

论文题目：

基于人工神经网络的大气红外探测仪 反演沙尘参数

作者姓名：	<u>孙祥龙</u>	入学时间：	<u>2013年9月</u>
专业名称：	<u>测绘工程</u>	研究方向：	<u>遥感科学技术与应用</u>
指导教师：	<u>许君一</u>	职 称：	<u>副 教 授</u>
副指导教师：	<u>李志楷</u>	职 称：	<u>高级工程师</u>

论文提交日期：2015年5月

论文答辩日期：2015年6月

授予学位日期：



**RETRIEVING DUST PARAMETERS FROM
ATMOSPHERIC INFRARED SOUNDER MEASUREMENTS:
AN ARTIFICIAL NEURAL NETWORK APPROACH**

A Dissertation submitted in fulfillment of the requirements of the degree of

MASTER OF ENGINEERING

from

Shandong University of Science and Technology

by

Sun Xianglong

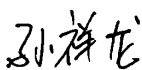
Supervisor: Associate Professor Xu Junyi

Geomatics College

May 2015

声 明

本人呈交给山东科技大学的这篇硕士学位论文，除了所列参考文献和世所公认的文献外，全部是本人在导师指导下的研究成果。该论文资料尚没有呈交于其它任何学术机关作鉴定。

硕士生签名 
日 期: 2015.06.09

AFFIRMATION

I declare that this dissertation, submitted in fulfillment of the requirements for the award of Master of Engineering in Shandong University of Science and Technology, is wholly my own work unless referenced of acknowledge. The document has not been submitted for qualification at any other academic institute.

Signature: 
Date: 2015.06.09

摘要

沙尘暴是我国西北和华北北部地区出现的强灾害性天气。近年来,沙尘暴天气频发,严重干扰着人们的正常生活,对社会经济和环境造成显著危害。基于遥感手段实现对沙尘暴实时、有效、大范围的监测已成为人们热切关注的话题。

为探究大气红外探测器(AIRS)在沙尘暴监测中的潜在应用,本文选择新疆塔克拉玛干沙漠作为研究区,利用 AIRS 测量值,借助人工神经网络方法反演沙尘相关的参数,得到沙尘暴空间和状态信息。主要完成以下工作:

1) 中分辨率成像光谱仪(MODIS)通过深蓝算法反演的气溶胶光学厚度(AOT)产品作为标准沙尘暴 AOT 参数,星载云气溶胶激光雷达(CALIPSO)反演的沙尘高度(Z_{dust})产品作为标准沙尘暴垂直高度参数,将 AIRS 中的数据集与标准沙尘参数进行相关性分析,选择与其相关性高的亮度温度、地表高度和卫星天顶角数据作为神经网络训练的样本数据和测试数据。

2) 对选取的 AIRS、MODIS 和 CALIPSO 数据集分别进行解析,并完成拼接、裁剪、投影转换、去条带效应等预处理。将预处理后的数据配准到同一地区、同一时间、同一坐标系下。

3) 选择 3 层 BP 神经网络,确定一种网络拓扑和训练参数,初始化网络权值。用配准后的样本数据进行训练,测试数据进行验证,反复迭代。通过误差分析与参数趋势分析确定神经网络的结构参数,使训练的网络在沙尘暴参数的反演中更加稳健。

4) 选择研究区沙尘爆发时期的 AIRS 测量值输入到训练完成的网络模型中,反演沙尘气溶胶光学厚度和沙尘高度,通过与标准沙尘参数进行相关性分析对反演结果进行评价,对塔克拉玛干地区沙尘的动态传输进行分析。

结果显示:训练后, AIRS 测量值反演的沙尘气溶胶光学厚度和沙尘高度与标准参数间相关系数分别为 0.88 和 0.79,表明利用 AIRS 测量值反演沙尘信息与 MODIS 深蓝算法反演的气溶胶光学厚度产品和 CALIPSO 反演的沙尘高度产品能达到相似的效果。基于红外高光谱测量得到全天时的沙尘参数为改善当地沙尘演变特征的监测和预警能力提供了巨大潜力。

关键词: 沙尘监测, 人工神经网络, 大气红外探测器, 气溶胶光学厚度, 沙尘高度

Abstract

Sandstorm is a strong disaster weather occurred in the northwest and north region of north China. In recent years, sandstorm weather occurred frequently, seriously interfere with people's normal life, causing significant harm to social economy and environment. Realizing the real-time, effective and large scale monitoring on sandstorm based on remote sensing method, has become a hot topic of people.

In order to explore the potential application of Atmospheric Infrared Sounder(AIRS) in sandstorm monitoring, this paper chooses Taklimakan Desert in Xinjiang as the study area to inverse parameters related to dust, uses the value measured by AIRS with the method of artificial neural network to get the space and state information of sandstorm. To this end, the main work needs to be done include the following:

1) Put the aerosol optical thickness (AOT) derived from Moderate Resolution Imaging Spectroradiometer (MODIS) using the algorithm of dark blue and the dust height(Z_{dust}) derived from Cloud Aerosol Lidar Infrared Pathfinder Satellite Observation (CALIPSO) as the standard AOT and vertical height parameters of dust storm, Respectively. Correlation analysis was carried out on the data sets of AIRS with the standard sand parameters. Choose the brightness temperatures, surface elevation and relative air mass that have high correlation to dust parameters as the sample data and test data for neural network training.

2) Parse the selected AIRS, MODIS and CALIPSO data sets, and complete the mosaic, cutting, projection transformation and removing the effect of the strip preprocessing. Match the data sets that after pretreatment to the same area, same time and same coordinate system.

3) Choose a 3 layer BP neural network, a kind of network topologies and training parameters, initialize the network weights. We use the sample data after registration for training, and test data for validation iteratively. Determine the structure parameters of the neural network through the error analysis and parameter trend analysis, to make the trained network more robust in a sandstorm parameter inversion.

4) Input the values from AIRS measurement in dust outbreak period to the training network model to retrieval the aerosol optical thickness and the height of the dust. Evaluate the inversion results through the correlation analysis to the standard sand parameters and

analyze the dynamic transmission of dust in Taklimakan desert.

Results showing that: after training, the correlation coefficient of 0.88 and 0.81 for AOT and Z_{dust} derived from AIRS measurement from standard parameters, respectively. Suggest that dust retrievals from hyperspectral IR sounder measurements are comparable to MODIS-derived AOT based on deep blue and CALIPSO-measured Z_{dust} . The retrievals of parameters related to dust during both day and night from IR hyperspectral measurements may offer great potential to improve our ability to monitor and forecast the evolving features of local dust.

Keywords: Dust Monitoring, ANN, AIRS, AOT, dust height

目录

1 绪论1

1.1 研究背景及意义 1

1.2 国内外研究现状 2

1.3 研究内容与技术路线 4

1.4 本章小节 6

2 研究区数据获取7

2.1 研究区概况 7

2.2 数据获取 9

2.3 本章小节 25

3 构建训练数据集 26

3.1 数据集的选取 26

3.2 数据集相关性分析 28

3.3 本章小节 32

4 建立人工神经网络模型 33

4.1 人工神经网络概述 33

4.2 网络模型的选择 37

4.3 网络层数的确定 37

4.4 隐层节点数的选取 38

4.5 激活函数的选定 39

4.6 神经网络模型的验证分析 41

4.7 本章小节 43

5 基于人工神经网络反演沙尘参数..... 44

5.1 2013 年 4 月 4 日案例 44

5.2 2013 年 4 月 7 日案例 48

5.3 本章小节 52

6 结论与展望..... 53

6.1 结论	53
6.2 展望	54
致谢	55
参考文献	56

Contents

1 Introduction 1

1.1 The research background and significance 1

1.2 The research status at home and abroad 2

1.3 The research contents and technical route 4

1.4 Summary of the chapter 6

2 The data obtained in study area 7

2.1 The general situation of the study area 7

2.2 Data acquisition 9

2.3 Summary of the chapter 25

3 Construction of the training datasets 26

3.1 Select datasets 26

3.2 Correlation analysis of the datasets 28

3.3 Summary of the chapter 32

4 Construct an artificial neural network model 33

4.1 Overview of artificial neural network 33

4.2 Select the network model 37

4.3 To determine the network layer 37

4.4 Select the number of hidden layer nodes 38

4.5 Select an activation function 39

4.6 Analysis and validation of the neural network model 41

4.7 Summary of the chapter 43

5 Dust parameters inversion based on artificial neural network 44

5.1 The April 4, 2013 case 44

5.2 The April 7, 2013 case 48

5.3 Summary of the chapter 52

6 Conclusion and Prospect 53

6.1 Conclusion 53

6.2 Prospect 54

Acknowledgements	55
Reference	56

1 绪论

1.1 研究背景及意义

沙尘是大气气溶胶的主要成分之一，对气候和环境有重要影响。沙尘粒子一方面通过直接散射及吸收太阳辐射和吸收及发射红外辐射影响地气辐射能量平衡^[1,2]，另一方面，可以作为云的凝结核，通过改变云滴和冰晶粒子的大小等微物理特性以及云的生命周期，从而间接影响气候系统的辐射能量收支^[3]。

沙尘暴天气是我国西北地区和华北北部地区出现的强灾害性天气。它能加剧土地沙漠化，对生态环境造成巨大破坏，对大气环境造成严重污染，对交通和供电线路产生重要影响，给人民生命财产造成严重损失^[4]。

近年来，源自中蒙地区的沙尘在东亚地区大范围传播，使得我国沙尘暴、扬尘和浮尘天气频繁发生，并对人类生活、健康、活动产生重大影响，因而受到越来越多人的关注，也加强了人们对于沙尘暴预报的研究，但由于沙尘暴预报的复杂性，一直未能取得比较满意的效果。近年来，受东蒙古地区干旱范围扩张的影响^[5,6]，韩国和日本遭受沙尘暴的袭击也越来越频繁^[7]。为了使用早期预警系统对沙尘暴作出较为准确的预报，减少沙尘暴对人类健康和社会经济带来的损失，对于亚洲地区沙尘的监测以及沙尘爆发后的灾害性评估就变得尤为重要。

沙尘暴的监测方法中，传统的地面监测方法受到许多因素的制约，不能很好地刻画沙尘暴过程。我国气象工作者针对沙尘暴展开过一些研究工作，但大多侧重于对产生沙尘暴天气的天气气候学特征进行研究和分析，并总结归纳其大气环流特征，或者通过对个别沙尘暴的特殊成因进行分析，找出沙尘暴预报指标。客观的、定量的、自动化的沙尘暴预报方法还很少，特别是针对我国大范围的预报就更少。遥感卫星基于大面积的同步观测和对于沙尘光学信息定量测量的能力可以实现对沙尘状态的高效观测。利用多种遥感数据监测沙尘暴，提取沙尘暴信息，定量分析沙尘暴的有关参数，已成为沙尘暴研究的热点课题。其中，紫外和可见光遥感技术在沙尘暴研究中发挥着越来越重要的作用^[8-11]，但由于受波段条件的限制，只能反演白天的数据。时间覆盖度的增加会改善沙尘监测的能力，因此，我们假设夜晚的观测数据会有助于沙尘信息的反演。然而，紫外和可见光波段在夜晚无法工作，所以为了增加时间覆盖度尤其是夜晚的覆盖，我们选择红外波段来进行夜晚的观测。

基于红外波段的遥感卫星探测沙尘信息已经取得一定成效。例如, 热分裂窗通道间的亮度温度(BT)差异已经可以用来作为指数来表示沙尘的严重程度, 尽管对于沙尘监测的定量化表示还有一定的限制。但值得一提的是: 如果能够在夜晚进行观测的话, 那么定量化的物理参数会比沙尘指数更有意义。而且, 如果能够从夜晚红外波段探测的数据中反演出红外气溶胶光学厚度的话, 依据亚洲沙尘地区米氏散射计算公式中的光谱变化消光系数就能够将红外气溶胶光学厚度转换为我们比较熟悉的可见光的气溶胶光学厚度。

因为红外波谱段上千个通道的数据包含了地表和大气红外辐射与沙尘辐射之间相互作用的信息, 使得高光谱红外探测应用于沙尘反演变得更加有效。实际上, 近期的研究已经表明: 高光谱分辨率的红外探测仪(比如大气红外探测仪和大气红外干涉仪)可以用来反演沙尘高度和红外气溶胶光学厚度^[12, 13, 14]。不过, 虽然高光谱红外探测仪在沙尘属性的反演上取得一定成效, 目前的研究工作也仅局限在海洋属性的反演或案例研究上, 针对陆地上空沙尘参数的反演工作比较少, 也比较困难。因此, 很有必要深入研究整体反演的精度或者利用其数据信息对于沙尘的监测和预警。

新疆塔克拉玛干沙漠作为中国最大的沙尘源, 其长距离传输对全球大气的气溶胶成分以及气候变化都有重要影响。本文希望通过对塔克拉玛干地区的沙尘参数进行研究, 能够对沙尘暴的预报以及环境改善有所贡献。

1.2 国内外研究现状

1.2.1 国外研究现状

国外从 20 世纪 20 年代开始就对沙尘暴的时空分布、成因机制以及监测与对策方面进行了研究: 20 年代初期, Hankin^[15]对发生在印度的“*Andhi*”沙尘暴进行了初步研究; 之后, Sutton^[16]对“*Haboob*”型沙尘暴也进行了分析, 结果表明哈布型沙尘暴是季风内出现的对流性沙尘暴; Joseph^[17]把发生在印度西北部地区的沙尘暴等级进行了划分, 认为沙尘暴等级应该分为以下 3 类: 即 4 级<风速≤6 级, 500m≤能见度<1000m, 称为弱沙尘暴; 6 级<风速≤8 级, 200m≤能见度<500m, 称为中等强度的沙尘暴; 风速≥9 级, 能见度<200m 等级范畴内, 又划分出了特强沙尘暴, 即 50m≤能见度<200m 时, 称为强沙尘暴, 当瞬时最大风速≥25m·s⁻¹, 能见度<50m, 甚至降到 0m 时, 称为特强沙尘暴(俗称“黑风”); 80 年代以来, Brazel^[18]对发生在美国亚利桑那州的沙尘暴天气类型作了统计分类: 把当

地形成沙尘暴的天气形势归结为四个天气类型：锋面型、雷暴和对流型、热带扰动型和上部低压或切断低压型；Jauregui^[19]对墨西哥城沙尘暴的时空分布进行系统研究后指出3月份沙尘暴发生频率最大；Gillette^[20]研究了美国沙尘暴的时空分布规律认为沙尘暴是由于风蚀引起的，并且通过对粒子质量通量与地面摩擦速度的轨迹分析得出：到达日本上空的沙尘主要源于40~50° N、100~110° E的中国沙漠戈壁、黄土高原和内蒙古地区；Bergam^[21]的研究表明：在全球范围内，沙尘暴主要的源地在北非的撒哈拉沙漠和中亚的戈壁；Tegen^[22]在全球模式中模拟了撒哈拉地区沙尘的起源、传输和分布；Ott.S.T.和AOtt(1991)^[23, 24, 25]曾利用卫星和GATE资料对撒哈拉沙尘暴的爆发进行过分析，立足从大风和气候两个方面研究沙尘暴的影响因素；Uemastu^[26]模拟了亚洲沙尘的跨太平洋传输；Shaw^[27]在夏威夷岛的Mauna Loa观测站，发现有来自亚洲的沙尘气溶胶个例；HSU^[8, 9]对于干旱半干旱地区的沙漠进行了研究，利用蓝光波段地物反射率低的特性通过建立查找表的方法提出了反演沙尘气溶胶光学厚度的深蓝算法。此外，Nilgun^[28]还分析了沙尘气溶胶的成分和含量。

1.2.2 国内研究现状

国内从20世纪70年代开始对沙尘暴天气进行研究。1993年9月在兰州召开“首届全国沙尘暴天气研讨会”。此后，沙尘暴研究成为中国科学界关注的重要问题。1997年，在中国气象局科技教育司和中国科学院自然与社会协调发展局的安排下，以1993年“5.5”黑风暴分析研究为基础，方宗义、朱福康等编写了《中国沙尘暴研究》。2000年，北京的沙尘暴天气频繁发生后，沙尘暴再次引起我国政府和学者的关注。王式功^[29]和全林生^[30]等研究了我国沙尘暴天气的时空分布特征、形成原因和发展趋势；胡隐樵等^[31]对沙尘暴形成的动力条件进行了初步探讨和分析；纪飞等^[32]从数值模拟和数值预报的角度，模拟了沙尘暴发生、发展的机制和演变结构；叶笃正^[33]等对2000年发生在我国华北地区沙尘天气的成因进行分析，并提出了相应的治理对策和措施；牛若芸、薛建军、周自江^[34]对2002年我国沙尘暴天气特征进行了分析，结果表明：2002年出现的12次沙尘暴过程中，以3月18~22日的强沙尘暴过程影响范围最广、强度最强；吕新苗、刘惠清、王文杰^[35]，对北京周边地区沙尘暴时空分布特征及其环境背景进行研究，结果表明：沙尘暴高值区位于浑善达克沙地和库布齐沙地的周边。从时间变化看，午后至傍晚是沙尘暴天气易发的高峰期；从季节变化的总体上看，春季较多，其次是夏季和初秋；王宁、张红^[36]对吉林省沙尘天气的时空分布特征及其分型预报进行了研究，通过普查历史天气图，

给出了区域性沙尘天气及区域性浮尘天气的预报指标,并在 2002 年春季沙尘预报服务中得到较好的应用;赵翠光^[4]使用人工神经网络方法建立了我国沙尘暴短期预测模型,该神经网络模型的输入因子是几个物理场的 REOF 展开的一些时间系数,输出为我国有无沙尘暴。结果表明 REOF 展开技术和人工神经网络方法两种方法的结合对于预测沙尘暴是可行的;牛生杰等^[37]对宁夏贺兰山地区沙尘暴发生频次进行了分析,结果表明:春季,尤其是 4,5 月份是该地区沙尘暴的高发期;李正明^[38]利用临夏州 1955~2004 年 50 年的气候资料,从地形、气候、环流特征等方面入手,全面系统地分析了临夏州大风、沙尘天气气候学成因,在此基础上,建立了临夏州春季大风沙尘天气短期实时预报系统;路志英,杨玉峰等^[39]针对现有 BP 神经网络的缺陷,提出了基于遗传算法的神经网络预测模型,并利用该模型研究了沙尘暴的预测问题,仿真预测研究表明:提出的方法无论是收敛速度还是预报准确率方面都取得了比较满意的结果;我国学者围绕沙尘暴问题还展开了更为广泛的研究,限于文章篇幅这里不再详述。

1.3 研究内容与技术路线

1.3.1 研究内容

本论文的研究内容及工作主要包括以下几个方面:

1) 资料收集与整理

在进行研究之前,首先明确沙尘暴的危害以及关于沙尘暴监测的国内外研究进展,对比分析各种沙尘暴监测方法的优缺点,论述新疆塔克拉玛干地区地形地貌以及气候特点等,提出人工神经网络的监测方法,并对 MODIS、CALIPSO 提取沙尘信息的数据集产品进行研究,搜集关于其反演算法的资料。

2) 利用 MODIS 数据反演沙尘气溶胶参数算法研究

简述气溶胶光学厚度遥感监测的原理,对比分析利用 MODIS 数据进行气溶胶光学厚度反演的传统算法,选择深蓝算法(deep blue)作为研究对象反演研究区的沙尘气溶胶光学厚度,并简述其原理和具体的实现流程。

3) 利用 CALIPSO 数据反演沙尘参数

利用 CALIPSO 数据依据测得的退偏振比区分沙尘与其他类型气溶胶,通过接收大气中不同层的后向散射,计算出沙尘气溶胶的垂直高度信息,并简要介绍其实现流程。

4) 选择特征数据集

为实现人工神经网络的训练,有效监测沙尘参数信息,选择合适的数据集产品进行训练和验证必不可少。这一部分论述了如何从 AIRS、MODIS、CALIPSO 数据中选择有效产品,并对 AIRS 的数据集与 MODIS 气溶胶光学厚度产品和 CALIPSO 沙尘高度产品进行对比和相关性分析。

5) 数据预处理及配准。对选取的数据集分别进行解析,并完成拼接、裁剪、投影转换、去条带效应等预处理。将预处理后的数据配准到同一地区、同一时间、同一坐标系下。

6) 建立 BP 神经网络沙尘暴预报模型

分析研究塔克拉玛干地区地理状况、气候状况以及沙尘分布特点,选取适当的 MODIS、CALIPSO、AIRS 训练数据集,经过网络模型的选择、参数的选取以及模型的训练和验证等工作,得出适合该地区沙尘暴预测的神经网络模型。

7) 对人工神经网络模型的反演结果进行评价

通过训练完成的人工神经网络模型反演的研究区沙尘参数结果分别与 MODIS、CALIPSO 反演的沙尘气溶胶数据进行对比分析,并得出适用性评价,并对塔克拉玛干沙漠沙尘的动态传输进行分析。

1.3.2 研究方法

由于本文是基于 MODIS 以及 CALIPSO 反演的数据集产品来与 AIRS 中的相关数据集一起训练构造人工神经网络模型的,因此,高品质的 MODIS、CALIPSO 数据反演算法提取沙尘信息就变得尤为重要,否则,基于训练成功的 AIRS 数据反演出的沙尘参数毫无意义。所以,本文首先针对 MODIS、CALIPSO 两种数据的反演算法进行研究,然后分析并提取出 AIRS 数据中与大气沙尘相关的数据集,提出三层 BP 人工神经网络进行训练。神经网络建模过程是一个反复训练和测试的过程,其目的是为了不断优化网络权值并选择更好的模型以达到最好的预测效果。在构造神经网络模型之前,首先应用 SPSS 统计软件对原始资料进行统计分析,并利用 SURFER 软件进行空间特征分析。建模过程中先选定一种训练集合和测试集合,确定一种网络拓扑、迭代次数、动量项、学习速率等训练参数作为初始模型。在训练完成之后,用测试样本集合测试模型的反演效果,如果结果不好,则重新选定网络拓扑、训练参数和迭代次数,再进行训练、测试和反复试探,直到测试结果满意为止。

1.3.3 技术路线

具体的技术路线和实现流程如下图所示：

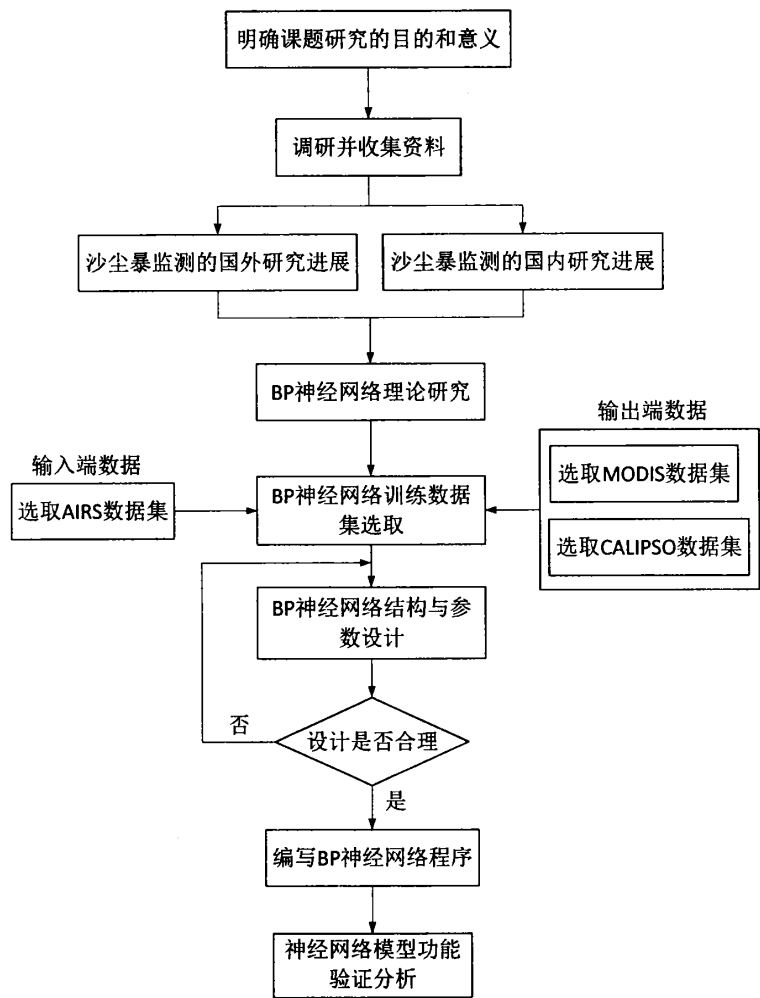


图 1.1 论文技术路线图
Fig. 1.1 The technical route of dissertation

1.4 本章小节

本章阐明了本论文的研究背景及意义，综合论述了国内外的研究进展状况，在此基础上提出人工神经网络的沙尘暴监测方法，结合研究内容制定了具体的实施方案和研究计划，并不断进行改进和修正。

2 研究区数据获取

2.1 研究区概况

塔克拉玛干沙漠位于新疆塔里木盆地中心（见图 2.1），东西长约 1000 千米，南北宽 400 千米，总面积为 337600 平方千米，是中国面积最大的沙漠，也是世界第二大流动沙漠。位于 $36.26^{\circ} \sim 42.10^{\circ} \text{ N}$ 、 $74.88^{\circ} \sim 90.00^{\circ} \text{ E}$ 的区域。地处欧亚大陆中心，四面为高山所环绕，地形闭塞，降水稀少，气候极端干旱。是全球同纬度最干旱的地区之一，也是生态环境严重敏感和脆弱的地区。

研究分析在全球变化背景下塔克拉玛干沙漠及周边地区的沙尘传输及状态变化，对于保护和改善当地的生态环境，促进当地社会经济的持续稳定发展具有重要意义。

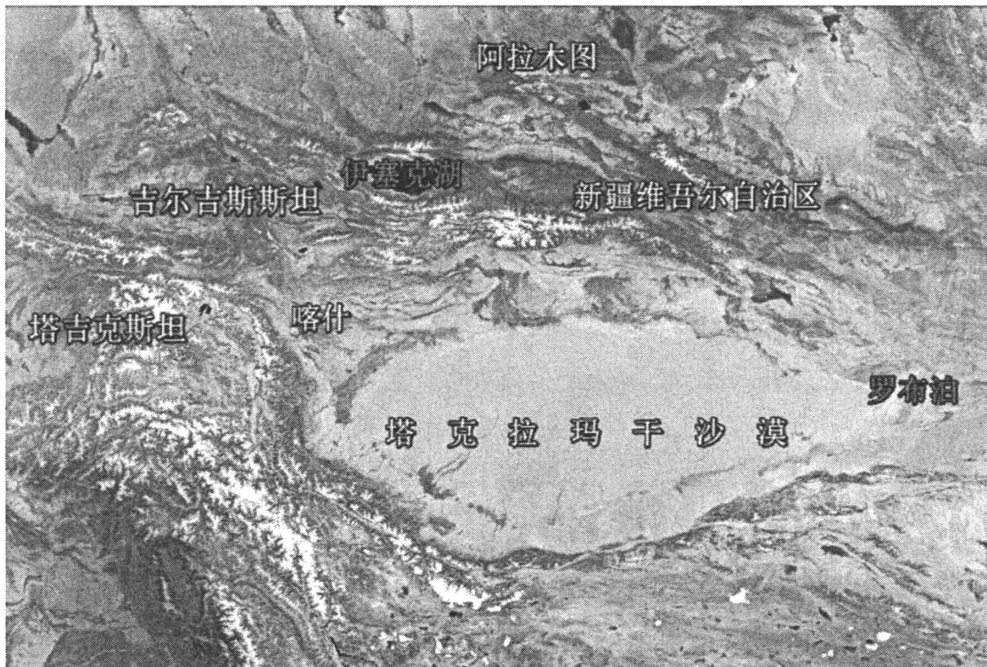


图 2.1 塔克拉玛干沙漠卫星视图

Fig. 2.1 Satellite view of the Taklimakan Desert

2.1.1 地形地貌

塔克拉玛干沙漠西部和南部海拔高达 1200~1500 米，东部和北部则为 800~1000 米，沙丘最高达 200 米。也有资料显示，塔克拉玛干沙漠，是全世界第一大流动沙漠。塔克拉玛干腹地被评为中国五个最美的沙漠之一。由于地处欧亚大陆的中心，四面为高

山环绕，塔克拉玛干沙漠充满了奇幻和神秘的色彩。

塔克拉玛干沙漠流动沙丘的面积很大，沙丘高度一般在 100 ~ 200 米，最高达 300 米左右。沙丘类型复杂多样。

塔克拉玛干沙漠的侧翼为雄伟的山脉：天山在北面，昆仑山在南面，帕米尔高原在西面。东面逐渐过渡，直到罗布泊沼盆。在南面和西面，在沙漠和山脉之间，则是由卵石碎屑沉积物构成的一片坡形沙漠低地。

塔克拉玛干沙漠的地表是由几百米厚的松散冲积物形成的。这一冲积层受到风的影响，其为风所移动的沙盖厚度达 300 米。风形成的地形特征多种多样，各种形状与大小的沙丘均可见到。在沙漠的东部和中部，以中间凹陷的沙丘和巨大、复杂的沙丘链形成的网为主。

2.1.2 气候

塔克拉玛干沙漠，系暖温带干旱沙漠，酷暑最高温度达 67.2°C ，昼夜温差达 40°C 以上；平均年降水不超过 100 毫米，最低只有四五毫米；而平均蒸发量则高达 2500 ~ 3400 毫米。主要特征可总结为：冷热剧变、降水量少而集中，变率大和风沙频繁等。

西部地区夏季盛行北风和西北风。这两种气流在位于克里雅河最北端的沙漠中心附近相会后，造成复杂的环流系统，这一点清楚地反映在沙丘的形状上。春季，当地表沙变暖时，上升气流发展，东北风特别强烈。在此期间，强飓风尘暴常常发生，使大气充满沙尘，可高达海拔 3962 米。从其他方向来的风也将尘雾扬入空中，几乎终年笼罩着塔克拉玛干沙漠。

白天，塔克拉玛干赤日炎炎，银沙刺眼，沙面温度有时高达 70~80 度，旺盛的蒸发，使地表景物飘忽不定。下垫面性质既有沙漠和戈壁的影响，又有一定数量的小盆地和地形的微起伏的作用，形成更为复杂的局地气候特色。

南疆的太阳总辐射可达 140~150 千卡/厘米².年，散射辐射为 70~80 千卡/厘米².年，大于直接辐射，原因在于沙漠地区多浮尘，明显地反映了其沙漠气候特征。在沙漠下垫面的平均反射率为约为 30%。有效辐射值较大，约为 60 千卡/厘米².年，反映了本区干旱、高温和云量不多的基本事实。沙漠地区乱流热交换的年平均比蒸发耗热年平均大 10 倍之多，在暖季，下垫面所获得热量的 80%是以乱流热交换的方式用于空气增温。因此，沙漠夏季特别炎热。

2.2 数据获取

本文采用人工神经网络的方法，利用 MODIS 反演的气溶胶光学厚度和 CALIPSO 卫星反演的沙尘高度数据作为沙尘参数样本进行训练，得到神经网络模型。因此，由 MODIS 和 CALIPSO 数据对于沙尘参数准确有效地反演，得到沙尘参数样本数据，对于神经网络模型的构建非常重要。这也是决定网络模型反演精度的重要影响因素。

本节将对 MODIS 反演气溶胶光学厚度和 CALIPSO 反演沙尘高度等参数的理论进行研究，对 AIRS 传感器的特点以及数据获取方法进行论述，为后续提取和选择训练数据集提供理论依据和数据支持。

2.2.1 MODIS 数据获取

2.2.1.1 MODIS 数据简介

MODIS 全称 Moderate-Resolution Imaging Spectroradiometer，即中分辨率成像光谱仪。是搭载在 TERRA 和 AQUA 卫星上的一个重要传感器，也是卫星上唯一将实时观测数据通过 x 波段向全世界直接广播，并可以免费接收数据并无偿使用的星载仪器，全球许多国家和地区都在接收和使用 MODIS 数据。

MODIS 是当今世界上新一代“图谱合一”的光学遥感仪器，属被动式成像分光辐射计，携带 490 个探测器，分布在 36 个光谱波段（其中第 13 和 14 波段各由两个通道组成，分为高值通道和低值通道。所以实际上一共有 38 个光谱通道），覆盖从可见光到红外（0.4~14μm）波段，其主要特性和用途见下表：

表 2.1 MODIS 数据技术指标
Table 2.1 Technical indicators of MODIS

项目	指标
轨道	705km，降轨上午 10:30 过境，升轨下午 1:30 过境，太阳同步，近极地圆轨道
扫描频率	每分钟 20.3 转，与轨道垂直
测绘宽带	2330km × 10km
望远镜	直径 17.78cm
体积	1.0m × 1.6m × 1.0m
重量	250g
功耗	225w
数据率	11Mbit/s
量化	12bit
星下点空间分辨率	250m,500m,1000m
设计寿命	5 年

表 2.2 MODIS 仪器特性和主要用途
Table 2.2 Band feature and characters of MODIS instrument

通道	波谱范围 (nm)	信噪比 NEΔT	主要用途	分辨率 (m)
1	620~670	128	陆地、云边界	250
2	841~876	201		250
3	459~479	243	陆地、云特征	500
4	545~565	228		500
5	1230~1250	74		500
6	1628~1652	275		500
7	2105~2135	110		500
8	405~420	880	海洋水色、浮游植物、生物地 理、化学	1000
9	438~448	8380		1000
10	483~493	802		1000
11	526~536	754		1000
12	546~556	754		1000
13	662~672	910		1000
14	673~683	1087		1000
15	743~753	586		1000
16	862~877	516		1000
17	890~920	167	大气水汽	1000
18	931~941	57		1000
19	915~965	250		1000
20	3.660~3.840	0.05	地球表面和云顶温度	1000
21	3.929~3.989	2.00		1000
22	3.929~3.989	0.07		1000
23	4.020~4.080	0.07	同上	1000
24	4.443~4.498	0.25	大气温度	1000
25	4.482~4.549	0.25		1000
26	1.360~1.390	1504	卷云、水汽	1000
27	6.535~6.895	0.25		1000
28	7.175~7.475	0.25		1000
29	8.400~8.700	0.05		1000
30	9.580~9.880	0.25	臭氧	1000
31	10.780~11.280	0.05	地球表面和云顶温度	1000
32	11.770~12.270	0.05		1000
33	13.185~13.485	0.25	云顶温度	1000
34	13.485~13.785	0.25		1000
35	13.785~14.085	0.25		1000
36	14.085~14.385	0.35		1000

MODIS 仪器的地面分辨率分别为 250m、500m 和 1000m，视幅宽度为 2330km，在对地观测过程中，每秒可同时获得 6.1 兆比特来自大气、海洋和陆地表面信息，每日或每两日可获取一次全球观测数据。多波段数据可以同时提供反映陆地、云边界、云特性、海洋水色、浮游植物、生物地理、大气中水汽、地表、云顶温度、大气温度、臭氧和云顶高度等特征的信息，用于对陆表、生物圈、固态地球、大气和海洋进行长期全球观测。为了提高数据精度，MODIS 系统的传感器除发射前进行定标外，还在飞行过程中进行定标。校正系统共有 4 个部件，其中包括黑体（black body-BB）、阳光漫射器（solar diffuser-SD）、阳光漫射稳定性监测器（solar diffuser stability monitor-SDSM）和光谱辐射校正器（spectra radio metric calibration assembly）。黑体用于对热红外波段（3.5~14 μm ）校准；阳光漫射器用于对可见光、近红外和短波红外波段（0.4~2.2 μm ）等太阳反射波段的校正；阳光漫射稳定性监测器用于检测阳光漫射器本身的稳定性；光谱辐射校正器是一个非常复杂的多功能仪器，用于提供在轨光谱、辐射和空间校正。这就使 MODIS 卫星数据相对于其它卫星数据有了非常高的可靠性、连续性，最大限度地减小了由于系统定标而引起的误差。每一个 MODIS 仪器的设计寿命为 5 年，1998 年到 2006 年将计划发射 4 颗卫星。由此估计，利用 MODIS 仪器至少将获得 15 年 36 个光谱波段的地球综合信息，这些数据对于开展自然灾害与生态环境监测、全球环境和气候变化研究以及进行全球变化的综合性研究等将是非常有意义的。MODIS 探测器有助于我们从整体上理解地球，提高预测将来变化以及区分人类活动和自然活动对环境的影响的能力，其目标是构建全球动力模型，包括大气、海洋和陆地，在地球发生变化之前就能预测到。

按数据产品特征划分：主要产品包括校正数据产品、陆地数据产品、海洋数据产品和大气数据产品；若按处理级别划分，又可以分为以下 6 种：

0 级产品：也称原始数据；

1 级产品：指 L1A 数据，已经被赋予定标参数；

2 级产品：指 L1B 级数据，经过定标定位后数据，本系统产品是国际标准的 EOS-HDF 格式。包含所有波段数据，可能是应用比较广泛的一类数据；

3 级产品：在 1B 数据的基础上，对由遥感器成像过程产生的边缘畸变（Bowtie 效应）进行校正，产生 L3 级产品；

4 级产品：由参数文件提供的参数，对图像进行几何纠正，辐射校正，使图像的每一点都有精确的地理编码、反射率和辐射率。L4 级产品的 MODIS 图像进行不同时相的匹配时，误差小于 1 个像元。该级产品是应用级产品不可缺少的基础；

5 级及以上产品：根据各种应用模型开发 L5 级产品。

2.2.1.2 MODIS 数据反演沙尘气溶胶参数

利用卫星观测数据反演气溶胶光学厚度主要利用太阳反射波段的数据，以气溶胶粒子对太阳辐射的散射为物理机制实现气溶胶信息获取。为了实现沙尘气溶胶光学厚度反演，需要选择气溶胶散射强，地表反射率低的光谱通道，且卫星观测通道设计上要求波段宽度窄、信噪比高。

卫星遥感反演陆地气溶胶光学厚度(AOT)主要基于可见光和近中红外波段，其难点在于气溶胶模式的确定和地表反射率贡献的去除。在沙漠或戈壁区，由于较高的地表反射率，对大气散射造成干扰，给气溶胶参数的反演带来很大困难。由于卫星接收到的上层辐射（表观反射）一方面来源于太阳光未到达地表直接被大气散射的贡献值（大气路径辐射值），另一方面来源于地表反射的贡献值，因此，卫星探测的大气与地表间相对辐射亮度值决定了气溶胶参数反演的效果。

下图所示为气溶胶单次散射反照率 ω_0 分别为 0.91、0.96 和 1.0 时蓝光波段(490nm)在不同地表反射率下大气上层的表观反射率。黑色实线表示无气溶胶条件下的表观反射率， $\tau_a = 0$ 。黑色、绿色和红色虚线分别表示 $\tau_a = 1.0$ 时在不同单次散射反照率下的表观反射率；竖直虚线表示地表反射率的临界值。从图中可以看出：在亮地区，沙尘爆发区与非爆发区（对比黑色虚线与黑色实线）得到的表观反射率之间的差异并不明显。这表明，在亮地区，卫星探测的表观反射率对于气溶胶厚度的变化并不敏感。另一方面，图中绿色虚线（ $\omega_0 = 0.96$ ）和红色虚线（ $\omega_0 = 0.91$ ）与黑实线的对比表明：气溶胶对于太阳光的散射作用，会增大暗地区的总反射率或减小高反射地区的总反射率。在地表反射的临界值处（竖直虚线表示），由于气溶胶的吸收作用，使得总的反射值未发生变化^[40]。在这些临界值附近反演气溶胶参数会造成很大的误差。先前，有一些算法，用大于 600nm 的波长来反演气溶胶，在亮地区（高地表反射区）会存在很大的困难。

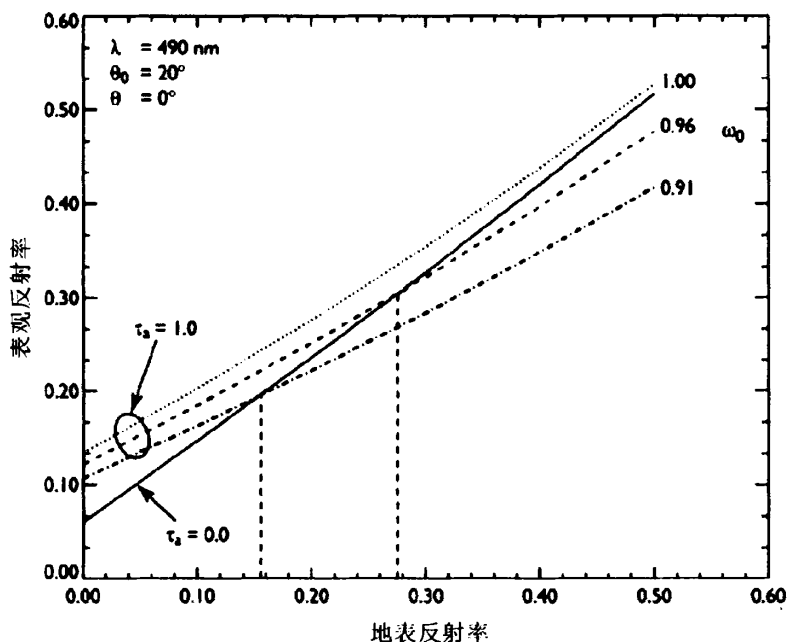


图 2.2 在不同的气溶胶光学厚度 τ_a 和单次散射反照率 ω_0 下, 蓝光波段(490nm)探测的大气顶部的表观辐射率随地表反射率变化的曲线

Fig. 2.2 Simulated apparent reflectance at the top of the atmosphere at 490 nm, as a function of surface reflectance for various values of the aerosol optical thickness τ_a and single-scattering albedo ω_0

由于地表辐射的影响, 在干旱、半干旱的沙漠或戈壁区探测沙尘气溶胶光学厚度存在一系列问题, 很多学者也针对这些问题进行了大量的探索和研究, 并取得实质性进展。在如何去除地表反射率贡献方面, 目前应用比较广泛的有暗目标法、结构函数法以及深蓝算法等。不同的算法在不同的应用领域有不同的优缺点, 下面就这些算法的原理以及优缺点进行简要介绍。

1) 暗目标法

陆地上的稠密植被、湿土壤及水体覆盖区在可见光波段反射率很低, 在卫星图像上称为暗像元。模拟及观测研究表明, 在晴空无云的暗像元上空, 卫星观测反射率随大气气溶胶光学厚度单调增加, 利用这种关系反演大气气溶胶光学厚度的算法, 称为暗像元方法, 它是Kaufman和Sendra^[41]在反演稠密植被上空气溶胶光学厚度时建立的。暗像元方法利用大多数陆表在红(0.60~0.68 μm)和蓝(0.40~0.48 μm)波段反射率低的特性, 根据植被指数(NDVI)或短波红外通道(2.13 μm 和3.8 μm)观测值进行暗像元判识, 并依据一定的关系假定这些暗像元在可见光红或蓝通道的地表反射率, 用于反演气溶胶光学厚度。暗像元算法基于表观反射率的大气贡献项, 即利用卫星观测的路径辐射反演气

溶胶光学厚度。它是目前陆地上空气溶胶遥感应用最为广泛的算法。但这种方法只对植被覆盖度高或土壤含水量高的暗区域比较有效,在干旱、半干旱的沙漠或戈壁区,由于地表反射率比较高,对于大气散射的干扰大,由此给亮地区沙尘气溶胶参数的反演造成许多困难,业务运行上会出现许多的不确定性。

2) 结构函数法

针对亮地表的陆地气溶胶反演,Tanré等假定一定范围内复杂地表的大气透过率相同提出了结构函数法^[42],这是早期研究陆地污染气溶胶采用的卫星遥感算法,该算法假设同一个地区一段时间内地表反射率是不变的,利用“清洁日”大气作为参考,反演“污染日”大气的气溶胶光学厚度。利用结构函数法可以反演城市地区的气溶胶分布状况。该算法利用 TM、AVHRR、SPOT 等数据也取得了较好效果^[42, 43],但由于对几何校正要求较高,业务化运行比较困难。

3) 深蓝算法

Hsu 等人^[8, 9]根据在红光和蓝光波段 AOT 对天顶辐亮度有显著的贡献提出了基于地表反射率库的深蓝(Deep-Blue)算法,利用 SeaWIFS 图像建立了地表反射率库,在 AOT 较小时仅使用蓝光数据进行反演,AOT 较大时则综合使用红光和蓝光数据进行反演,此方法已成功应用于撒哈拉沙漠、阿拉伯半岛等干旱、半干旱地区。本文选择塔克拉玛干地区来反演沙尘气溶胶光学厚度,基于研究区高的地表辐射值,因此选择深蓝算法来进行后续的研究。

(1) 深蓝算法原理

为了反演亮地区的气溶胶参数,我们探究偏振辐射传输模型来计算反射强度场,如下所示:

$$R_{(\mu,\mu_0,\varphi)} = \frac{\pi I(\mu,\mu_0,\varphi)}{\mu_0 F_0} \quad (2.1)$$

式中: R 代表卫星接收到的辐射(或表观反射); F_0 为太阳辐射通量; I 为大气顶层辐射; μ 为观测天顶角余弦; μ_0 为太阳天顶角余弦; φ 为散射辐射方向与太阳入射方向间的方位角。

在公式(2.1)中,我们假设下垫面为均匀的朗伯面,因此,大气顶层总的辐射值可以以地表反射率方程的形式表示为:

$$R_{(\mu,\mu_0,\varphi)} = R_{0(\mu,\mu_0,\varphi)} + \frac{T A_s}{1 - s A_s} \quad (2.2)$$

式中, $R_{0(\mu,\mu_0,\varphi)}$ 表示观测方向的路径辐射项; T 为传输方程,描述大气顶层与底部间

辐射值的变化； A_s 表示朗伯体表面辐射率； s 为大气整层向下的半球反射率。基于此公式，就可以在不同的观测几何(太阳天顶角、观测天顶角、相对方位角)和卫星观测的辐射值之间建立查找表，计算这些参数之间的关系，去除地表贡献，实现地气解耦，得到所需的大气参数。

陆地气溶胶光学厚度反演，首先需要解决地表反射率和气溶胶模型这两个未知参数的确定问题。为了从表观反射率反演气溶胶光学厚度，需要合理假定气溶胶模型，以提供单次散射反照率 ω_0 值和气溶胶相函数 P 。在确定了气溶胶模型和地表反射率后，根据公式(2)，可从表观反射率反演得到气溶胶光学厚度。

(2) 技术路线

基于深蓝算法来获取沙尘气溶胶光学厚度等参数的具体实现流程如下所示：

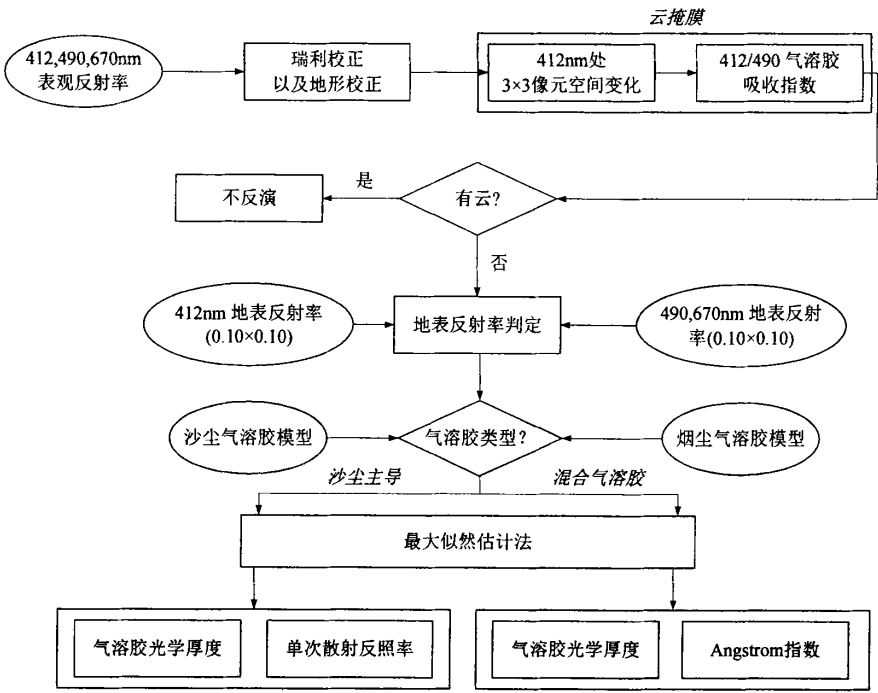


图 2.3 深蓝算法流程图

Fig. 2.3 Flow chart of the Deep Blue algorithm

上图所示为深蓝算法的组织结构和具体实现流程。首先，将测得的地表反射率进行瑞利散射和吸收的校正以及地形变化的校正去除地表气压变化的影响。分别建立 1atm 地表气压和 0.4atm 地表气压的查找表，计算真实地表高程和观测视角海平面之间表面压力的瑞利散射值，基于瑞利散射的差异来校正卫星像元地表辐射率。然后，在反演气溶胶

之前,通过计算 3×3 像元范围内的空间一致性和气溶胶吸收指数 (AI) 值筛选出有云的子像素,并将其去除。这是因为当前在有云污染的区域反演气溶胶由于其复杂性,还没有取得实质性的进展。在无云场景的查找表数据库中依据地理位置来判定给定像元的地表反射率。在 MODIS 第 412nm、470nm 和 650nm 通道测量的反射率与由太阳天顶角、观测天顶角、相对方位角、地表反射率、气溶胶光学厚度和单次散射反照率构成的查找表中的辐射值进行对比。基于最大似然估计法对沙尘和烟尘模式进行线性拟合,计算混合比,直到计算的光谱反射率与实测值达到最佳一致时停止。不断重复此过程,获取不同气溶胶以及地表条件的最佳信息。对于混合气溶胶,一旦气溶胶模式和最佳拟合的混合比被确定下来,气溶胶光学厚度值和 Angstrom 指数也将被计算出来。在沙尘气溶胶占主导地位的情况下,单次散射反照率值也能被计算出来。

反演的沙尘气溶胶光学厚度将作为人工神经网络模型训练的输出样本和验证样本对 AIRS 中的数据集(人工神经网络模型的输入样本)进行训练和验证,这将在后面的章节中进行论述。

2.2.2 CALIPSO 数据获取

2.2.2.1 CALIPSO 数据简介

CALIPSO 卫星由装备垂直偏振通道的云层-气溶胶激光雷达(CALIOP)、宽视角相机(WFC)和红外成像辐射仪(IIR)三部分组成。CALIPSO 数据是由美国航天局(NASA)兰利研究中心(Langley Research Center, LaRC)大气科学数据中心(ASDC)处理、存储及发布的。这些数据通过分级数据格式(HDF)输出。CALIPSO 数据主要由两级数据组成,其中第一级数据提供了卫星激光雷达测量的基本数据和系统的校准信息等,具有很高的准确度和可信度,包括 3 个雷达通道(532nm 的垂直和平行通道以及 1064 nm 的红外通道)的衰减后向散射系数 ABS 序列(Attenuated backscattering coefficient profile)以及其不确定度。命名方式为: CAL_LID_L1-Prov-V2-01.2008-07-17T01-08-00ZN.hdf。其意义分别为: CALIPSO_激光雷达_第一级数据.年-月-日-代号.hdf^[44,45]。其中 ABS 是一个二维矩阵,其横坐标为垂直高度,纵坐标为地理坐标。通过确定 CALIPSO 数据中距离目标地点最近的经纬度坐标,就可以把所需的相对应参数导出。第一级数据的垂直和水平空间分辨率如下表所示。可见对于 532nm 的 CALIPSO 信号,根据垂直分辨率,在海拔-2~40.1km 的高度范围内一共有 583 个高度坐标对应有 583 个样本。由于陆地海拔最低的死海海拔为海平面下 392m,因此全分辨率区域扩展至-0.5km,而垂直分辨率为 300m 的-2.0~-0.5km

区域是为了回复监视表面返回的潜在延迟^[45]。在海拔-2~40.1km 的高度范围内，可以根据研究需要选择适当的垂直高度范围进行目标的研究分析。

表 2.3 CALIPSO 数据空间分辨率
Table 2.3 Spatial resolution of CALIPSO

海拔/km	水平分辨率/km	垂直分辨率（532nm）/m	垂直分辨率（1064nm）/m
30.1/40.1	5.0	300	---
20.2/30.1	1.67	180	180
8.2/20.2	1.0	60	60
-0.5/8.2	0.33	30	60
-2.0/-0.5	0.33	300	300

2.2.2.2 CALIPSO 数据反演沙尘参数

针对气溶胶的观测，地基观测数据网数据精度较高，然而空间局限性大；传统被动遥感数据，如 MODIS、MISR、TOMS 等，可以实现大范围空间覆盖，并对沙尘气溶胶的水平分布和传输有较好的观测，却无法提供气溶胶在垂直层面上的分布信息；由于沙尘气溶胶的垂直分布对大气的垂直加热进而对大气热力结构的影响和改变^[46]、对气溶胶-云相互作用扮演着重要角色^[47]以及决定沙尘气溶胶通过不同云微物理过程中的间接效应，使得沙尘垂直分布的研究变得十分必要。以往一些工作通过飞机实验和地面激光雷达等手段获取了一些沙尘气溶胶的垂直廓线，但这种方式能够涵盖的时空范围依然非常有限。而星载激光雷达监测系统具有长时间大范围连续观测云和气溶胶的优势，在对气溶胶和云垂直方向参数的反演和研究中发挥了重要作用。

2006 年 4 月发射的 CALIPSO 卫星是对地观测系统(Earth Observing System, EOS)的 A-Train 卫星群的成员之一，由美国 NASA 的兰利研究中心和法国国家空间研究中心(CNES)联合研制。它所携带的双偏振激光雷达 CALIOP（Cloud-Aerosol Lidar with Orthogonal Polarization），能远距离探测各种地形、光亮地表、薄云和晴空条件下气溶胶的垂直分布，提供的颗粒物退偏振信息可以区分冰云和水云，还可以辨识非球形气溶胶粒子。旨在提高人们对气溶胶和云在气候系统中的作用的认知水平。后期发展了一系列算法用以识别气溶胶与云层，反演垂直高度和粒子的物理学特性。

对于沙尘气溶胶在垂直结构和传输的研究主要围绕以下几个方面来进行：首先，通过 CALIPSO 卫星接收的大气中不同层的后向散射，准确测量云和气溶胶的垂直分布状况；其次，依据 CALIPSO 测得的退偏比数据，并结合气溶胶层的高度、地理位置、消光后向散射比等信息区分沙尘与其他类型的气溶胶。最后，消光系数混合反演算法利用

每种气溶胶的雷达比反演气溶胶的消光廓线，再通过积分得到 532nm 波段的 AOD 值。

1) 原理

对于利用 CALIPSO 数据来计算沙尘气溶胶垂直和平行衰减后向散射系数、退偏比、衰减颜色比等可用如下的公式来表示：

首先，CALIPSO 激光雷达方程可以表示为：

$$X(d) = E(\xi)\beta(d)T^2(d)/d^2 \quad (2.3)$$

式中， $X(d)$ 表示接收到的信号； E 为平均激光能量； d 为激光雷达到达测量点的距离； β 为 d 范围内的后向散射系数； ξ 为激光雷达系统参数； T 为从激光雷达到测量点的单程透射率，它与光学厚度有关，并符合比尔朗勃定律，而光学厚度又可以表示为：

$$T(d) = \exp\left(-\int_{d_0}^d \alpha(d')d_d'\right) \quad (2.4)$$

式中， α 为激光雷达到任意高度 d' 的大气单位体积消光系数，进行能量、增益和距离归一化后，可以将激光雷达简化为：

$$X(r) = C\beta(r)T^2(r) \quad (2.5)$$

在计算双向透射率时，大气中激光散射受空气中气溶胶粒子和分子两部分的影响，所以总的后向散射系数是二者后向散射系数的加和，即：

$$\beta(r) = \beta_a(r) + \beta_m(r) \quad (2.6)$$

而 CALIPSO 原始数据中包含 532nm 通道平行分量、532nm 通道垂直分量，将各自进行归一化可得：

$$X_{532,par}(r) = C_{532,par}(r)\beta'_{532,par}(r) \quad (2.7)$$

$$X_{532,per}(r) = C_{532,per}(r)\beta'_{532,per}(r) \quad (2.8)$$

$$X_{1064}(r) = C_{1064}(r)\beta'_{1064}(r) \quad (2.9)$$

式中， $\beta'_{532,par}(r)$ 、 $\beta'_{532,per}(r)$ 、 $\beta'_{1064}(r)$ 分别为 532nm 平行衰减后向散射系数和垂直衰减后向散射系数以及 1064nm 总的衰减后向散射系数。

后向散射系数反映了大气组分对太阳光的散射能力，是 CALIPSO 产品中重要的剖面数据，也是计算各层气溶胶垂直高度分布状况的基础，分别为：

$$\beta'_{532,par}(r) = \beta_{532,par}(r)T_{532}^2(r) \quad (2.10)$$

$$\beta'_{532,per}(r) = \beta_{532,per}(r)T_{532}^2(r) \quad (2.11)$$

$$\beta'_{532,total}(r) = \beta'_{532,par}(r) + \beta'_{532,per}(r) \tag{2.12}$$

$$\beta'_{1064}(r) = \beta_{1064}(r)T_{1064}^2(r) \tag{2.13}$$

进一步地，对 532nm 垂直衰减后向散射系数与 532nm 平行衰减后向散射系数进行比值运算，就得到体退偏比，它表示被测颗粒物的不规则程度，其数值越大表明颗粒物形状越不规则。体退偏比是区分沙尘与其它气溶胶的基础，在此基础上我们才能在沙漠等的干旱半干旱地区提取出沙尘信息，并进行后续的处理。体退偏比的求解公式表示为：

$$VDR(r) = \beta'_{532,per}(r)/\beta'_{532,par}(r) \tag{2.14}$$

对 1064nm 总衰减后向散射系数与 532nm 总衰减后向散射系数进行比值运算，则得到颜色比，它反映了被测颗粒物的大小，数值越大表示粒径越大。颜色比也是区分沙尘与其他气溶胶非常重要的参数，其求解公式表示为：

$$CR(r) = \beta'_{1064}(r)/\beta'_{532,total}(r) \tag{2.15}$$

2) 技术路线

CALIOP 提供 Level 1B、level 2 和 level 3 三级产品。本文采用 Level 1B 产品中版本为 3.01 的数据和 Level 2 产品中分辨率为 5 千米的气溶胶层数据以及垂直特征数据。

表 2.4 CALIOP 激光雷达数据参数
Table 2.4 The data parameter of CALIOP

数据名称	参数	数据集
Level 1B	经度	Longitude
	纬度	Latitude
	532nm 总衰减后向散射系数	Total_Attenuated_Backscatter_532
	532nm 垂直分量衰减后向散射系数	Perpendicular_Attenuated_Backscatter_532
	1064nm 总衰减后向散射系数	Total_Attenuated_Backscatter_1064
Level 2 Alay	经度	Longitude
	纬度	Latitude
	气溶胶光学厚度	Column_Optical_Depth_Aerosols_532
	气溶胶层数	Number_Layers_Found
	气溶胶层顶高度	Layer_Top_Altitude
	气溶胶层底高度	Layer_Base_Altitude
	气溶胶层体退偏比	Integrated_Volume_Depolarization_Ratio
	气溶胶层衰减色比	Integrated_Attenuated_Total_Color_Ratio
	气溶胶层 532nm 衰减后向散射系数	Integrated_Attenuated_Backscatter_532
	特征标识符	Feature_Classification_Flags
Level 2 VFM	经度	Longitude
	纬度	Latitude
	特征层标识符	Feature_Classification_Mask

Level 2 VFM 垂直特征数据产品主要用于描述由 CALIPSO 卫星所观测到的气溶胶层/云层的垂直分布和水平分布。

本文针对所用到的 CALIOP 产品其处理过程进行简要论述：

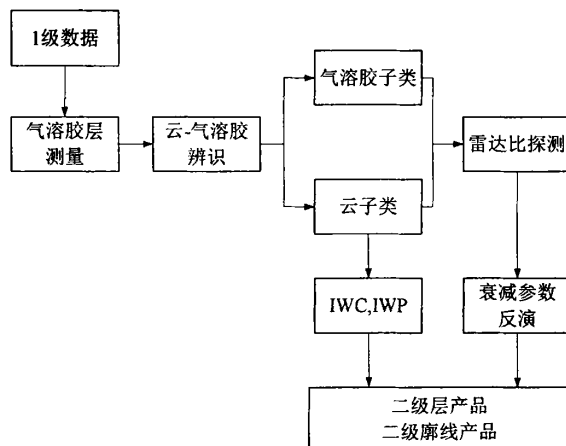


图 2.4 CALIOP Level 2 数据处理流程

Fig. 2.4 The data processing flow char of CALIOP Level 2

首先，依据测得的后向散射信号，区分探测特征层以及地表边界。然后，利用云-气溶胶判别算法（CAD）将特征层具体分为云、气溶胶以及平流层大气。此算法是一个多维的概率函数，由 532nm 平均衰减后向散射、1064/532 平均衰减后散射比（色比）、中间高度层、体积极化率和纬度构成，经归一化处理后，负值代表气溶胶，正值代表云。CAD 的绝对值越大，表示划分的可信度就越高。被判定为气溶胶的数据又综合考虑气溶胶层的退偏振比、后散射系数、色比和地理位置等信息，将气溶胶细分为沙漠沙尘、污染沙尘、烟尘、洁净陆地气溶胶、洁净海洋气溶胶和污染陆地气溶胶 6 个类型（Level 2 5km 气溶胶层数据可以为我们提供这六种气溶胶类型），同时对每种气溶胶给定雷达比。最后，基于消光系数混合反演算法，利用每种气溶胶的雷达比反演气溶胶的消光廓线，再通过积分得到 532nm 和 1064nm 波段的 AOD 值。

CALIPSO 的 Level2 产品数据由 Level 1 数据采用相关算法计算而得，由于算法的不确定性，使得 Level2 中的气溶胶光学厚度、消光系数、后向散射系数等参数存在一定误差。因此，在利用 CALIPSO 数据时需开展质量控制以便减少误差，本文进行数据质量控制参数设置见下表：

表 2.5 数据质量控制参数设置

Table 2.5 The parameter setting of data quality control

Level2 气溶胶廓线数据	Level2 层产品
Extinction_Coefficient_532≠-9999 及 Extinction Coefficient_532_Uncertainty≤10	Feature_Optical_Depth_53≠-9999 及 Feature_Optical_Depth_532_Uncertainty≤0.1
Atmospheric volume description(AVD)第 1-3 位值为 3,且第 10-12 位值不为 0	Feature Classification Flags (FCF)第 1-3 位值 为 3,且第 10-12 位值不为 0
-100< CAD_SCORE< -50	-100< CAD_SCORE< -50
Extinction_QC_532 = 0 或 Extinction_QC_532 = 1	Extinction_QC_532 = 0 或 Extinction_QC_532 = 1

其中，Level2 气溶胶廓线数据中 Extinction_Coefficient≠-9999 用以去除结果中无效值；设定 Extinction Coefficient_Uncertainty≤10 表示获取的数据不确定性较小；Atmospheric volume description 为 16 位整型数，该参数包含了产品数据中特征层的类别信息（云、气溶胶及无信号情况）及判别情况的置信度水平，当 Atmospheric volume description 第 1-3 位值设置为 3 时表示当前获取的数据集含气溶胶相关信息，而第 10-12 位值的设置情况代表了当前数据集中含的气溶胶子类型信息，主要分为 7 种情况（分别为清洁海洋型、沙尘、污染大陆型、清洁大陆型、污染沙尘、烟尘、无法判别、其他情况等），将 10-12 位值设为 0 可以排除当前数据子类型无法判别的情况；Level2 层产品层产品中 Feature_Optical_Depth_532 的设置同于 Extinction_Coefficient，而 FCF 则为该产品中包含探测到的特征层信息的参数，其设置方法也类似于 AVD；CAD_SCORE 一般变化在[-100,100]之间，用来表示区分云与气溶胶数据集结果的可信程度，且该值绝对值越大表示获取的数据集判别结果越可靠。CAD_SCORE 为正值时代表结果数据集为云的信息，而为负值时则代表结果数据集为气溶胶的信息。设置-100<CAD_SCORES<50 一方面保留了大量有效数据，另一方面也确保了结果数据的准确性。Extinction_QC_532 是一个反映消光系数反演过程中激光雷达比(Sa，消光系数与散射系数之比)的预设和变化情况的参数，通常，该值设为 0 或 1 时，反演结果最准确。

2.2.3 AIRS 数据获取

大气红外探测仪（AIRS，Atmospheric Infrared Sounder），是根据美国宇航局 EOS 计划，与中分辨率成像光谱仪（MODIS）、先进微波扫描辐射计（AMSR-E）、先进微波探

测器（AMSU）、云和地球辐射能量系统（CERES）、微波湿度探测器（HSB）等一同安装于 EOS PM-1 卫星（下午星，即 EOS Aqua，2002 年 5 月 4 日发射）上的新型红外遥感仪，资料开始于 2002 年 8 月。Aqua 上其它仪器及其主要应用见下表。大气红外探测仪提供了前所未有的更新、更准确的陆地、海洋和大气红外光谱数据，主要应用于探测大气湿度，以推动全球气候研究和天气预报的进展。其最大的特色即在于拥有 2378 个连续红外光谱通道（3.7-15.4 μm ），其所提供的高光谱精度大气温湿度资料及云、地表、臭氧等参数资料在美国各科研及业务机构、大学、公司已被广泛应用于全球气候研究和天气预报。它极大地提高了对流层温度廓线测量准确度及大气湿度测量准确性，使 1 公里对流层温度准确率达到 1K，在晴空或部分云覆盖条件下大气湿度廓线准确率达到 10%，同时使地表温度反演的平均准确度达到 0.5K。随 Aqua 卫星一日过境两次，AIRS 可提供地上每一点白天/夜间的大气三维结构数据。

表 2.6 Aqua 遥感器技术特性
Table 2.6 The technical characteristics of remote sensor Aqua

型号	名称	制造者	技术特性	主要应用
MODIS	中分辨率成像光谱仪	NASA	0.405~14.385 μm ,36 通道	大气/海洋/陆地监测
AIRS	大气红外探测器	NASA	0.4~15.5 μm ,2382 通道	大气温湿度
AMSR-E	先进微波扫描辐射计	NASDA	6.9~89GHz,12 通道	水汽、降水、温湿度
AMSU	先进微波探测器	NASA	23.8~89GHz,15 通道	水汽、地表温度
CERES	云和地球辐射能量系统	TRW	0.3~5.0 μm ,8~12 μm ,0.3~100 μm	地球反射/发射辐射
HSB	微波湿度探测器	巴西	1 个通道 150GHz + 3 个通道 180GHz	大气湿度廓线

AIRS 仪器的功率、重量、数据率分别为 225.5 瓦、140.6 公斤、1.44 兆/秒（15.5GB/日）。设计寿命五年。其所搭载的 Terra 卫星由美国 Lockheed Martin 公司研制。

AIRS/AMSU/HSB 是目前最先进的一套从红外到微波波段探测大气垂直廓线的观测仪，用于测量大气温度，提供大气的水汽分布信息，云、海/陆地表面温度及大气湿度数据，是作为气候研究的 EOS 系统中的重要组成部分，以获取大气、陆地、海洋目标的多光谱、高光谱数据。通过 AIRS/AMSU/HSB 我们可以得到^[48]：

- （1）海表面温度和大气温度廓线；

- (2) 水汽廓线及总体的可沉降水汽量;
- (3) 红外地表比辐射率和陆面温度;
- (4) 云顶气压、云覆盖率、云顶温度、云的红外辐射率;
- (5) 微量气体密度以及微量气体可能出现的高度;
- (6) 臭氧廓线及总体臭氧柱密度^[48]。

表 2.7 大气红外探测仪 AIRS 仪器基本参数
Table 2.7 The basic parameters of the atmospheric infrared sounder AIRS instrument

扫描宽度: 1650 公里
水平分辨率: 2.3 公里 (可见光/近红外通道) 1.35 公里 (红外通道)
垂直分辨率: 1 公里 (红外通道)
波段数: 2382 (可见光: 4; 红外: 2378)
可见光谱覆盖范围: (0.4~1.0) μm 红外光谱覆盖范围: (3.74~4.61) μm (6.20~8.22) μm (8.80~15.4) μm
光谱分辨率: $\lambda/\Delta\lambda>1200$ (标称)
灵敏度 (NEΔT): 4.2μm 时 0.14K 3.7~13.6μm 时 0.20K 13.6~15.4μm 时 0.35K
扫描角天底±49.5 度, 视场 (FOV) 圆形 1.1 度

AIRS L2 级产品以 EOS-HDF 格式存放, 每 6min 生成一个数据文件, 每天需处理经过中国区域的 2 轨 6 个文件。AIRS-L2 标准产品文件包含有地面气温及 28 层的温、湿度廓线等, 地面气温存放在 30×45 大小的数据集中, 温、湿度廓线存放在 30×45×28 大小的数据集中。AIRS 仪器采用连续扫描的方式, 完成一次横向跨轨扫描需 2.67 秒, 为方便处理, 把每 6 分钟的观测范围定义为一个景(即 granule), 一天 24 小时含 240 个景, 每个景由 135 条扫描线组成, 每条扫描线有 90 个观测视场(footprint), 每个视场包含 2378 个光谱取样, 约有 120 个景是白天观测得到, 而另外 120 个景是晚上观测得到。

图 2.5 为 AIRS 红外光谱示意图, 表 2.8 是 AIRS 的通道特征, 包括硬件模块标识、频率范围、通道范围、光谱分辨功率、灵敏度、通道数、主要探测目的等。从中我们可以看到, AIRS 在整个波段包含了 CO₂ 的强吸收带(15.5μm 和 4.3μm), 主要用来探测大气温度; 水汽的吸收带(6.3μm) 主要用来探测云和大气湿度的特性; CO 吸收带(4.67μm)、N₂O 吸收带(4.5μm 和 7.8μm)、CH₄ 吸收带(7.7μm) 和 O₃ 吸收带(9.6μm), 可以探测相应的臭氧和其他微量气体; 另外, 还包含一些窗区的波段。

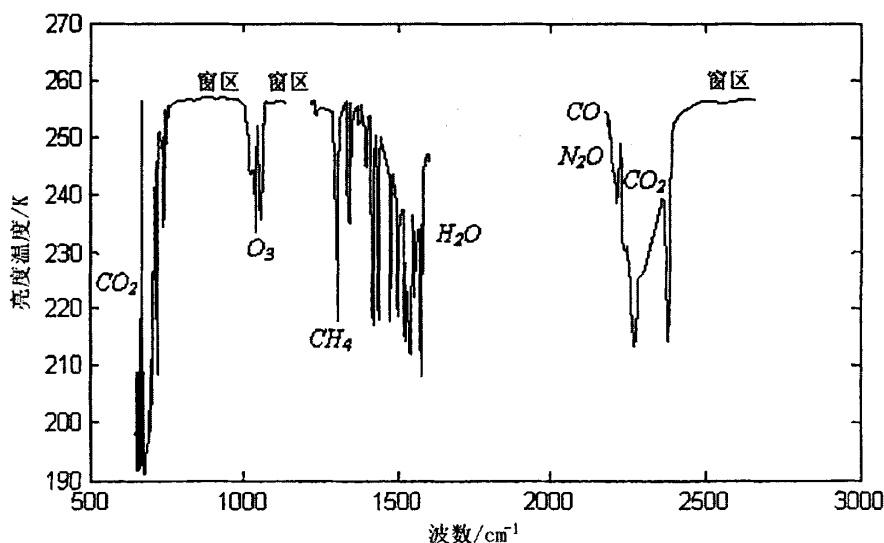


图 2.5 AIRS 红外光谱示意图
Fig. 2.5 Schematic diagram of Infrared Spectrum

表 2.8 AIRS 通道特征
Table 2.8 Channel characteristics of AIRS

模块	探测范围 (cm ⁻¹)	通道号范围	光谱分辨率 (cm ⁻¹)	灵敏度 (K)	通道数目	探测内容
M-12	649.61-681.99	1 - 130	0.41-0.46	0.321-0.697	130	温度
M-11	687.60-728.44	131 - 274	0.47-0.53	0.211-0.365	143	温度
M-10	728.06-781.88	275 - 441	0.55-0.63	0.219-1.113	166	表面
M-09	789.26-852.43	442 - 608	0.65-0.76	0.246-1.285	166	表面
M-08	851.49-903.78	609 - 769	0.57-0.64	0.125-0.544	160	表面
M-07	911.24-974.29	770 - 936	0.65-0.75	0.081-1.512	166	表面
M-06	973.82-1046.20	937 - 1103	0.76-0.88	0.072-0.266	166	臭氧
M-05	1056.10-1136.66	1104 - 1262	0.90-1.04	0.108-0.869	158	地表、臭氧
M-04d	1216.97-1272.59	1263 - 1368	0.95-1.04	0.076-0.33	105	水汽
M-04c	1284.35-1338.86	1369 - 1462	1.06-1.16	0.101-0.278	93	水汽、甲烷
M-03	1338.16-1443.07	1463 - 1654	0.95-1.11	0.072-0.148	191	水汽
M-04b	1460.27-1527.00	1655 - 1760	1.15-1.25	0.071-0.22	105	水汽
M-04a	1541.10-1613.86	1761 - 1864	1.28-1.41	0.120-0.47	103	水汽
M-02b	2181.50-2325.06	1865 - 2014	1.75-1.98	0.086-0.248	149	温度、CO
M-01b	2299.80-2422.85	2015 - 2144	1.74-1.94	0.13-0.243	129	温度
M-02a	2446.20-2569.75	2145 - 2260	1.98-2.18	0.177-0.58	115	表面
M-01a	2541.90-2665.24	2261 - 2378	1.94-2.13	0.302-1.026	117	表面

2.3 本章小节

本章综合论述了塔克拉玛干地区的地理概况。从地形地貌和气候两个方面阐明了研究区的地理状况。简要介绍了 MODIS、CALIPSO、AIRS 数据,论述了 MODIS 和 CALIPSO 反演沙尘气溶胶参数的算法研究以及具体的实现流程。针对 MODIS 反演沙尘气溶胶光学厚度的暗像元法、结构函数法和深蓝算法进行对比分析,选择深蓝算法来进行干旱半干旱的亮地区气溶胶光学厚度的反演。对基于 CALIPSO 激光雷达方程推导出的后向散射消光系数、体退偏比以及色比来得到沙尘各层的高度,并将沙尘粒子与其它气溶胶粒子区分开来。为后续的研究工作提供了理论依据和数据准备。

3 构建训练数据集

本文我们采用人工神经网络的方法来进行后续的研究工作，它通过训练可以支持对任意函数的学习，由此可以建立一种模型来描述这种非线性关系^[49, 50]。由于神经网络中记忆的规律，是从已知样本和已知变量组成的数据集中学习训练得到的。所以很明显，只有当已知样本的数据可靠时，学习到的规律才可能是正确的；只有当选取的样本在表达的规律中具有代表性时，才能得到对预报有用的数学模型。所以，对于人工神经网络来讲，输入、输出层节点的选取是影响神经网络训练精度的重要一步，如果选取的节点参数与我们所要探求的沙尘属性参数之间相关性不高，那么训练出的网络模型得到的结果必然不会有很高的精度，或者说，基于这些节点训练出的神经网络模型并不能生产我们所需的产品，那么，接下来的工作也将毫无意义。因此，为了训练神经网络模型，将预报量（气溶胶光学厚度 AOT、沙尘高度 Z_{dust} ）和预报变量（AIRS 亮度温度 BTs、地表高度和相关的大气质量数据）联系起来进行训练的一系列数据就十分必要。近期 NASA 的 A-Train 任务为构建这些必要的数据集提供了有效方法，从中可以完成人工神经网络的训练工作。

3.1 数据集的选取

AIRS 作为 A-Train 任务的一部分，搭载于 Aqua 卫星上，共有 2378 个通道覆盖在红外波谱的 3.8-4.6 μm 、6.2-8.2 μm 、8.8-15.4 μm 波段范围，用来测量亮度温度 BTs。虽然在白天和夜晚都可以进行探测，但我们选择 6.2-15.4 μm 的长波波段探测的数据来构建训练数据集，排除受太阳光影像严重的波段（3.8-4.6 μm ）在白天的数据。另一方面，由于沙尘高度信息应该与沙尘层间辐射标识有关，以及短波波段的散射影响不可忽略，都是排除短波波段（3.8-4.6 μm ）数据的重要原因。AIRS 在 6.2-15.4 μm 波段范围的通道大致可分为 4 个部分：大气吸收区（WIN；8.15-12.8 μm ）、CO₂ 吸收区（CO₂；12.9-15.0 μm ）、水蒸气吸收区（WV；6.3-8.0 μm ）和 O₃ 吸收区（O₃；9.2-10.1 μm ）。关于红外谱段吸收的详细信息见参考文献^[51]中的详细介绍。本文仅仅考虑低噪声当量温度差的通道（NE Δ T），并去除 WIN 波段弱气体吸收的通道，一共选择 234 个通道（WIN、CO₂、WV 和 O₃ 波谱段分别为 130、31、44 和 29 个通道）来进行人工神经网络模型的构建。

下图所示为所选通道的波谱分布。其中绿色、红色、黄色和蓝色的点分别代表 CO₂、WIN、O₃、WV 波谱段的通道数据。值得指出的是水蒸气连续吸收(尤其是近地表的水蒸

气连续吸收)会对 WIN 通道传统吸收波段间的亮度温度 BTs 产生影响。虽然水蒸气的持续吸收对于本项研究影响不大，但却期望后续的工作通过将 130 个 WIN 通道的 TBs 映射为目标参数间接反映出来。

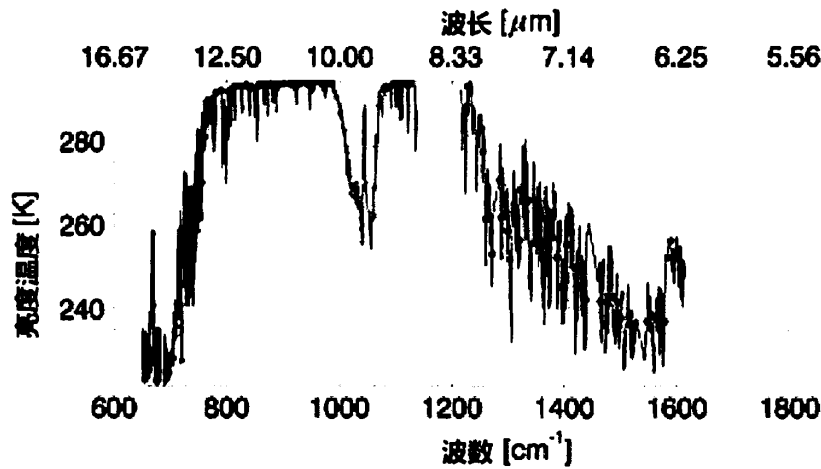


图 3.1 在本次研究中选择的用来进行人工神经网络模型构建的 AIRS 234 个通道数据
Fig. 3.1 Selected AIRS 234 channels used for the ANN model development in this study

MODIS 传感器也搭载于 Aqua 卫星，以 250 米、500 米、1000 米的空间分辨率依据波长来探测可见光和红外（0.415-14.235μm）波谱段的 36 个通道的地物反射。MODIS 2 级气溶胶产品提供了 10 千米分辨率基于可见光测量的气溶胶属性信息^[52]。在本次研究中，针对陆地，我们选择 MODIS 的气溶胶光学厚度 AOTs 产品与 AIRS 的地表亮度温度数据一起进行训练，这是因为深蓝算法（利用比 0.5μm 短的波长）在探测干旱半干旱地区的沙尘气溶胶时会更加有效^[9]。针对海洋，我们将通过建立查找表方法得到的 AOTs^[53]与 AIRS 的亮度温度数据一起进行训练。

CALIPSO 卫星也是作为 A-Train 任务的一部分，搭载了具有两个波长（532nm 和 1064nm）的激光雷达传感器（云气溶胶激光雷达，CALIOP），提供了沿卫星轨道的气溶胶和薄云横切面视图。CALIPSO 激光雷达 2 级气溶胶层顶和底部高度数据以及 532nm 的气溶胶光学厚度产品可以用来展示沙尘层的垂直分布，我们选择将其作为训练数据集来构建人工神经网络模型。为了识别沙尘负荷的像元，被判识为沙尘或污染沙尘气溶胶的 CALIPSO 气溶胶产品被用来与 AIRS 的 BTs 一起进行训练。CALIOP 测量的 532nm 的总的后向散射衰减廓线用来对本次研究中 AIRS 反演得到的沙尘高度进行验证。这将在后续部分进行讨论。表 3.1 概括了本次研究中将会用到的数据集。

表 3.1 本研究选取的数据集

Table 3.1 Description of Data Sets Used in This Study

产品	分辨率	参数	参数描述
AIRS Level 1B	13.5km	Radiances	AIRS 234 个通道的亮度温度
		Topog	地表地形数据
		Satzen	卫星天顶角数据
MODIS Level 2	10km	Optical Depth Land and Ocean	海洋上空的气溶胶光学厚度
		Deep Blue Aerosol Optical	陆地上空的气溶胶光学厚度
CALIPSO Level 2	5km	Layer Top Altitude	沙尘层顶部高度
		Layer Base Altitude	沙尘层底部高度
		Feature Optical Depth 532	532nm 波段测量的气溶胶光学厚度
		Feature Classification Flag	特征分类标识
CALIPSO	1/3km	Total Attenuated Backscatter	532nm 波段测量的总的后向散射衰减

许多人可能会更加关注于 MODIS 反演气溶胶光学厚度的不确定性在人工神经网络反演中的影响。为了减少这些不确定性的干扰，我们采用最好质量的海洋的 MODIS 气溶胶光学厚度和 CALIPSO 场景分类标识数据。而地表的 MODIS 气溶胶光学厚度（利用深蓝算法得到）和 CALIPSO 的沙尘高度数据不作深入处理，这是因为质量控制信息对它们的影响不大。

地表地形（ Z_{sfc} ）和相关的地表到传感器的大气质量数据（ $1/\mu$ ； $\mu=\cos\theta$ ； θ 是卫星的天顶角）也作为模型的输入。这两种数据从 AIRS L1B 产品中获取。11 μm 和 12 μm 间的亮度温度差异(BTDs)也作为输入数据来检验人工神经网络模型对于给定输入的敏感度。取 11.8~12.2 μm 和 10.8~11.2 μm 波段范围的差异计算 BTD。

3.2 数据集相关性分析

由于卫星遥感监测的红外波段辐射率携带了与大气参数相关的信息，而且有着强烈的非线性关系。例如，测量的大气顶部（TOA）的红外辐射率与用各层平均温度的发射率表示的大气辐射以及地表辐射都有很强的相关性。因此，借助解释沙尘属性和红外辐射率非线性关系的工具探究基于高光谱测量获取沙尘信息的程度变得尤为重要。使用配置数据集，人工神经网络首先探究给定输入参数的模型表现。虽然人工神经网络方法获取的通道权重可以在调节 AIRS BTs 和沙尘属性关系时直接使用，但我们探究为什么给定输入节点参数表现更好的物理原因。

3.2.1 对比 AIRS 数据与当量 MOIDS 产品数据

对于由大气柱计算得到的沙尘数据，光学厚度的光谱变化可以用光谱消光系数变化来表示，消光系数的计算由给定沙尘属性的米散射计算公式得到^[51]。

$$\tau_{\lambda 1} = \frac{Q_{\text{ext}, \lambda 1}}{Q_{\text{ext}, \lambda 2}} \tau_{\lambda 2} \quad (3.1)$$

式中， $Q_{\text{ext}, \lambda 1}$ 和 $Q_{\text{ext}, \lambda 2}$ 分别代表两个波长的消光系数。假设 $\lambda 1$ 和 $\lambda 2$ 分别代表可见光和红外波段的波长，那么基于红外气溶胶光学厚度就可以用公式（3.1）转为可见光的气溶胶光学厚度。当沙尘在气溶胶成分中占主导地位时，公式（3.1）就可以表示为一阶近似，将不同波段的红外气溶胶光学厚度转为可见光光学厚度。然而，本次研究中并不是简单地将亮度温度 BTs 转为红外气溶胶光学厚度，再通过公式（3.1）将红外光学厚度转为可见光光学厚度。在人工神经网络模型中，通过训练过程使得红外波谱的亮度温度是与可见光气溶胶光学厚度 AOTs 具有一定的相关性。值得指出的是公式（3.1）只有在气溶胶性质同样对于红外波谱敏感的情况下才会有效。

训练时，预报量（目标输出）在 AOT 反演模型中被定义为与 MODIS 在 550nm 反演的气溶胶光学厚度相当的量。在这里，AIRS 亮度温度 BTs 是与 550nm 的气溶胶光学厚度 AOTs 间接相关的，这是因为 AIRS 像元大小（最低分辨率为 13.5km）与 10km 分辨率的 MODIS 气溶胶光学厚度产品相似，MODIS AOT 目标位置与用来训练的 AIRS 视场（FOV）类似。AOT 的训练数据使用 2007 年 2 月到 5 月间塔克拉玛干地区（25-50N，63-90E）的数据，数据比较充足。之所以选择这个时间段的数据进行训练，是因为春季沙尘发生的频率比较高。这个过程中，在使用 DeSouza-Machado 等人参考沙尘探测函数（DDA）得到的基于 AIRS 的沙尘探测方法后^[54]，只有沙尘负荷的 AIRS 像元被保留下来。考虑到低的仪器噪声和红外波谱段的沙尘吸收，为避免水汽吸收波段的影响，沙尘探测函数（DDA）基于 BTD 方法，选择 5 个位于大气窗口的波谱段（8.12、8.86、10.40、11.11 和 12.6 μm ）的 AIRS 通道数据。

为了验证选取的 AIRS 数据集参数与 MODIS 反演的沙尘气溶胶光学厚度 AOT 之间的相关程度，我们将不同的输入参数与 46073 组当量 MODIS AOTs 联系起来，并展开有效的测试来显示模型的表现。测试工作通过有序添加参数的方式来展开（例如，按照累计的方式首先添加 WIN 波段的 BTs，然后添加 CO₂、WV 和 O₃ 波段的 BTs，最后按序添加 Zsfc、1/u 和 BTD）。不考虑不同输入层的影响，并且保持隐层神经元的数量一直不

变。

图 3.2 展示了对于每一组输入参数反演 AOT 得到的相关系数和均方根误差(RMSE)的统计, 分别由实线和虚线来表示。横坐标上的 WIN、CO₂、WV、O₃ 分别代表 AIRS BTs 在大气、CO₂、水蒸气和 O₃ 波段的输入参数; Zsfc 表示地表地形; 1/u 为从地表到达传感器的相关的大气质量数据; BTD 为 11 μ m 和 12 μ m 之间的亮度温度差异。输入由累积的方式进行展开。可以看出: 基于 WIN 波段 BTs 的训练结果与 MODIS AOT 的反演结果具有较好的一致性, 相关系数大于 0.86, 这表明: 高光谱 WIN 通道的测量值携带了丰富的关于沙尘 AOTs 的信息。增加更多其它波段的通道数据(如 CO₂、WV、O₃ 等)进行训练, 随着通道数的递增统计结果会不断改善。结合不同波段的 BTs, 添加 Zsfc 和 1/u 数据也会有助于提高相关性, 相关系数最终达到 0.90 左右, 最小均方根误差在 0.33 左右。但 BTD 数据的增加会使得统计结果变差, 这可能是因为 BTD 数据表征了 WIN 通道的平均信息, 而这些信息在 130 个 WIN 通道中已经被使用过的原因。

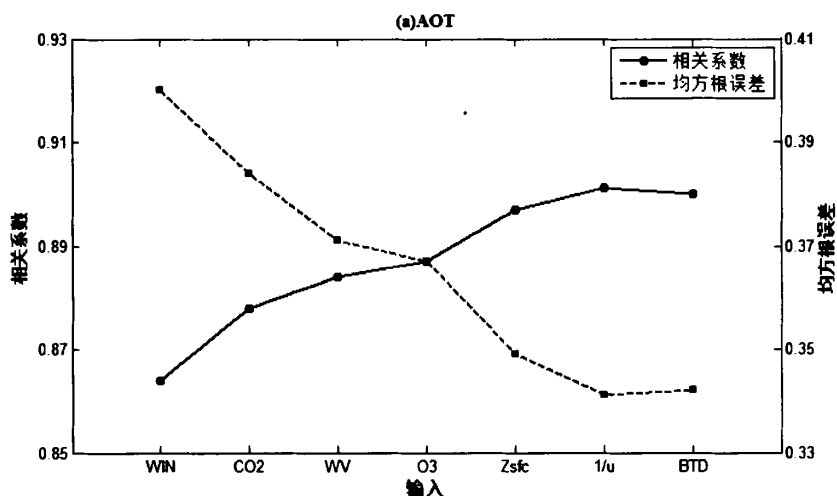


图 3.2 AIRS 反演的 AOTs 与 MODIS 获取的 AOTs 之间的相关系数和均方根误差统计图

Fig. 3.2 Correlation coefficients and RMSEs between AIRS-derived and MODIS-estimated AOTs

3.2.2 对比 AIRS 数据与当量 CALIPSO 产品数据

对于给定的大气气柱, 当 CALIPSO 特征分类标识指示为沙尘或污染沙尘时使得沙尘像元才被最终确定下来。然后, 沙尘层的垂直分布由 CALIPSO 反演得到的单个或多个沙尘层顶部和底部高度来确定。对于人工神经网络训练模型来得到沙尘高度, 一个 CALIPSO 像元中所有沙尘层的高度以响应层的气溶胶光学厚度来衡量, 从而计算给定区域内单个沙尘高度。

$$Z_{\text{dust}} = \sum_{i=1}^N z(i) \frac{\text{AOT}(i)}{\text{AOT}(\text{total})} \quad (3.2)$$

这里 $z(i)$ 和 $\text{AOT}(i)$ 分别代表海平面以上的平均高度和第 i 沙尘层的气溶胶光学厚度。 N 代表 CALIPSO 产品中总的沙尘层数。因此，在公式 (3.2) 中， $\text{AOT}(\text{total})$ 代表 N 个独立沙尘层的总的气溶胶光学厚度。由于权重的调整， Z_{dust} 可以被认为是有效沙尘高度，但我们为了简洁起见仍然采用“沙尘高度”的概念。

由于不同的空间分辨率（比如 AIRS 和 CALIPSO 的空间分辨率分别为 13.5km 和 5km），多个 CALIPSO 像元被包含于一个 AIRS 视场中。如果至少一个 CALIPSO 像元被确定为沙尘负荷，那么在同一个 AIRS 视场中所有的 CALIPSO 气溶胶测量值将被平均地构建为一对一的匹配数据。取 AIRS 视场中 CALIPSO 像元的平均值，沙尘高度就可以通过公式 (3.2) 中总的 AOT 气柱来衡量，在通过亮度温度计算 MODIS 气溶胶光学厚度的同一地理区域进行计算。然而，由于沿着 CALIPSO 测量轨道，使得相同的 4 个月的 AOT 反演的训练周期难以满足高效的 Z_{dust} 匹配数据的构建。所以，为了在训练中获得充足的样本，选择两年（2007 年 1 月到 2008 年 12 月）周期的数据。

使用 50976 组配置数据，借助与反演 AOT 同样的方法我们来探究沙尘高度 Z_{dust} 反演对于给定输入参数的敏感程度。图 3.3 展示了这些敏感度测试的统计特性。对于只有 WIN 通道的 BTs 数据，相关系数能达到 0.68，虽然比 AOT 实验中的值小，但也表明 Z_{dust} 同样地对 WIN 通道最为敏感。值得指出的是：增加 CO_2 通道的 BTs 数据也会使得统计结果发生实质性变化，从图中可以看出，相关系数增加到 0.74，均方根误差从 0.94km 减小到 0.86km。增加 CO_2 通道数据使得统计结果实质性改善表明：在 WIN 通道数据基础上增加 CO_2 的通道数据可以有效改善高光谱测量中 Z_{dust} 的反演结果。这可能是因为从 CO_2 通道数据中获取的温度分布信息可以添加到基于 WIN 通道获取的沙尘辐射信息中，从而导致对于大气顶部 TOA 辐射更好的描述。

相比来讲，其它的输入参数比如 WV、 O_3 、 $1/u$ 和 BTD 虽然可以少量地改善统计结果，但整体看起来对 Z_{dust} 的反演结果影响不大。而且 BTD 数据的参与也会对 Z_{dust} 的反演成果造成不利影响。

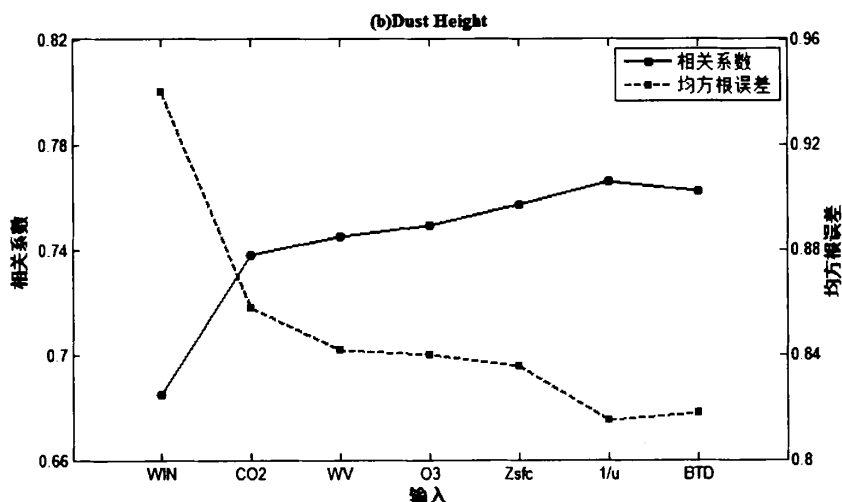


图 3.3 AIRS 反演的 Z_{dust} 与 CALIOP 测量的 Z_{dust} 之间的相关系数和均方根误差统计图

Fig. 3.3 Correlation coefficients and RMSEs between AIRS-derived and CALIPO-measured dust height

基于敏感度测试的结果我们可以得出结论：WIN 通道的数据对于 AOT 和 Z_{dust} 的反演最为重要。但添加 CO_2 通道的数据会大大改善 Z_{dust} 的反演成果。然而，在本次研究中，为了通过增加参数来改善统计结果，我们使用除 BTB 之外的所有参数来进行 AOT 和 Z_{dust} 人工神经网络模型的建立。

3.3 本章小节

本章通过对 MODIS、CALIPSO 数据集进行简要介绍，选择能表征沙尘状态属性的数据集产品作为标准沙尘参数。对 AIRS 中的数据集与标准沙尘参数进行相关性分析，选择与其相关性高的亮度温度、地表高度和卫星天顶角数据作为后续人工神经网络训练的样本数据和测试数据，为后续的研究工作提供数据准备。

4 建立人工神经网络模型

4.1 人工神经网络概述

人工神经网络（Artificial Neural Networks, ANN）起源于 20 世纪 40 年代，距今发展已经半个多世纪。它是一种数据挖掘工具，是由大量的、简单的处理单元（称为神经元）广泛的互相连接而形成的复杂网络系统。它反映了人脑功能的许多基本特征，是一个高度复杂的非线性动力学系统。神经网络具有大规模并行、分布式存储和处理、自适应、自组织和自学习的能力，非常适合处理需要同时考虑许多因素和条件的、不精确和模糊的信息处理问题。正是由于这些特性，人工神经网络的应用与研究正成为人工智能、认知科学、神经生理学、非线性动力学等相关专业的热点。

4.1.1 神经网络发展历史

自从上世纪 40 年代初，美国的 McCulloch 和 Pitts 从信息处理的角度，研究神经细胞行为的数学模型表达提出二值神经元模型以来，人们对神经网络进行了大量的研究。其中经历了 40 年代末心理学家 Hebb 提出著名的 Hebb 学习规则，50 年代 Rosenblatt 提出感知器模型，60 年代初期 Grossberg 对感知器模型的发展，人工神经网络发展到一个高潮期，60 年代末神经网络研究的低潮，研究人员开始了一系列反思，80 年代提出新的神经网络 HNN 和 Boltzmann 机，神经网络再一次复兴等一系列的过程。目前，神经网络的发展非常迅速，从理论上对它的计算能力、对任意连续函数的逼近能力、学习理论以及动态网络的稳定性分析上都取得了丰硕的成果。

4.1.2 生物神经元结构

神经元就是神经细胞，它是动物的重要特征之一，在人体内从大脑到全身存在大约 10^{10} 个神经元。每个神经元与大约 $10^3 \sim 10^5$ 个其他神经元相连接，构成一个极为庞大而复杂的网络，即“生物神经网络”。生物神经网络中各神经元之间连接的强弱，按照外部的激励信号作适应性变化，每个神经元又随着所接受的多个激励信号的综合结果呈现出兴奋和抑制状态。

人脑的学习过程就是神经元之间连接强度随外部激励信息作适应变化的过程，人脑处理信息的结果由各神经元状态的整体效果来确定。显然，神经元是人脑信息处理系统

的最小单元。简单的神经元由细胞体、树突、轴突和突触组成。细胞体是神经元的本体，内有细胞核和细胞质，完成普通细胞的生存功能。树突有大量的分枝，多达 10^3 数量级，长度较短（通常不超过 1 毫米），用以接受来自其它神经元的信号。轴突用以输出信号，有些较长（可达 1 米以上），轴突的远端也有分枝，可与多个神经元相连。突触是一个神经元与另一个神经元相联系的特殊部位，通常是一个神经元轴突的端部靠化学接触或电接触将信号传递给下一个神经元的树突或细胞体。神经元间的信号通过突触传递。通过它，一个神经元内传送的冲击信号将在下一个神经元内引起响应，使下一个神经元兴奋，或阻止下一个神经元兴奋。

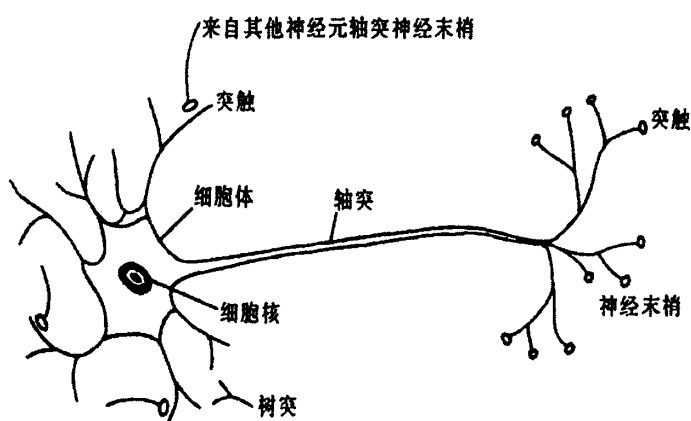


图 4.1 生物神经元示意图

Fig. 4.1 Biologic Neuron Sketch Map

4.1.3 人工神经元模型

人工神经网络是在现代神经生物学基础上提出的模拟生物过程用以反映人脑某些特性的计算结构，它不是人脑神经系统的真实描写，而只是它的某种抽象、简化和模拟。数学上的神经元模型和生物学上的神经细胞相对应，在人工神经网络中，神经元常被称为“处理单元”，或者从网络的观点出发，把它称为“节点”。按照生物神经元的结构和工作原理构造一个人工神经元如下图所示：

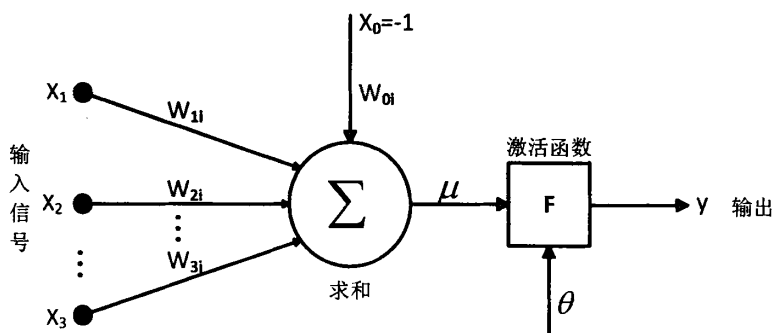


图 4.2 人工神经元模型

Fig. 4.2 Model of Artificial Neuron

作为人工神经网络基本单元的神经元模型，有三个基本要素：

- 1) 一组连接，连接强度由各连接上的权值 W_i 表示，权值为正表示激活，为负表示抑制。
- 2) 一个求和单元，用于求取 n 个输入信号的加权和（线性组合），对人工神经元的全部输入信号进行整合。
- 3) 一个非线性激活函数 F ，起非线性映射作用，并将神经元输出幅度限制在一定的范围之内。常见的激活函数 F 有阶跃函数、段线性函数和 Sigmoid 函数等。

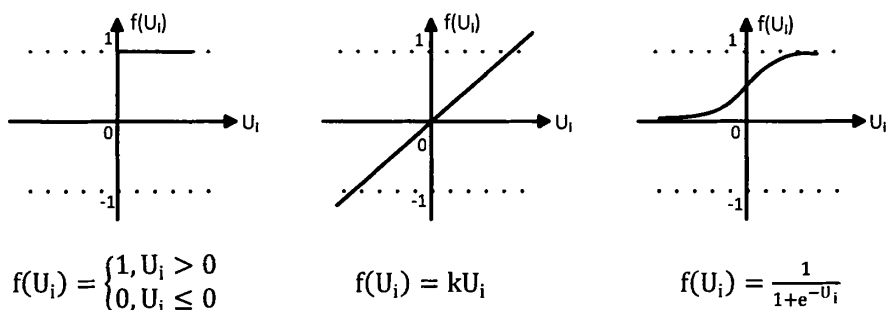


图 4.3 (a)阶跃函数

Fig. 4.3 (a)Step Function

(b)线性函数

(b)Linear Function

(c)S 型函数

(c)Sigmoid Function

此外，还有一个阈值 θ （或偏置 $-\theta$ ）。

4.1.4 网络结构及工作方式

除单元特性外，网络的拓扑结构也是人工神经网络的一个重要特性。从连接方式看，人工神经网络主要有两种。

- 1) 前馈型网络。各个神经元接受前一层的输入，并输出给下一层，没有反馈。节点

分为两类，即输入单元和计算单元，每一计算单元可有任意个输入，但只有一个输出。通常前馈网络可分为不同的层，第 i 层的输入只与第 $i-1$ 层的输出相连，输入和输出节点与外界相连，而其他中间层则称为隐层。前馈网络中有明显的层次关系，信息单方向地从输入层向输出层流动，常用的有多层感知器、BP 网络等。

2) 反馈型网络。所有节点都是计算单元，同时也可接收输入，并向外界输出，其中每个连接弧都是双向的。若总单元数为 n ，则每一个节点有 $n-1$ 个输入和一个输出，常用的 Hopfield 网络就属于这种网络。

人工神经网络的工作过程主要分为两个阶段：第一个阶段是学习期，此时各个计算单元状态不变，各连线上的权值可通过学习来修改；第二个阶段是工作期，此时各连接权固定，计算单元状态变化，以达到某种稳定状态。

5.1.5 人工神经网络的特点

1) 固有的并行结构和并行处理。

人工神经网络与人类的大脑类似，不但结构上是并行的，其处理顺序也是并行的和同时的。在同一层内的处理单元都是同时工作的，即神经网络的计算功能分布在多个处理单元上，而传统的计算机通常只有一个处理单元，其处理顺序是串行的。

2) 知识的分布存储。

在神经网络中，知识不是存储在特定的存储单元，而是分布在整个系统中，要存储多个知识就需要很多连接。当一个神经网络输入一个激励时，它要在已存储的知识中寻找与该输入匹配最好的知识存储为其解。

3) 容错性。

人类大脑具有很强的容错能力，这正是由于大脑中知识是存储在很多处理单元和它们的连接上的。人工神经网络可以从不完善的数据和图形进行学习和做出决定。由于知识存在整个系统中，而不是在一个存储单元内，因此一定比例的结点不参与运算，对整个系统的性能不会产生重大影响。

4) 自适应性。

人类有很强的适应外部的学习能力，比如小孩在周围环境的熏陶下可以学会很多事情，通过学习可以认字、说话、走路、思考、判断等。人工神经网络也具有学习能力，根据学习的类型可分为：有指导的训练：将输入样本加到网络输入并给出相应的输出，通过多次训练迭代获得连接权值。无指导的训练：网络通过训练自行调节连接加权，从

而对输入样本分类。神经网络的综合推理的能力使其具有正确响应和分辨从未见过的输入样本的能力。

4.2 网络模型的选择

人工神经网络通过将输入空间划分到可能与期望输出值相关的子域中，提供了一种方便有力的展示分类与回归非线性关系的方法。我们用一种叫做多层感知器（MLP）的人工神经网络模型来建立 550nm 的 MODIS AOT 和 CALIPSO 的 Z_{dust} 与 AIRS 在 WIN、CO₂、WV、O₃ 波段的亮度温度 BTs、Zsfc、1/u 以及 BTD 的关系，这是一种前馈的神经网络。

在本文的工作中，要解决的问题是对于大气中沙尘的监测，利用历史监测数据来训练网络，使其掌握从输入空间向输出空间的映射关系，这属于监督性学习，在监督性学习的网络中，误差反向传播网络（BP 网络）是最经典的。这种网络是多层的，由一个输入层、一个输出层和一个以上的隐蔽层构成，并由权重来调节各层之间的关系，它的功能很强，可以解决高度非线性化的问题。因此，本文采用 BP 网络作为预测网络，具体实现时，需要先构造一个网络框架，对于 BP 网络，Matlab 的 NN Toolbox 提供了建立神经网络的专用函数 `newff()`。用 `newff` 函数来确定网络层数、每层中的神经元数和传递函数，其语法为：

$$\text{net} = \text{newff}(\text{PR}, [\text{S1}, \text{S2}, \dots, \text{SN}], \{\text{TF1}, \text{TF2}, \dots, \text{TFN}\}, \text{BTF}, \text{BLF}, \text{PF})$$

其中，PR 表示由每个输入向量的最大最小值构成的 $R \times 2$ 矩阵；Si 表示第 i 层网络的神经元个数；TF 表示第 i 层网络的传递函数，缺省为 `tansig`；BTF 表示字符串向量，为网络的训练函数名，可在如下的函数中进行选择：`traingd`、`traingdm`、`traingdx`、`trainbfg`、`trainlm` 等，缺省为 `trainlm`；BLF 表示字符串变量，为网络的学习函数名，缺省为 `learnsgdm`；BF 表示字符串变量，为网络的性能函数，缺省为均方差“mse”。`newff()` 在确定网络结构后会自动调用 `init` 函数用缺省参数来初始化网络中各个权重和阈值，产生一个可训练的前馈网络，即该函数的返回值为 `net`。对于 `newff()` 函数中各参数的选择，将在下面的论述中进行介绍。

4.3 网络层数的确定

除了一个输入层和一个输出层，BP 网络可以包含不同数目的隐层。隐含层数的合理选取是网络取得良好决策性能的一个关键步骤。隐层层数应根据问题的复杂性，综合考

虑网络的精度和训练时间。增加网络的层数虽然可以进一步降低误差，提高精度，但同时也使得网络更加复杂，从而增加了网络训练时间。当映射关系比较简单，网络精度满足的条件下，可以选择较少的层数，这样可加快训练，当模式样本数很多时，为了减小网络规模，增加一个隐层是必要的，但一般 BP 网络隐层数不超过两层。

针对本文的研究工作，由于我们所要寻求的输入输出之间是一种有理函数的近似，非线性关系比较简单，因此，选择较为经典的 3 层 BP 神经网络，即一个输入层，一个输出层和一个隐层，来进行后期的训练工作。

4.4 隐层节点数的选取

隐层节点的个数是网络构造中的关键问题，节点数太少不能满足拟合精度，且训练次数多、训练时间长。它对神经网络所起的作用就相当于光学中的分光镜，将混杂于输入信号中的相互独立的基本信号分离出来，再组合出新的向量，来实现网络由输入向输出的映射^[55]。理论上讲，对于一个具体的应用问题应该存在一个最佳的隐含层节点个数，这个最佳个数与问题的重复性和网络的具体结构形式以及各隐层节点函数的特性有关。为了尽可能地避免训练时出现“过拟合”现象，保证足够高的网络性能和泛化能力，确定隐层节点数的最基本原则是：在满足精度要求的前提下选取尽可能紧凑的结构，即选取尽可能少的隐层节点数。

有一个确定最佳隐层节点数的常用方法，称为试凑法。这种方法先设置较少的隐节点训练网络，然后逐渐增加隐节点数，再用同一样本集进行训练，从中确定网络误差最小时对应的隐节点数。在使用此方法时，可以先借助一些经验公式。这些公式计算出来的隐节点数只是粗略的估计值，可作为试凑法的初始值。比如：

$$m = \sqrt{n+1} + \varepsilon \quad (4.1)$$

式中： m 表示初始的隐层节点数； n 为输入层的节点数； l 为输出层的节点数； ε 为 0-10 之间的一个常数。

或者：

$$m = \log_2 n \quad (4.2)$$

式中： m 表示初始的隐层节点数； n 为输入层的节点数。

或者：

$$m = \sqrt{0.43nl + 0.12l^2 + 2.54n + 0.77l + 0.35} + 0.51 \quad (4.3)$$

式中： m 表示初始的隐层节点数； l 为输出层节点数； n 为输入层的节点数。

还有一些隐层节点的确定方法，在这里不再详述。

本文选取 2007-2010 年的数据集来作为训练样本，2011-2013 年的数据集作为检验样本，通过变换隐层节点数及迭代次数进行多次试报，预报的成功界限指数为：

$$CSI = \frac{c_f}{c_f + w_f} \times 100\% \tag{4.4}$$

式中： c_f 为正确预报的结果数； w_f 为试漏报或空报之和。得到的实验结果如下表所示：

表 4.1 实验结果
Table 4.1 The experimental results

模型	迭代次数	隐层节点数	CSI (%)	准确率 (%)
模型 1	100	5	43	83
模型 2	300	5	41	85
模型 3	500	5	39	81
模型 4	300	6	34	79
模型 5	300	4	25	78

从表中可以看出：隐层节点数并不是越多越好，因为如果隐层节点数过多，虽然有利于精确地逼近每一个样本点，使网络识别率高且记忆性好，但其振荡剧烈会使得网络推广能力变差。而过少的隐层节点数，将会导致网络缺少灵活性和分类能力。所以我们需要在实际应用中，经多次试验后才能确定网络的隐层节点数。迭代次数也不是越高越好，迭代次数过高会使得过于迁就某些不典型、不精确的样本，这样虽然拟合较好但会导致偏离总体规律。但是如果网络训练次数过少，识别率太低，同样会造成推广能力较差。

表中模型 2 的预报效果较好，预报的成功界定指数达到 41%，准确率达到 85%，其中，错误的预报跟其它影响沙尘暴的因子没有考虑进训练网络以及网络参数的选取有关。因此，进一步考虑其它因素的影响，不断优化 BP 神经网络，理论上会取得更好的预报效果，在这里不再详述，我们最终选定隐层节点数为 5 进行后续的工作。

4.5 激活函数的选定

不同的激活函数可以反映样本输入与输出之间不同的对应关系，因为 BP 网络的非

线性逼近能力是通过 S 型传递函数体现出来的，所以，隐层中一般采用 S（sigmoid）型的传递函数，如双曲正切 S 型（tansig）传递函数、对数 S 型（logsig）传递函数。输出层的传递函数一般采用纯线性函数（purelin）或 S 型函数。传递函数表达式分别为：

$$A(n) = \text{tansig}(n) = \frac{1-e^{-n}}{1+e^{-n}}$$

(4.5)

$$A(n) = \text{logsig}(n) = \frac{1}{1+e^{-n}}$$

(4.6)

$$A(n) = \text{purelin}(n) = n$$

(4.7)

在相同的训练次数和学习速率条件下，选定 5 组输入层—输出层不同的激励函数，通过对三层 BP 网络中输入、输出的激活函数进行不同组合的训练对比，得到实验结果如下表所示：

表 4.2 不同激励函数组合结果对比

Table 4.2 Combination results of different excitation functions

模型	激活函数组合	最大拟合误差（%）	收敛情况	检验误差（%）
模型 1	Tansig-tansig	0.36	收敛	1.86
模型 2	Tansig-logsig	84	不收敛	32
模型 3	Logsig-logsig	253	不收敛	67
模型 4	Logsig-purelin	236	不收敛	75
模型 5	Tansig-purelin	7.8	收敛	2.01

可以看出：使用 Tansig-logsig、Logsig-logsig、Logsig-purelin 传递函数都出现了不收敛的现象，Tansig-tansig 传递函数组合模型收敛，且训练和检验效果最好，又考虑到 BP 网络的非线性逼近能力是通过 S 型激活函数体现出来的，因此输入层一般采用 S 型的激活函数，这里，我们就选用 tansig 函数作为输入层的激活函数。模型 5 的 Tansig-purelin 函数组合也能达到收敛的效果，虽然检验效果略逊于 Tansig-tansig 函数组合，但相比于 S 型函数输出范围的限制，purelin 函数可以不受约束，因此，我们选定 purelin 函数作为输出层的激活函数。更多详细的训练过程解释见 Lee and Sohn[2012]^[56]。

通过上述神经网络结构、网络模型、模型参数的选定等工作后，我们可以得到沙尘暴预报模型建模框架，如下图所示：

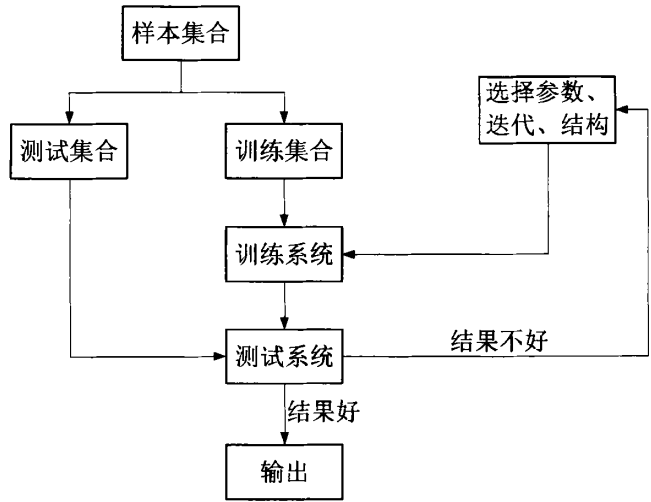


图 4.4 建模过程流程图

Fig. 4.4 The flow chart of modeling process

4.6 神经网络模型的验证分析

为了验证上述神经网络模型的成果，我们比较基于 AIRS 像元反演的气溶胶光学厚度 AOT 和沙尘高度 Z_{dust} 与 MODIS AOT 以及 CALIPSO Z_{dust} 之间的差异。为了验证基于 AIRS 测量值反演的 AOT 成果，我们选用 2008 年 2 月至 5 月间 AIRS 反演的 AOT 值与当量 MODIS 反演的 AOTs 值进行比较。对于验证 AIRS 测量值反演沙尘高度的成果，我们选用 2009 年 1 月至 12 月一整年的 AIRS 反演的 Z_{dust} 数据与当量 CALIPSO 测量值进行对比。两组对比实验选用同一地理区域来进行数据集的训练工作。

选用 2008 年 2 月到 5 月时间内，共 57384 组当量 AOT 的对比统计结果以二维直方图的形式来给出（见图 4.5），间距选择 0.1 个单位。从图中可以看出：基于人工神经网络方法从 AIRS 数据中反演得到的 AOTs 与 MODIS 反演的 AOTs 具有较好的一致性。相关系数能达到 0.84，平均偏差为 0.03，均方根误差为 0.37。统计结果表明：在 550nm 处 MODIS 反演得到的沙尘气溶胶光学厚度 AOT 可以被高光谱红外测量值准确地表征，这可能是基于公式 3.1 中波长和消光系数的关系来进行连结的。换句话说：基于高光谱红外测量的数据可以准确获取红外沙尘气溶胶光学厚度 AOTs 数据，从而通过公式 3.1 得到光学的气溶胶光学厚度 AOTs。

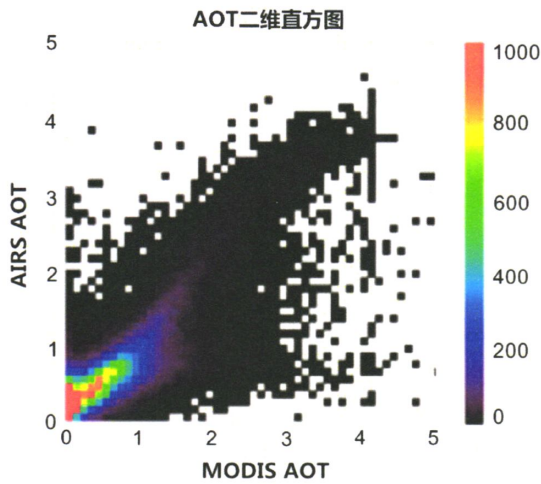


图 4.5 MODIS 反演的 AOTs 和人工神经网络反演的 AOTs 对比的二维直方图
Fig. 4.5 Two-dimensional histograms of MODIS-derived versus ANN-retrieved AOTs

为了验证沙尘高度 Z_{dust} 反演的效果，基于 AIRS 反演的像元级别的 Z_{dust} 数据被分配到与一年周期的 CALIPSO 测量值进行对比。图 4.6 展示了 2009 年 1 月到 12 月时间内，一共 20272 个 AIRS 沙尘像元的验证统计结果。结果显示：其反演的沙尘高度 Z_{dust} 与 CALIPSO 测量值之间也具有较好的一致性。相关系数为 0.79，平均偏差为 0.02km，均方根误差为 0.54km。值得指出的是：从权重上考虑，大部分的沙尘负荷位于 3km 以下，沿 CALIPSO 轨道方向的 AIRS 视场 FOV 约为 13.5km，这将导致不到一半的像元区域位于 CALIPSO 像元中，而且这些像元也并非全部都是沙尘像元。因此，可以认为当使用准确的全部为沙尘像元当量的数据集时会得到更好的反演效果。

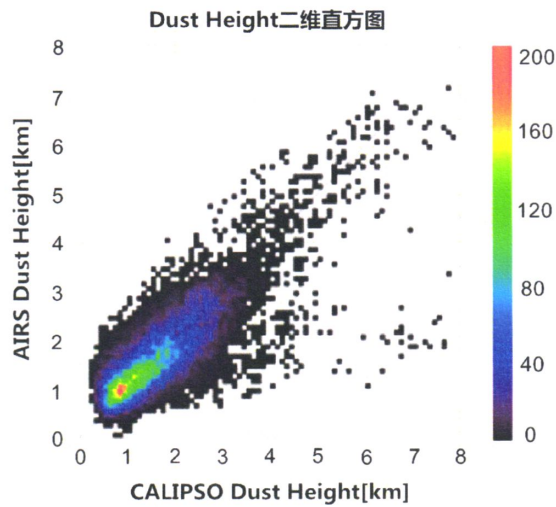


图 4.6 CALIPSO 反演的 Z_{dust} 与人工神经网络反演的 Z_{dust} 对比的二维直方图
Fig. 4.6 Two-dimensional histograms of CALIPSO-derived versus ANN-retrieved Z_{dust}

4.7 本章小节

本章通过论述神经网络的发展历史、生物神经元结构、人工神经元模型以及神经网络的工作方式和特点等，介绍了人工神经网络的概要情况。在后续的工作中，通过对网络模型的选择，网络层数的确定，隐层节点数的选取以及激活函数的选定等建立起人工神经网络模型。通过对网络模型的反演结果进行验证分析，不断调整网络结构参数，使得网络模型在沙尘暴预报中更加稳定和精确。

5 基于神经网络反演沙尘参数

除了利用统计数据验证神经网络模型之外，我们还在二维空间探究利用神经网络模型的反演结果与 MODIS、CALIPSO 测量值的对比分析。分别选取 2013 年 4 月 4 日和 7 日沙尘爆发时期塔克拉玛干沙漠地区的影像进行研究，一方面，对神经网络反演气溶胶光学厚度和沙尘高度的结果进行验证分析，另一方面，对塔克拉玛干地区沙尘暴的动态传输以及运动进行分析。在这里，我们进行神经网络模型训练得到 Z_{dust} 时并没有使用 $1/u$ 数据，这是因为 CALIPSO 的视场只位于 AIRS 10~19 度的天顶角范围内，超出这个范围的 AIRS 视场数据并无意义，因此，天顶角数据 $1/u$ 对于模型反演沙尘高度的影响可以忽略不计。下图（图 5.3 和图 5.6）AIRS 反演的沙尘高度 Z_{dust} 通过 234 个通道的 BTs 数据和 Z_{sfc} 数据来得到。

5.1 2013 年 4 月 4 日案例

图 5.1 展示了 2013 年 4 月 4 日塔克拉玛干沙漠地区亚洲沙尘爆发时期 MODIS 真彩色影像、AIRS 亮度温度和亮度温度差异的分布。在真彩色影像中（图 5.1a）青藏高原北部（位于影像南部）亮黄色的沙尘区清晰可见，暗表面上被残云覆盖着。通过 AIRS 亮度温度影像（图 5.1b）可以看出，研究区整体辐射强度比较大，亮度温度值在以北纬 39 度，东经 81 度为中心的椭圆状区域附近达到最大，南侧有一个较宽的区域，亮度温度的值都在 270 左右，说明辐射强度相对较小。在亮度温度差异的影像中（图 5.1c）可以明显看出：AIRS 亮度温度值达到最大的区域附近与影像其它区域差异值也较大。加上南侧低亮温的情况可以推测出：在北纬 39 度，东经 81 度附近可能有沙尘暴的存在，而且有向东南发展的趋势。

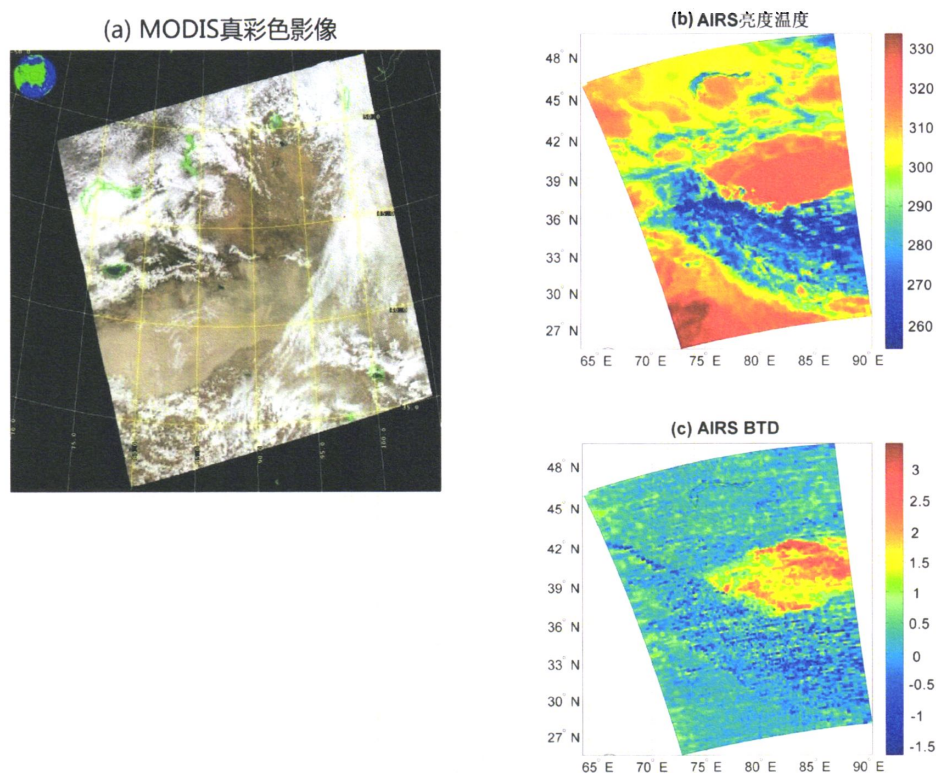


图 5.1 2013 年 4 月 4 日塔克拉玛干沙漠地区 (a)MODIS 真彩色影像,(b)AIRS 亮度温度影像,(c)AIRS 亮度温度差异影像

Fig. 5.1 (a) MODIS true color image, (b) AIRS bright temperature image, and (c) AIRS bright temperature difference image on 4 April 2013 over the Taklimakan desert area

图 5.2 展示了同一地区和时期利用神经网络模型基于 AIRS 测量值数据反演得到的气溶胶光学厚度 AOT 和 MODIS 反演的 AOT 的二维分布。AIRS 基于像元的 AOT 反演值（图 5.2a 所示）展示了沙尘从东北部到西南部的发展状况，其中，在以北纬 39 度、东经 81 度地区（塔克拉玛干沙漠的东部）为中心的区域，气溶胶光学厚度值能达到 4.0，说明可能有沙尘暴的存在，这也验证了上文对于亮度温度以及亮度温度差异影像的分析。在大范围区域下，AIRS 的反演值所显示的一般模式和幅度分布与基于深蓝算法得到的 MODIS AOTs 值十分相似。MODIS 深蓝算法探测到北纬 43.5 度以北的地区有弱气溶胶光学厚度 AOTs，这片区域在 MODIS 真彩色影像中清晰可见。然而，CALIPSO 的测量值（图 5.3 所示）在北纬 43.5 度以北的地区（天山山脉东部末尾区域）并未显示出沙尘的存在，所以可以推断由深蓝算法判定的低 AOT 值区域不一定与沙尘有关。AIRS 反演的 AOTs 与 MODIS AOTs 之间当量像元的相关系数为 0.87（这里并未显示）再一次表明

基于像元的结果与人工神经网络模型反演的结果具有很强的一致性。

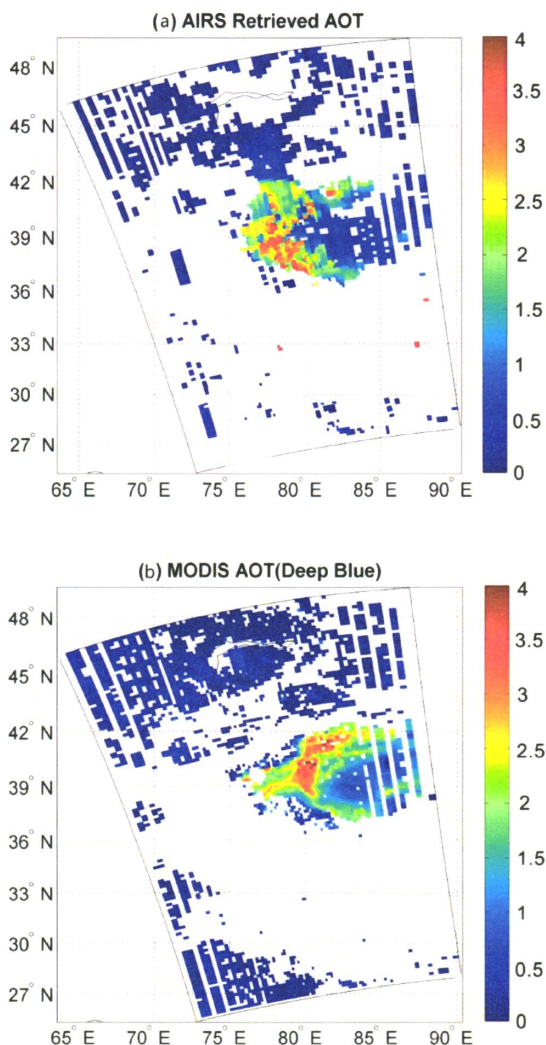


图 5.2 2013 年 4 月 4 日塔克拉玛干沙漠地区 (a)基于 AIRS 测量值反演的气溶胶光学厚度 AOT, (b)MODIS 反演的气溶胶光学厚度

Fig. 5.2 (a) AOT retrieved from AIRS measurements, (b) MODIS derived AOT on 4 April 2013 over the Taklimakan desert area

图 5.3a 显示了本案例中, AIRS 反演的 Z_{dust} 数据的二维分布。黑色的实线指示了 CALIPSO 卫星从南到北的轨道方向。图 5.3b 展示了 CALIOP 在 532nm 处沿卫星轨道方向总衰减后向散射的垂直视图和 AIRS 测量值反演的 Z_{dust} (+号表示), 后向散射单位为 $\text{km}^{-1}\text{sr}^{-1}$ 并乘以 1000, 在图 5.3b 中数值小于 $4.0 \times 10^{-3} \text{km}^{-1}\text{sr}^{-1}$ 的单元被掩盖掉。图 5.3a 中显示沙尘高度值在 38°N , 81°E 区域附近达到最大, 其基本模式与气溶胶光学厚度的

分布基本一致，这也从侧面验证了在此区域附近沙尘暴爆发的可能性。图 5.3b 显示的高度的一致性也表明 AIRS 反演的结果可以很好地表征沙尘层的高度。一个值得注意的有趣现象是：图 5.3a 中，小的 Z_{dust} 值（小于 3km）区域大致对应于低于海平面 1000m 的较低的沙漠盆地区域。

虽然图 5.3b 展示了 AIRS Z_{dust} 和 CALIPSO 总的衰减后向散射之间有较好的一致性，但是在 42.5~43 度之间的纬度带还是有明显偏差。在这个地带，单个的强沙尘层被分成多个层。在人工神经网络沙尘高度的训练模型中，在给定像元位置被各层气溶胶光学厚度权衡过的所有的 CALIPSO 沙尘层被转化为单个 AOT 值权衡的沙尘高度，导致对于多个沙尘层，差异性会被加大。除了反演的错误，图 5.3b 中 AIRS Z_{dust} 和 CALIOP 后向散射之间的差异部分是因为沙尘高度的定义差异造成的，尤其与多个沙尘层有关。

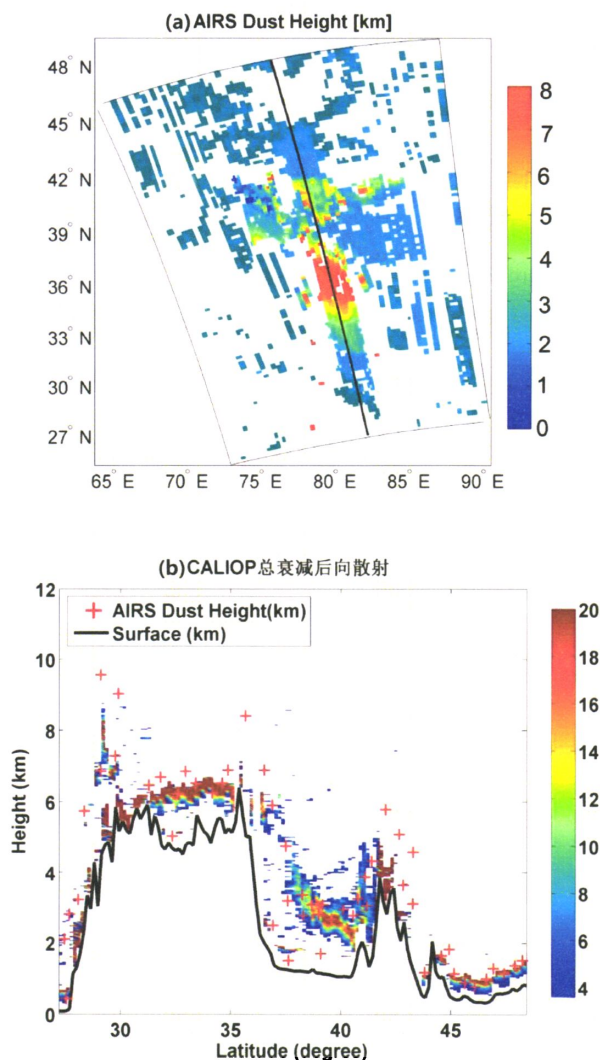


图 5.3 2013 年 4 月 4 日塔克拉玛干沙漠地区 (a)AIRS 测量值反演的沙尘高度二维分布, (b)532nm 处沿 CALIPSO 轨道方向 CALIOP 测量的总衰减后向散射垂直视图(图 5.3a 黑线)

Fig 5.3 (a) Two-dimensional distributions of dust heights retrieved from AIRS measurements and (b) vertical view of CALIOP-measured total attenuated backscatter at 532nm on 4 April 2013 over the Taklimakan desert area, along the CALIPSO track (black line in Figure 5.3a)

5.2 2013 年 4 月 7 日案例

为了展示塔克拉玛干地区沙尘的运动状态以及传输过程,我们选取 2013 年 4 月 7 日(间隔三天)同一地区的相同影像来进行对比分析。在 MODIS 真彩色影像中(图 5.4a),北纬 35 度到 40 度条带间的地物清晰可见,东北以及西北部地区被云覆盖。在 AIRS 亮

度温度影像中(图 5.4b), 北纬 30 度到 36 度之间的宽条带区域地物辐射值(亮度温度值) 都比较大。在此条带附近向两侧延伸时, 亮度温度值逐渐减小。从亮度温度差异的影像 中(图 5.4c) 可以看到, 亮度温度差异的最大值转移到 35°N , 72.5°E 的区域, 说明在 此区域附近可能有强辐射的沙尘暴出现, 而且经过三天的时间, 沙尘由东北部地区向 西南部方向进行转移(对比图 5.1c 和图 5.4c)。

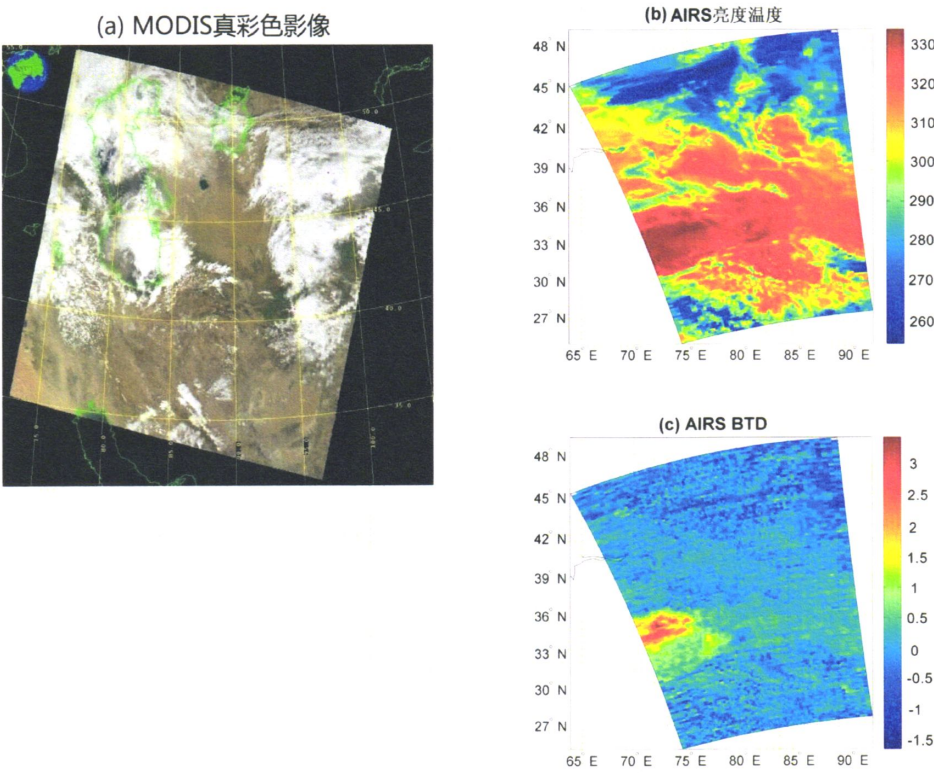


图 5.4 2013 年 4 月 7 日塔克拉玛干沙漠地区 (a)MODIS 真彩色影像, (b)AIRS 亮度温度影像, (c)AIRS 亮度温度差异影像

Fig. 5.4 (a) MODIS true color image, (b) AIRS bright temperature image, and (c) AIRS bright temperature difference image on 7 April 2013 over the Taklimakan desert area

图 5.5 展示了同一地区和时间基于 AIRS 测量值反演的气溶胶光学厚度 AOT 与 MODIS AOT 二维分布间的对比影像。两幅影像无论在 AOT 的分布模式还是值的大小上都极为相近, 这也表明利用训练出的神经网络模型基于 AIRS 测量值反演的结果可以有效逼近 MODIS 基于深蓝算法得到的气溶胶光学厚度。两幅影像中, AOT 的极大值都出现在 36°N , 72°E 区域附近, 这说明在此出现的沙尘暴极有可能为 2013 年 4 月 4 日由东北部地区传输而来, 并且被 AIRS 和 MODIS 测量值有效地捕捉到, 同时也验证了对于

图 5.4 的论述和分析。

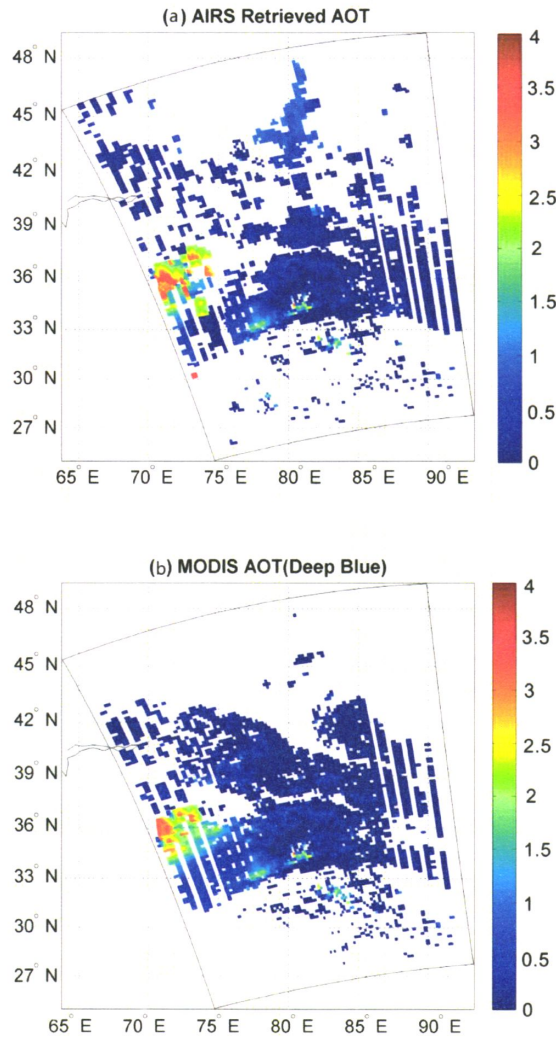


图 5.5 2013 年 4 月 7 日塔克拉玛干沙漠地区 (a)基于 AIRS 测量值反演的气溶胶光学厚度 AOT, (b)MODIS 反演的气溶胶光学厚度

Fig. 5.5 (a) AOT retrieved from AIRS measurements, (b) MODIS derived AOT on 7 April 2013 over the Taklimakan desert area

图 5.6a 展示了本案例中 AIRS 反演得到的 Z_{dust} 值的二维分布, 并且在图 5.6b 中用沿着自南向北轨道 (图 5.6a 的黑线部分) CALIOP 探测的总衰减后向散射的垂直视图与其作对比。可以看出 CALIPSO 沿轨道测量的沙尘高度分布可以用 AIRS 的反演值很好地表示。然而, 在北纬 30 度和 44 度附近的高海拔地区, 沙尘高度值都被过高地估计。虽然并不清楚是不是复杂地形导致了过高的估计值, 但并不奇怪的是在复杂的高海拔地区遥

感探测受到的的干扰也会比较多。仔细观察可以发现北纬 39 度以南的地区高的后向散射衰减值可能并不与沙尘的存在相关，而是与云有关（见 MODIS 真彩色影像）。经过与 4 月 4 日（图 5.3b）的沙尘层对比发现在垂直方向上沙尘也存在不断扩张的趋势。在自然情况下，因为 AIRS Z_{dust} 代表了沙尘层平均气溶胶光学厚度权重的高度，所以如果总的沙尘层很厚或者扩散严重的话反演结果可能会与 CALIOP 测量值之间存在差异。

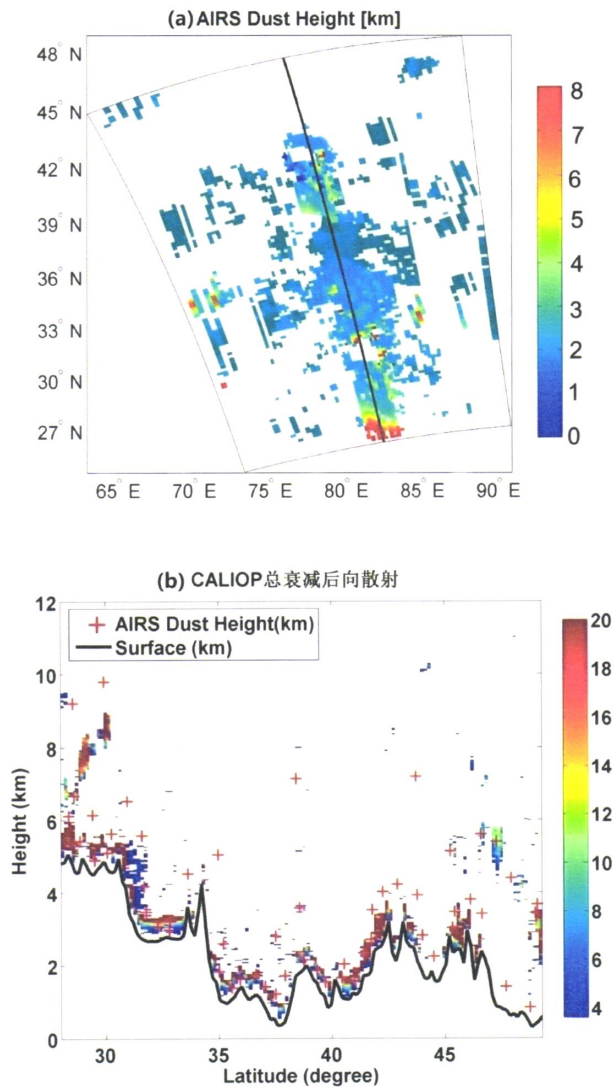


图 5.6 2013 年 4 月 7 日塔克拉玛干沙漠地区 (a)AIRS 测量值反演的沙尘高度二维分布, (b)532nm 处沿 CALIPSO 轨道方向 CALIOP 测量的总衰减后向散射垂直视图(图 5.6a 黑线)

Fig 5.6 (a) Two-dimensional distributions of dust heights retrieved from AIRS measurements and (b) vertical view of CALIOP-measured total attenuated backscatter at 532nm on 7 April 2013 over the Taklimakan desert area, along the CALIPSO track (black line in Figure 5.6a)

5.3 本章小节

本文选取 2013 年 4 月 4 日与 2013 年 4 月 7 日沙尘爆发时期的影像进行对比分析,一方面验证了利用神经网络得到的 AIRS 测量值反演的结果可以有效表示 MODIS 深蓝算法反演的气溶胶光学厚度和 CALIPSO 测量的沙尘高度。另一方面,经过分析可以知道 4 月 7 日监测到的沙尘是由 4 月 4 日影像中部的沙尘传输而来,并且被 AIRS、MODIS 和 CALIPSO 有效捕捉,这对于研究沙尘的动态传输和运动提供了有利条件。

案例研究表明:在沙尘爆发区,当研究区沙尘气溶胶占主要成分时,利用神经网络得到的 AIRS 的反演值可以很好地与 MODIS 监测的 AOTs 和 CALIPSO 测量的总衰减后向散射相契合。然而,当远离沙尘爆发源区域时,由于在传输过程中沙尘粒子会与其它气溶胶粒子结合,会使得沙尘粒子在光学属性上发生变化。在海洋上部,除了将沙尘粒子与混合物分离带来的一些潜在问题外,沙尘粒子还会与比它大许多的海洋盐气溶胶粒子相融合。尤其当沙尘气溶胶与其它类型的气溶胶粒子相结合时,我们的重点应该放在如何进行基于红外波段的沙尘粒子反演上面。

6 结论与展望

6.1 结论

本文利用人工神经网络方法从 AIRS 高光谱红外测量值中反演塔克拉玛干地区沙尘的气溶胶光学厚度 AOT 和沙尘高度 Z_{dust} 。为了这种方法的开展,我们将两种训练数据集分开进行:为了获取沙尘气溶胶光学厚度 AOT,将当量 AIRS BTs (选择 WIN、CO₂、WV 和 O₃ 分别为 130、31、44、29 个通道的数据)和 550nm 处 MODIS 深蓝算法反演的气溶胶光学厚度 AOTs 来进行训练;为了获取沙尘高度 Z_{dust} ,将 AIRS 相同通道的 BTs 和 CALIPSO 沙尘高度 Z_{dust} 来进行训练。与卫星观测量相关的 Z_{sfc} 值和 $1/u$ 值也加入到当量数据集中。在人工神经网络模型中,不管是将 AIRS BTs 与 MODIS AOT 还是与 CALIPSO Z_{dust} 进行训练,大量的敏感度测试都表明:在高光谱红外测量中,WIN 通道的亮度温度数据在反演 AOT 和 Z_{dust} 时都起着至关重要的作用,这是因为沙尘状态底部表面的辐射会与沙尘层发生作用,包括沙尘层的吸收和红外波段的发射作用。因此,表面温度、表面辐射和沙尘层红外波段的辐射对于大气顶部辐射影响很大,而这些影响在近乎透明的窗口通道被很好地解决。另一方面,在 CO₂ 通道的亮度温度 BTs 对于 Z_{dust} 的反演也起到重要作用。基于敏感度测试结果表明,将 234 个通道的 BTs、 Z_{sfc} 和 $1/u$ 数据作为输入,建立起了反演气溶胶光学厚度 AOT 和沙尘高度 Z_{dust} 的人工神经网络模型。

为了评估人工神经网络模型反演沙尘气溶胶光学厚度 AOT 的效果,将不同于训练数据集时间段的 AIRS 测量值反演的 AOTs 与 MODIS 反演的 AOTs 进行比较。像元级别的 AOT 比较结果显示:相关系数为 0.88,平均偏差为 0.03,均方根误差为 0.37,这表明基于人工神经网络模型反演的气溶胶光学厚度 AOTs 与 MODIS 反演的 AOTs 基本一致。基于 AIRS 测量值反演的像元级别的沙尘高度 Z_{dust} 与 CALIPSO Z_{dust} 也保持了较好的一致性,其中,相关系数为 0.79、平均偏差为 0.02km,均方根误差为 0.54km。

然而,值得指出的是:这只是一个表明高光谱红外探测在监测沙尘属性(AOT 和沙尘高度)能力方面的遥感实验。事实上,由于神经网络自身功能的限制,使得如果要将此方法应用于其它方面,比如在不同季节,不同地域来进行沙尘的监测或者其它信息的监测还需要用到当时当地的数据集来进行训练。

6.2 展望

在这两个中国沙漠地区反演沙尘的案例中，二维的沙尘反演量与 MODIS AOT 和 CALIPSO Z_{dust} 都能很好地契合，这表明人工神经网络算法可以应用于沙尘监测过程中。考虑到 AIRS 测量可以不受时间的限制，利用高光谱红外探测来改善监测和预测沙尘信息的能力具有巨大潜力。未来也可以将此项技术更加深入地应用到其它传感器，比如 IASI 和国家极地轨道环境卫星系统预备项目的 CrIS (Crosstrack Infrared Sounder)，或者其它沙尘地区，比如撒哈拉沙漠和阿拉伯沙漠地区，或者结合更加精确的地基测量数据训练 AIRS 相关的数据集得到沙尘信息（如 CE318）会取得更加有效的结果。深入一点来讲，将人工神经网络应用于其它遥感监测与分类等领域，也会有十分广阔的前景。

致谢

时光飞逝，转眼间两年的研究生生涯即将落下帷幕。回忆起这两年的点点滴滴，有痛苦，有欢笑，有失败的泪水，也有成功的喜悦。我很珍惜这段来之不易的人生旅程，它让我不断成长，使我更加完善。

几个月的论文写作过程，从刚开始的茫然不知所措，到后来慢慢找到研究方向，一步一步走过，可谓是酸甜苦辣、五味俱全。期间，除了自己努力的汗水，还有许多人给予的帮助，在他们的支持和鼓励下我才能静下心来专心完成整个研究过程并取得满意的结果。在此，我很想对他们说一声谢谢。

我要感谢的人很多。

许君一老师给予我生活和学习上的许多帮助令我毕生难忘。他的热心，和蔼，善良和对于学生的关怀让我对老师这个职业充满敬佩和仰慕。他对我论文提出的建设性意见更对我写好本次论文起到巨大的催化作用。许老师为人严谨，学识渊博。整个研究期间，许老师给我许多的指导，帮助我走出研究困境，让我少走了许多弯路。论文不断整改，每次修改许老师都会重新检查一遍，字字斟酌，这种认真负责的态度让我深受感动。

很感谢王佳同学长期以来的陪伴和支持，她积极乐观的态度、热情似火的性格、勤奋努力的个性，激励着我不断进步，让我对未来充满希望。

感谢师门的小伙伴们给我带来的许多的快乐，感谢师哥、师姐、师妹还有小王子陪我度过整个研究生生涯。这个春天，风中飘散的，满是幸福的气息。

青春岁月就这样一闪而过。走出校门，走入社会，生活将会是另外的一番风景。

校园中的人、事、一草一木都令我眷恋与不舍。

也许，多年过后还会回忆起曾经的那份激情与浪漫。

参考文献

- [1] 夏祥鳌, 王明星. 气溶胶吸收及气候效应研究的新进展[J]. 地球科学进展, 2004, 19(4): 630-635.
- [2] 沈志宝, 魏丽. 我国西北大气沙尘气溶胶的辐射效应[J]. 大气科学, 2000, 24(4): 541-548.
- [3] 石广玉, 王标, 张华, 等. 大气气溶胶的辐射与气候效应[J]. 大气科学, 2008, 32(4): 826-840.
- [4] 赵翠光. 人工神经网络方法在沙尘暴短期预报中的应用[J]. 气象, 2004, 30(4): 39-41.
- [5] Park, H. -S., B. J. Sohn. Recent trends in changes of vegetation over East Asia coupled with temperature and rainfall variations[J]. Geophys. Res., 2010, 115, D14101, doi:10.1029/2009JD0.
- [6] Lee, E. H., B. J. Sohn. Recent increasing trend in dust frequency over Mongolia and Inner Mongolia regions and its association with climate and surface condition change[J]. Atmos. Environ., 2011, 45(27), 4611–4616, doi:10.1016/j.atmosenv. 2011. 05. 065.
- [7] Kim, J. Transport routes and source regions of Asian dust observed in Korea during the past 40 years (1965–2004)[J]. Atmos. Environ., 2008, 42(19), 4778–4789, doi:10.1016/j.atmosenv. 2008. 01. 040.
- [8] Hsu, N. C., J. R. Herman, and C. Weaver. Determination of radiative forcing of Saharan dust using combined TOMS and ERBE data[J]. Geophys. Res., 2000, 105(D16), 20649–20661, doi:10. 1029/2000jd900150.
- [9] Hsu, N. C., J. R. Herman, et al. Aerosol properties over bright-reflecting source regions, IEEE Trans[J]. Geosci. Remote Sensing., 2004, 42(3), 557–569, doi:10.1109/TGRS.2004.824067.
- [10] Remer, L. A., et al. The MODIS aerosol algorithm, products, and validation[J]. Atmos. Sci., 2005, 62(4), 947–973, doi:10.1175/JAS3385.1.
- [11] Sinyuk, A., O. Torres, and O. Dubovik. Combined use of satellite and surface observations to infer the imaginary part of refractive index of Saharan dust[J]. Geophys. Res. Lett., 2003, 30(2), 1081, doi:10.1029/2002GL016189.

- [12] DeSouza-Machado, S. G., et al. Infrared retrievals of dust using AIRS: Comparisons of optical depths and heights derived for a North African dust storm to other collocated EOS A-Train and surface observations[J]. *Geophys. Res.*, 2010, 115(D15), D15201, doi:10.1029/2009jd012842.
- [13] Peyridieu, S., A. Chédin, et al. Saharan dust infrared optical depth and altitude retrieved from AIRS: A focus over North Atlantic Comparison to MODIS and CALIPSO[J], *Atmos. Chem. Phys.*, 2010, 10(4), 1953–1967, doi:10.5194/acp-10-1953-2010.
- [14] Yao, Z., J. Li, et al. Asian dust height land infrared optical depth retrievals from hyperspectral longwave infrared radiances[J]. *Geophys. Res.*, 2012, 117, D19202, doi:10.1029/2012JD017799.
- [15] Hankin E. h., On dust raising winds and descending currents[J]. part VI, *Indian met Memoirs*, 1921, 22(1): 80-84.
- [16] Sutton L J, Haboobs[J], *Quart. J. Roy. Meteor. Soc.* 1925, 51(2): 25-30.
- [17] Joseph P V, Raipal D K, Deka S N, “Andhi” the convective dust storm of Northwest India[J]. *Mausam*, 1980, 31(3): 431-442.
- [18] Brazel, A. J. , W. C. Nicking, The relationship of weather types dust storm generation in Arizona[J]. *JClimotoloey*. 1986, 6(3). 255-275.
- [19] Auregui E. The dust storms of Mexico City[J]. *Inter. J. Climatology*, 1989, 9(2): 169-180.
- [20] Gillttte, D. A. J. Adams, A. Endo, D. Smith and A. kill, Threshold velociter for input of soil particles into air by desert soils[J]. *Geophys. Res.*, 1980, 85: 5621-5623.
- [21] Bergam etti G. *Encyclopedia of Earth System Science*[M]. San Diego: Academic Press., 1992. 171-192.
- [22] Tegen I, Fung I. Modeling of mineral dust in the atmosphere: Source, transport, and optical thickness[J]. *Geophysical Research.*, 1994, 99(22): 897-914.
- [23] Ikuko Mori, Masataka Nishikawa, Toshifumi Tanimura. Change in size distribution and chemical composition of kosa (Asian dust) aerosol during long-range transport[J]. *Atmospheric Environment*, 2003(37): 4253-4263.
- [24] Youngsin Chun, Jiyoung Kim, Jae Cheon Choi. Characteristic number size distribution of aerosol during Asian dust period in Korea[J]. *Atmospheric Environment*, 2001 (35): 2715-2721.

- [25] Yehouda Enzel, Rivka Amit. The climatic and physiographic controls of the eastern Mediterranean over the late Pleistocene climates in the southern Levant and its neighboring deserts[J]. *Global and Planetary Change*, 2008(60): 165-192.
- [26] Uemastu M, Duce P, Prospero J, et al. Transport of mineral aerosol from Asia over north Pacific ocean[J]. *Geophysical Research.*, 1993, 88(c9): 5343-5 352.
- [27] Shaw G. E. transport of Asian desert to the Hawian islands[J], *J. Appl. Meteor*, 1980(19): 1254-1259.
- [28] Nilgun Kubilay, Slobodan Nickovic, Francois Dulac, et al. An illustration of the transport and deposition of mineral dust onto the eastern Mediterranean[J]. *Atmospheric Environment*, 2000, 34: 1293-1303.
- [29] Wang Shigong, Dong Guangrong, Yang Debao, et al. A study on sandstorm over the desert region in north China[J]. *Journal of Natural Disasters*, 1996, 6(2): 31-37. [王式功, 董光荣, 杨德保, 等. 中国北方地区沙尘暴变化趋势初探[J]. *自然灾害学报*, 1996, 6(2): 31-37.].
- [30] Quan Linsheng, Shi Shaoying, Zhu Yafen, et al. Temporal spatial distribution characteristics and causes of dust day in China[J]. *Acta Geographica Sinica*, 2001, 56(4): 477-485. [全林生, 时少英, 朱亚芬, 等. 中国沙尘天气变化的时空特征及其气候原因[J]. *地理学报*, 2001, 56(4): 477-485.].
- [31] Hu Yinqiao, Y Mitshuta. Micrometeorological characteristics and local triggering mechanism of strong dust storm[J]. *Chinese Journal of Atmospheric Sciences*, 1997, 5: 581-589. [胡隐樵, 光田宁. 强沙尘暴微气象特征和局地触发机制[J]. *大气学*, 1997, 5: 581-589.].
- [32] Ji Fei, Qin Yu. The numerical simulation on dust storm over East Asia: a case analysis[J]. *Acta Scientiarum Naturalium Universitatis Pekinensis*, 1998, 34(5): 639-645. [纪飞, 秦瑜. 东亚沙尘暴数值模拟: 个例分析[J]. *北京大学学报(自然科学版)*, 1998, 34(5): 639-645.].
- [33] 叶笃正, 丑纪范, 刘纪远, 等. 关于我国华北沙尘大气的成因与治理对策[J], *地理学报*, 2000, 55(5): 513-521.
- [34] 牛若芸, 薛建军, 周自江. 2002 年我国沙尘暴天气特征分析[J]. *南京气象学院学报*, 2004, 27(2): 178-184.

- [35] 吕新苗, 刘惠清, 王文杰. 北京周边地区沙尘暴时空分布特征及其环境背景[J]. 地理研究, 2004, 23(1): 38-43.
- [36] 王宁, 张红. 吉林省沙尘天气的时空分布特征及其分型预报研究[J]. 吉林气象, 2003, (2): 24-26.
- [37] 牛生杰, 孙继明, 桑建人. 贺兰山地区沙尘暴发生次数的变化趋势[J]. 中国沙漠, 2000, 20(1): 55-58.
- [38] 李正明. 临夏州沙尘天气短期实时预报系统[J]. 甘肃农业, 2005, 2: 26-27.
- [39] 路志英, 张启孟, 赵智超. 基于 SVM 的沙尘暴预测模型[J]. 天津大学学报, 2006, 39(9): 1110-1114.
- [40] Y. J. Kaufman. Satellite sensing of aerosol absorption[J]. Geophys. Res., vol. 92, pp. 4307-4317, 1987.
- [41] Kaufman Y J and Sendra C. Algorithm for automatic atmospheric corrections to visible and near IR satellite imagery[J]. International Journal of Remote Sensing, 1988, 9(8): 1357-1381.
- [42] Holben B N, Vermote E, Tanré D, et al. Aerosol retrieval over land from AVHRR data Application for atmospheric correction[J]. IEEE Transactions Geoscience and Remote Sensing, 1992, 30(2): 212-222.
- [43] Liu G R, Chen A J, Liu T H, et al. Applying SPOT data to estimate the aerosol optical depth and air quality[J]. Environmental Modelling and Software, 2002, 17(1): 3-9.
- [44] David M. Winker, C. Hosteler, M. Vaughnet, et al. CALIOP algorithm theoretical basis document. part1: CALIOP instrument, and algorithms overview, September 9, 2006: 4-26.
- [45] Chris A. Hostetler, Z. Liu, J. Reaganet, et al. CALIOP algorithm theoretical basis document, calibration and level 1 data products, 7 April 2006: 5-64.
- [46] 王宏, 石广玉, 王标, 等. 中国沙漠沙尘气溶胶对沙漠源区及北太平洋地区大气辐射加热的影响[J]. 大气科学, 2007, 31(3): 515-526.
- [47] 段婧, 毛节泰. 气溶胶与云相互作用的研究进展[J]. 地球科学进展, 2008, 23(3): 252-261.
- [48] 王根. 主成分累计影响系数法一种用于高光谱资料的通道选择新技术[D].南京: 南京信息工程大学. 2011.

- [49] Krasnopolsky, V. M., H. Schiller. Some neural network applications in environmental sciences. Part I: Forward and inverse problems in geophysical remote measurements[J]. Neural Networks, 2003, 16(3-4), 321-334, doi:10.1016/S0893-6080(03)00027-3.
- [50] Mas, J. F., and J. J. Flores, The application of artificial neural networks to the analysis of remotely sensed data, Int[J]. Remote Sens, 2008, 29(3), 617-663, doi:10.1080/01431160701352154.
- [51] Liou K. An Introduction to Atmospheric Radiation[J]. Physics Today, 2008.
- [52] Remer, L. A., et al. The MODIS aerosol algorithm, products, and validation[J]. Atmos. Sci., 2005, 62(4), 947-973, doi:10.1175/JAS3385.1.
- [53] Levy, R. C., L. A. Remer, et al. Evaluation of the Moderate Resolution Imaging Spectroradiometer (MODIS) retrievals of dust aerosol over the ocean during PRIDE[J]. Geophys. Res., 2003, 108(D19), 8594, doi:10.1029/2002jd002460.
- [54] DeSouza-Machado, S. G., et al. Infrared retrievals of dust using AIRS: Comparisons of optical depths and heights derived for a North African dust storm to other collocated EOS A-Train and surface observations[J]. Geophys. Res., 2010, 115(D15), D15201, doi:10.1029/2009jd012842.
- [55] 崔利群. 人工神经网络在数据挖掘中的应用研究[D]. 沈阳:辽宁工程技术大学, 2004.
- [56] Lee, S.-S., B. J. Sohn. Nighttime AOT for Asian dusts MODIS IR measurements: An Artificial Neural Network approach[J]. Meteorol. Soc. Jpn., 2012, 90(2), 163-177, doi:10.2151/jmsj.2012-201.
- [57] 矫梅燕, 周自江. 中国沙尘暴及其监测预警[J]. 中国工程科学, 2003, 5(9): 1-6.
- [58] 王汉芝. 基于模糊神经网的沙尘暴预报模型[D]. 天津: 天津大学, 2002.
- [59] 陈守余, 周梅春. 人工神经网络模拟实现与应用[M]. 北京: 中国地质大学出版社, 2000.
- [60] 向国全, 董道珍. 1997. BP 模型中的激励函数和改进的网络训练法[J]. 计算机研究与发展, 34(2): 113-11.
- [61] Friedrich B K, Barmann F. A Learning Algorithm for Multilayered Neural Networks Based on Linear Square Problems[J]. Neural Networks, 1993, 4(6): 127-602.
- [62] 尹宏, 韩志刚. 气溶胶大气对太阳辐射的吸收[J]. 气象学报, 1989, 47(1): 118-123.

- Yin Hong, Han Zhigang. Absorption of solar radiation by aerosol atmosphere[J]. Acta Meteorologica Sinica, 1989, 47(1): 118-123.
- [63] 刘玉洁, 杨忠东等. MODIS 遥感信息处理与算法[M]. 北京: 科学出版社, 2001.