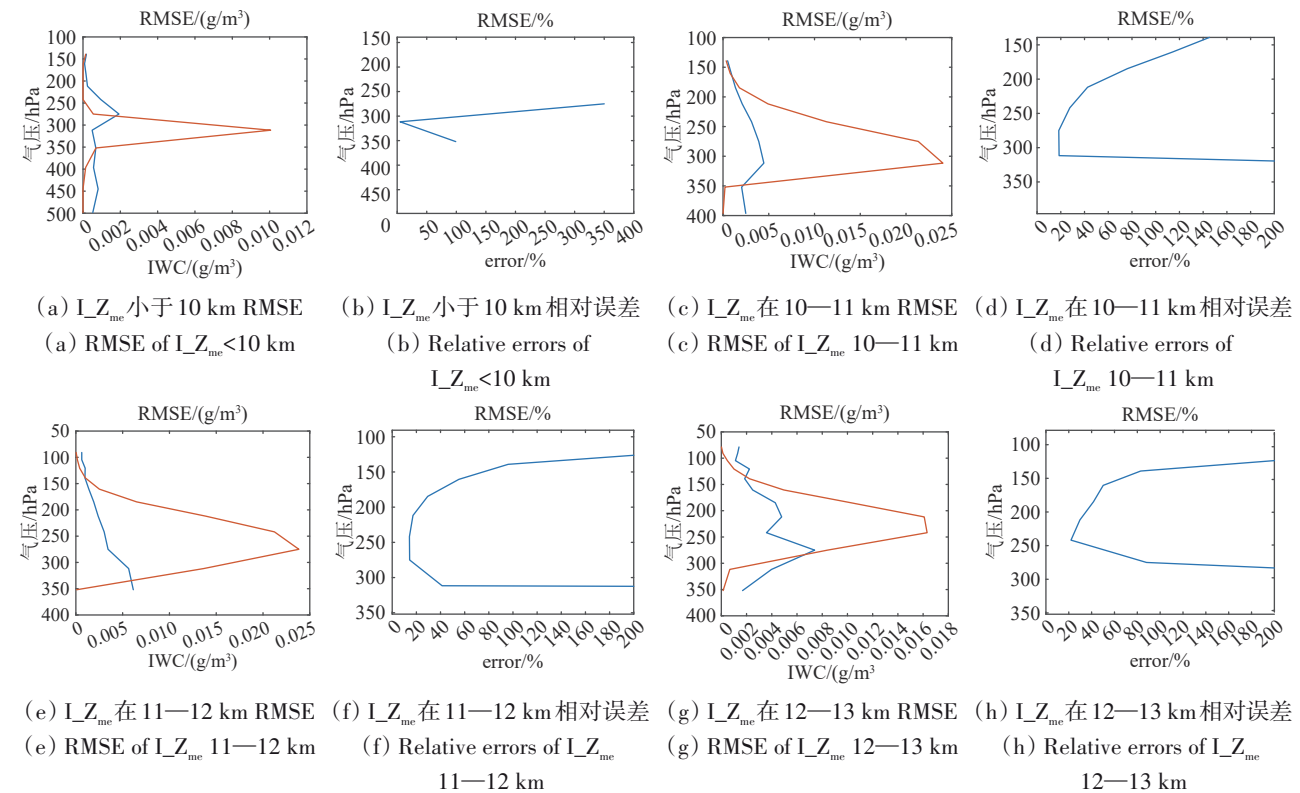
Table 4 Comparison of the retrieval errors of ice parameters with various frequency combinations

方案	IWP RMSE/(g/m ²)	$I_{D_{me}}$ RMSE/um	$I_{Z_{me}}$ RMSE/m
频率方案 1	26.10	14.83	777.55
频率方案 2	15.31	11.44	697.28
频率方案 3	17.58	15.73	919.48
频率方案 4	11.34	10.24	575.69
频率方案 5	33.52	22.75	1073.11

加入数据预分类的冰廓线神经网络反演试验流程如下: (1) 对训练数据集中的冰云廓线样本根据上述 I_Z 的 6 种阈值 (即: $I_Z < 10$ km、 10 km $< I_Z < 11$ km、 11 km $< I_Z < 12$ km、 12 km $< I_Z < 13$ km、 13 km $< I_Z < 14$ km、 14 km $< I_Z < 15$ km) 进行分类; (2) 使用 6 种类型的训练样本分别训练 6 个神经网络; (3) 对测试数据集中的所有亮温差数据使用本文 4.2 节神经网络算法反演得到 I_Z ; (4) 将测试数据集中的亮温差数据样本根据 I_Z

的 6 种阈值进行分类; (5) 对分类的亮温差测试数据使用对应的网络反演得到冰廓线。

预分类的 IWC 反演试验结果如图 14 所示, 纵坐标是大气层数, 横轴为冰粒密度。图 14 (a)、14 (c)、14 (e)、14 (g)、14 (i)、14 (k) 中红线是 6 种不同高度样本反演冰廓线的均值, 蓝线是均方根误差 RMSE; 图 14 (b)、14 (d)、14 (f)、14 (h)、14 (j)、14 (l) 是反演 RMSE 与廓线真值相比的相对百分比误差。表 5 列出了 6 种不同高度廓线在不同气压层的反演误差平均值和相对误差。从反演 IWC 的各层 RMSE 值和相对误差分布来看, 基本上也是在 IWC 值越大的高度, 反演绝对误差越大, 但是相对误差越小, 大部分冰云样本在冰云密集高度的相对误差都能控制在 50% 左右, 说明对 IWC 的反演精度较好; 但每种分类的云顶和云底位置相对误差显得偏大, 说明由于云层厚度的不同, 边界上的特征不明显, 冰云厚度的差异依旧会对反演结果造成影响。图 15 展示了几组 IWC 经预分类与未经预分类的结果, 可以看出, 常规神经网络反演的冰廓线 IWC 可能会在本没有冰分布的高度上引入较大的误差, 而预分类神经网络反演的冰廓线 IWC 更加符合实际分布。



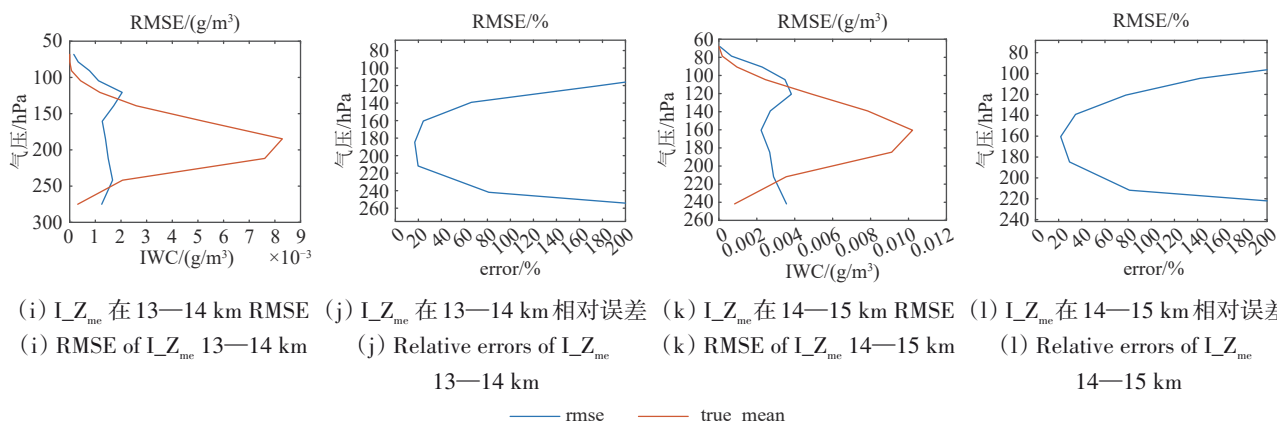


图14 采用预分类神经网络的冰廓线IWC反演结果
Fig. 14 Retrieval results of IWC using pre-classified neural network

表5 不同高度分类的冰廓线的分层反演误差

Table 5 Layered retrieval errors of IWC of ice clouds with various height

气压/hPa	RMSE/(10 ⁻³ g/m ³)						all
	<10 km	10—11 km	11—12 km	12—13 km	13—14 km	14—15 km	
500	0.52						0.09
400	0.58	2.52					1.78
300	0.95	4.28	4.94	5.07			5.06
200	0.21	1.78	2.17	4.56	1.46	2.80	2.85
150	0.13	0.70	1.18	2.17	1.49	2.47	1.51
100			0.63	1.17	1.01	3.09	1.26

气压/hPa	相对误差/%						all
	<10 km	10—11 km	11—12 km	12—13 km	13—14 km	14—15 km	
500							
400							
300	114	18.5	32.8				52.0
200		57.2	22.5	34.9	18.5	58.6	33.1
150		129	75.2	66.2	44.8	28.1	51.3
100							

结合表5中数据，从图14所示的不同高度的冰云反演误差来看，小于10 km分类采用了159个样本训练，13个样本测试，所选范围数据较少，RMSE范围为0—0.0019 g/m³，仅在IWC廓线峰值高度的反演误差<100%，此分类样本数据较少，但分布高度存在较大差异，因此反演误差较大；10—11 km分类采用了1337个样本训练，392个样本测试，反演RMSE范围为0—0.0045 g/m³，在冰云密集的高度范围350—200 hPa，相对误差在50%以内，在峰值对应高度约310 hPa，相对误差仅有18.9%；11—12 km反演RMSE范围为0—0.0061 g/m³，在冰云密集范围320—160 hPa，相对误差在50%以内，在IWC廓线峰值对应高度约270 hPa，相对误差仅为15%左右；12—13 km分类使用了813个样本训练，365个样本测试，反演RMSE范围为0—0.0074 g/m³，在廓线分布的主要高度250—160 hPa相对误差小

于50%，IWC廓线峰值对应高度约240 hPa，相对误差大小最小，为21.7%；13—14 km分类使用了458个样本训练，258个样本测试，反演RMSE范围为0—0.0021 g/m³，在IWC廓线主要范围230—150 hPa，相对误差在50%以内，在廓线峰值约180 hPa处，相对误差最小，约为17%；14—15 km使用了235个样本训练，133个样本测试，反演RMSE范围为0—0.0038 g/m³，在IWC廓线主要范围190—130 hPa，相对误差在50%以内，在廓线峰值约160 hPa处，相对误差最小，约为22%。从不同云高的反演结果可以看出，神经网络每一类型廓线主要分布高度范围相对误差都在50%以内，峰值对应高度约为20%。但是无论对哪种高度的冰云进行反演时，云顶与云底的反演误差都较大。这是因为虽然按照云高对冰云廓线样本进行了预分类，但冰云厚度的差异依旧会对反演结果造成影响。

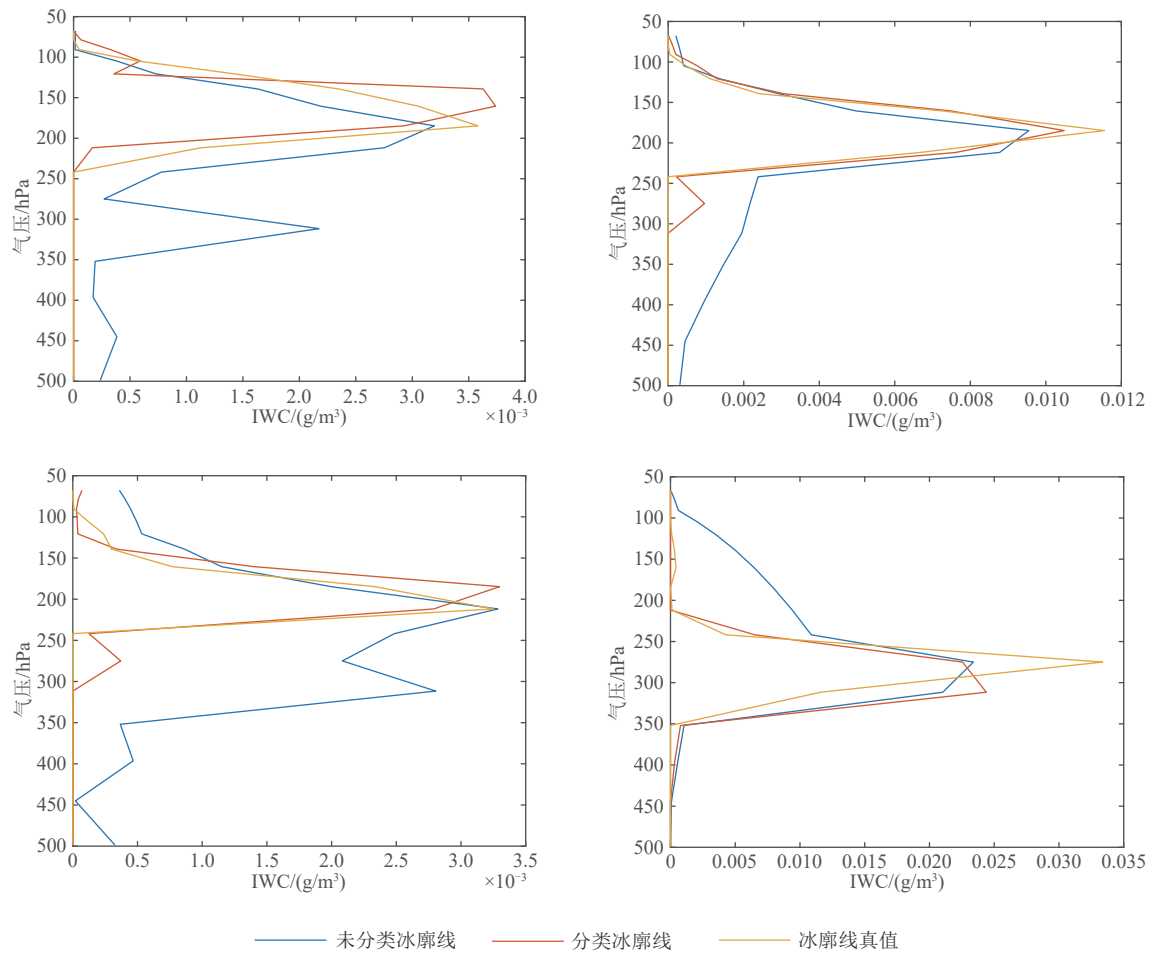


图 15 未分类与预分类神经网络的IWC反演廓线对比

Fig. 15 Comparison of retrieval profiles of IWC using unclassified neural network and pre-classified neural network

表 6 冰参数反演误差对比

Table 6 Comparison of the retrieval errors of ice parameters

变量	直接反演	间接反演
IWP RMSE/(g/m ²)	8.97	11.25
IWP error/%	19.8	24.9
I _{D_{me}} RMSE/μm	9.57	7.54
I _{D_{me}} error/%	11.5	9.9
I _{Z_{me}} RMSE/m	768.55	309.21
I _{Z_{me}} error/%	6.4	2.5

4.4 冰参数反演精度分析

与霰粒子类似，通过反演的冰密度廓线 IWC 也可以间接计算得到冰粒子参数。我们将相同冰云样本的间接反演冰参数误差与本文 4.2 节直接反演的结果进行对比。表 6 展示了两种情况下得到的 3 种冰参数的均方根误差 RMSE 和相对误差。从数值上来看，IWP 的间接反演 RMSE 从直接反演的 8.97 g/m² 上升到了 11.25 g/m²；而等效云高 I_{Z_{me}} 是间接反演误差更小，从 768.55 m 下降到 309.21 m，

降低了一半；等效粒径 I_{D_{me}} 也是间接反演的误差更小，RMSE 从 9.57 μm 下降到 7.54 μm。从这次试验结果来看，I_{D_{me}} 和 I_{Z_{me}} 采用间接反演方法的反演精度更高，IWP 采用神经网络直接反演算法精度更高。但是这一结论还需要更多样本测试来支持。

5 结 论

本论文基于 DOTLRT 辐射传输模式模拟星载太赫兹冰云探测亮温，使用神经网络算法开展了冰云参数反演仿真试验对冰云探测频率通道选择、冰云反演算法优化、冰霰参数和廓线的探测精度分析等太赫兹冰云探测的关键问题开展研究。本论文的创新之处在于探测和反演时将冰云粒子分为冰和霰两种类型粒子分开计算，突破现有研究仅仅计算冰云中单一冰粒子的局限，更加符合冰云真实情况。主要研究结论如下：

(1) 辐射特性表明冰、霰两种粒子的散射作用同时会对太赫兹频率亮温产生影响，同时反演

两种粒子参数模型太复杂, 因此必须利用两种粒子辐射特性的差异分开反演;

(2) 与 ATMS 真实观测亮温对比发现 WRF 模式预报的霰密度廓线 GWC 值准确性较差, 因此用 WRF 预报大气参数加上 ATMS 真实观测 GWC 构建的混合冰云数据库更接近真实冰云状态;

(3) 冰粒子对 183 GHz 频段亮温影响很小, 而该频段霰粒子对亮温散射效应明显, 因此可以使用 183 GHz 频段探测亮温单独反演冰云霰粒子参数和 GWC 廓线。而 GWC 廓线样本分布复杂, 本文基于 GWP 对样本预分类实现了较好精度的 GWC 神经网络反演算法; 对霰粒子参数, GWP 和 $G_{D_{me}}$ 采用间接反演算法精度更高, GWP 反演误差为 10.90 g/m^2 , $G_{D_{me}}$ 反演误差为 $25.38 \mu\text{m}$, $G_{Z_{me}}$ 采用直接反演算法反演精度更高, 反演误差为 513.62 m ;

(4) 以反演霰廓线为先验约束条件, 基于太赫兹频段冰云探测亮温差实现了对冰粒子参数和冰廓线 IWC 的神经网络反演算法。IWC 廓线样本分布复杂, 本文基于 $I_{Z_{me}}$ 对样本预分类实现了较好精度的 IWC 神经网络反演算法。对冰粒子参数, $I_{D_{me}}$ 和 $I_{Z_{me}}$ 间接反演精度更高, $I_{D_{me}}$ 反演误差为 $7.54 \mu\text{m}$, $I_{Z_{me}}$ 反演误差 309.21 m , 而 IWP 直接反演精度更高, 反演误差为 8.97 g/m^2 ;

(5) 频率通道对比试验表明, 更多的太赫兹探测频率通道能够实现更好的冰云参数探测精度。

志 谢 感谢美国科罗拉多大学博尔德分校 Al. Gasiewski 教授提供的 DOTLRT 模式。

参考文献 (References)

Boukabara S A, Garrett K, Chen W C, Iturbide-Sanchez F, Grassotti C, Kongoli C, Chen R Y, Liu Q H, Yan B H, Weng F Z, Ferraro R, Kleespies T J and Meng H. 2011. MiRS: an all-weather 1DVAR satellite data assimilation and retrieval system. *IEEE Transactions on Geoscience and Remote Sensing*, 49(9): 3249-3272. [DOI: 10.1109/TGRS.2011.2158438]

Brath M, Fox S, Eriksson P, Harlow R, Burgdorf M, Buehler S. 2018. Retrieval of an ice water path over the ocean from ISMAR and MARSS millimeter and submillimeter brightness temperatures. *Atmospheric Measurement Techniques*, 11(1): 611-632 [doi.org/10.5194/amt-11-611-2018]

Buehler S A, Jiménez C, Evans K F, Eriksson P, Rydberg B, Heymsfield A J, Stubenrauch C J, Lohmann U, Emde C, John V O, Sreerakha T R and Davis C P. 2007. A concept for a satellite mission to measure cloud ice water path, ice particle size, and cloud altitude. *Quarterly Journal of the Royal Meteorological Society*, 133(S2):

109-128 [DOI: 10.1002/qj.143]

Donovan D P. 2003. Ice-cloud effective particle size parameterization based on combined lidar, radar reflectivity, and mean Doppler velocity measurements. *Journal of Geophysical Research: Atmospheres*, 108(D18): 4573 [DOI: 10.1029/2003JD003469]

Evans K F, Wang J R, Racette P E, Heymsfield G and Li L H. 2005. Ice cloud retrievals and analysis with the compact scanning submillimeter imaging radiometer and the cloud radar system during CRYSTAL FACE. *Journal of Applied Meteorology and Climatology*, 44(6): 839-859 [DOI: 10.1175/JAM2250.1]

Evans K F, Wang J R, O' C Starr D, Heymsfield G, Li L, Tian L, Lawson R P, Heymsfield A J and Bansemmer A. 2012. Ice hydrometeor profile retrieval algorithm for high-frequency microwave radiometers: application to the CoSSIR instrument during TC4. *Atmospheric Measurement Techniques*, 5(9): 2277-2306 [DOI: 10.5194/amt-5-2277-2012]

Garrett T J, Navarro B C, Twohy C H, Jensen E J, Baumgardner D G, Bui P T, Gerber H, Herman R L, Heymsfield A J, Lawson P, Minnis P, Nguyen L, Poellot M, Pope S K, Valero F P J and Weinstock E M. 2005. Evolution of a Florida cirrus anvil. *Journal of the Atmospheric Sciences*, 62(7): 2352-2372 [DOI: 10.1175/JAS3495.1]

Heymsfield A J and Iaquinta J. 2000. Cirrus crystal terminal velocities. *Journal of the Atmospheric Sciences*, 57(7): 916-938 [DOI: 10.1175/1520-0469(2000)057<0916:CCTV>2.0.CO;2]

Jiménez C, Buehler S A, Rydberg B, Eriksson P and Evans K F. 2007. Performance simulations for a submillimeter-wave satellite instrument to measure cloud ice. *Quarterly Journal of the Royal Meteorological Society*, 133(S2): 129-149 [DOI: 10.1002/qj.134]

Kangas V, D'Addio S, Klein U, Loiselet M, Mason G, Orhlac J C, Gonzalez R, Bergada M, Brandt M and Thomas B. 2014. Ice cloud imager instrument for MetOp Second Generation//2014 13th Specialist Meeting on Microwave Radiometry and Remote Sensing of the Environment. Pasadena, CA: IEEE: 228-231 [DOI: 10.1109/MicroRad.2014.6878946]

Liou K N. 1986. Influence of cirrus clouds on ceather and climate processes: a global perspective. *Monthly Weather Review*, 114(6): 1167-1199 [DOI: 10.1175/1520-0493(1986)114<1167:IOCCOW>2.0.CO;2]

Liu L, Weng C S, Li S L, Hu S, Ye J, Dou F L, Shang J. 2020. Research status and progress of terahertz wave passive remote sensing of ice clouds. *Advances in Geoscience*, 35(12): 1211-1221 (刘磊, 翁陈思, 李书磊, 胡帅, 叶进, 窦芳丽, 商建. 2017. 太赫兹波被动遥感冰云研究现状及进展. *地球科学进展*, 35(12): 1211-1221) [DOI: 10.11867/j.issn.1001-8166.2020.103]

Mendrok J, Wu D L, Buehler S A, Jimenez C and Kasai Y. 2009. Submillimeter wave radiometer for observation of cloud ice: a proposal for Japanese mission//Proceedings Volume 7474, Sensors, Systems, and Next-Generation Satellites XIII. Berlin: SPIE: 249-257 [DOI: 10.1117/12.830704]

Piyush D N, Goyal J, Srinivasan J. 2017. Retrieval of cloud ice water path using SAPHIR on board Megha-Tropiques over the tropical ocean. *Advances in Geoscience*, 59(7): 1895-1906 [doi.org/10.1016/j.asr.2017.01.022]

Voronovich A G, Gasiewski A J and Weber B L. 2004. A fast multi-stream scattering-based Jacobian for microwave radiance assimilation. *IEEE Transactions on Geoscience and Remote Sensing*, 42(8): 1749-1761 [DOI: 10.1109/TGRS.2004.830637]

- Waliser D E, Li J L F, Woods C P, Austin R T, Bacmeister J, Chern J, Del Genio A, Jiang J H, Kuang Z M, Meng H, Minnis P, Platnick S, Rossow W B, Stephens G L, Sun-Mack S, Tao W K, Tompkins A M, Vane D G, Walker C and Wu D. 2009. Cloud ice: a climate model challenge with signs and expectations of progress. *Journal of Geophysical Research: Atmospheres*, 114(D8): D00A21 [DOI: 10.1029/2008JD010015]
- Wang D, Prigent C, Aires F, Jimenez C. 2016. A statistical retrieval of cloud parameters for the millimeter wave ice cloud imager on board MetOp-SG. *IEEE Access*, 5: 4057-4076 [doi: 10.1109/ACCESS.2016.2625742]
- Wang H, Duan C D, Lü R C, Lei H W, Zhu Z B and Chen G. 2017. Development of space borne Terahertz ice clouds measurement technology and existing technical problems. *Journal of Terahertz Science and Electronic Information Technology*, 15(5): 722-727 (王虎, 段崇棣, 吕容川, 雷红文, 朱忠博, 陈刚. 2017. 星载太赫兹冰云探测技术发展和面临问题. *太赫兹科学与电子信息学报*, 15(5): 722-727) [DOI: 10.11805/TKYDA201705.0722]
- Weng F, Zou X, Wang X, Yang S, Goldberg M. 2012. Introduction to Suomi national polar-orbiting partnership advanced technology microwave sounder for numerical weather prediction and tropical cyclone applications. *Journal of Geophysical Research Atmospheres*, 117(D19112): 1-14 [https://doi.org/10.1029/2012JD018144]
- Zhang R, Wang Y B. 2016. Research on machine learning and its algorithm and development, 23(2): 10-24 (张润, 王永滨. 2016. 机器学习及其算法和发展研究. *中国传媒大学学报自然科学版*, 23(2): 10-24) [DOI:10.16196/j.cnki.issn.1673-4793.2016.02.002]
- Zhang X, Hu W D, Liu R T, Si W K, Li Y D, Liu Y L and Lighthart L P. 2018. 874 GHz ice cloud detector design based on Cubesat platform. *Aerospace Shanghai*, 35(2): 144-150 (张欣, 胡伟东, 刘瑞婷, 司炜康, 李雅德, 刘莞喽, Lighthart L P. 2018. 基于 Cubesat 平台的 874 GHz 冰云探测仪设计. *上海航天*, 35(2): 144-150) [DOI: 10.19328/j.cnki.1006-1630.2018.02.018]

Research on retrieval algorithm of terahertz ice cloud sounding based on neural network

CHEN Ke¹, ZHANG Lan¹, ZHANG Youming², DONG Shanbin¹, LIU Yan¹, WU Qiong³, SHANG Jian³

1. Huazhong University of Science and Technology, Wuhan 430074, China;

2. Wuhan Maritime Communication Research Institute, Wuhan 430079, China;

3. Innovation Center for FengYun Meteorological Satellite (FYSIC), Key Laboratory of Radiometric Calibration and Validation for Environmental Satellites, National Satellite Meteorological Center (National Center for Space Weather), China Meteorological Administration, Beijing 100081, China

Abstract: Terahertz band has a number of potential advantages that complement existing visible and infrared techniques in ice cloud sounding application, but treating various phase ice particles (mainly ice and graupel) as single ice particles is a major limitation of current terahertz ice cloud retrieval algorithms. In this paper, a pre-classified neural network algorithm based on the terahertz radiation characteristics of ice cloud is proposed, which is able to retrieve the physical parameters of ice and graupel particles, respectively. The algorithm first uses a pre-classified neural network to retrieve the density profiles of graupel particles separately from the 183 GHz band brightness temperature data. The retrieved graupel profiles are then used as a priori constraint to calculate the higher frequency band brightness temperature difference due to ice particles only. Finally, another pre-classified neural network is used to retrieve the density profiles of ice particles separately from the preceding terahertz brightness temperature difference data. The proposed algorithm are evaluated through the end to end simulation experiments. Firstly, a hybrid ice cloud dataset including ice and graupel particle parameters is built based on the numerical weather prediction (NWP) model and the actual observation data. Then the synthetic ice cloud brightness temperature data from 183—874 GHz (i.e. 183 GHz, 243 GHz, 325 GHz, 448 GHz, 664 GHz and 874 GHz) are generated through Discrete-Ordinate Tangent Linear Radiative Transfer (DOTLRT) radiative transfer model with the hybrid ice cloud dataset. Finally, the parameters of ice and graupel are retrieved by the proposed algorithm from the simulated brightness temperature data, and compared with the input parameters to assess the retrieval accuracy. The simulation experiments show that the average Root Mean Square Errors (RMSE) of the retrieved IWP and GWP are 8.97 g/m² and 10.90 g/m² respectively, and the average RMSE of the retrieved I_{D_{mc}} and G_{D_{mc}} are 7.54 μm and 25.38 μm respectively, and the average RMSE of the retrieved I_{Z_{mc}} and G_{Z_{mc}} are 309.21 m and 513.62 m respectively, and the retrieved density profiles of ice and graupel particles also have high accuracy. The results indicate that the proposed algorithm can retrieve the total path amount, equivalent ice particle size, and equivalent ice cloud height and density profile of ice and graupel particles respectively with high accuracy, which is more consist with the real condition of ice cloud than the current retrieval algorithm.

Key words: terahertz, ice cloud sounding, neural network, ice and graupel particles, retrieval of ice cloud parameter

Supported by National Natural Science Foundation of China (No. 42275141); '13th Five-Year' Pre-research Project of Civil Aerospace of State Administration of Science of Technology and Industry for National Defense (SASTIND) (No. D040110); 2019 Shanghai Academy of Spaceflight Technology (SAST) Foundation (No. SAST2019-102)