

# 气象数据高精度融合技术研究

毛紫薇<sup>1</sup>, 章璇<sup>2</sup>, 唐加山<sup>1,3</sup>

<sup>1</sup>南京邮电大学通信与信息工程学院, 江苏 南京

<sup>2</sup>南京邮电大学理学院, 江苏 南京

<sup>3</sup>南京邮电大学随机数学研究中心, 江苏 南京

收稿日期: 2024年10月25日; 录用日期: 2024年11月18日; 发布日期: 2024年11月25日

## 摘要

降水观测在气象领域至关重要, 本文利用变分同化方法结合雷达与雨量计数据提升降水场估测精度, 通过构建变分同化模型, 并优化数据融合过程, 实现了高精度降水估计。研究成果为气象预报、水文学及洪水预警等领域提供了可靠数据支持, 并为未来研究提供了参考价值。

## 关键词

变分同化, 降水观测, L-BFGS-B迭代算法, 雷达, 雨量计

# Research on High-Precision Fusion Technology for Precipitation Observation

Ziwei Mao<sup>1</sup>, Xuan Zhang<sup>2</sup>, Jiashan Tang<sup>1,3</sup>

<sup>1</sup>School of Communication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing Jiangsu

<sup>2</sup>School of Science, Nanjing University of Posts and Telecommunications, Nanjing Jiangsu

<sup>3</sup>Research Center of Stochastic Mathematics, Nanjing University of Posts and Telecommunications, Nanjing Jiangsu

Received: Oct. 25<sup>th</sup>, 2024; accepted: Nov. 18<sup>th</sup>, 2024; published: Nov. 25<sup>th</sup>, 2024

## Abstract

Precipitation observation is crucial in meteorology. This study explores the strategy of enhancing precipitation field estimation accuracy through variational data assimilation (VDA) methods, combining radar and rain gauge data. By constructing a variational assimilation model and optimizing the data fusion process, high-precision precipitation estimation is achieved. The research findings

provide reliable data support for meteorological forecasting, hydrology, and flood warning, and point to future research directions.

## Keywords

Variational Data Assimilation, Precipitation Observation, L-BFGS-B Iteration Algorithm, Radar, Rain Gauge

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

在气象领域,降水观测的精确性对于水文学、洪水预报等应用至关重要[1][2]。雨量计,作为点观测的典范,以其高精度著称,能够精确捕捉局部地区的降水情况[3]。张欣[4]等使用对比实验验证了雨量计监测的合理性,雷梦佳[5]等制作了智能防堵磁致式雨量计原型样机,能达到与人工完全一致的雨量监测结果。但雨量计覆盖范围始终有限且数据不连续,难以满足大尺度降水估计的需求。相比之下,雷达观测具有大范围、连续性和快速响应的特点[6]。王文丽[7]将经多点校准法校准的雷达估测降雨数据、卡尔曼滤波校准后的雷达估测降雨数据输入到模型中进行径流模拟,评价其适用性。但雷达的估计精度存在明显不足,难以满足准确预测降雨量的需求。

为了克服单一观测手段的不足,学者们致力于发展数据同化技术,黄静[8]采用间接同化方式来同化雷达反射率因子,使得结果更接近真实值。张新[9]采用卡尔曼滤波算法对测雨雷达数据进行同化分析,降水误差得到有效降低,相比较而言,李梦迪[10]等验证了雷达融合雨量计降水估测方法,其准确性比单一手段同化技术高,且对极端降水有较好的捕捉和监测能力。

综上所述,本文综合分析雨量计及雷达的观测数据,结合 L-BFGS-B 迭代算法,为气象预报、水资源管理及防灾减灾提供更加准确、可靠的降水信息支持。

## 2. 模型基础

### 2.1. 模型背景与假设

背景 1: 现有两部具有不同探测精度雷达(文中表示为左雷达和右雷达)对同一区域的观测数据。左边雷达的降水测量误差为 50%, 结构估计误差为 15%; 右边雷达的降水测量误差为 20%, 结构估计误差为 10%。

背景 2: 在背景 1 的基础上增加 N 点雨量计观测数据, 建立变分同化模型。

设立背景的主要目的是应对一种特定情境, 当已有观测数据受到限制, 数据稀缺或质量不高的实际情况下, 能够借助文中提出的方法来有效解决问题。这样的简化不仅有助于提升方法的普适性和易用性, 还能促进其在不同领域和场景下的广泛应用。

假设: 降水场结构仅考虑降水场的梯度, 两部雷达资料区域均为正方形, 来自同一种降水观测手段的权重是相等的。

### 2.2. 符号说明

本文所需的变量及其含义如表 1 所示。

**Table 1.** Variables and their meanings  
**表 1.** 变量及其含义

符号	含义说明
$x$	需要优化的降水场
$x_b$	雷达观测区域的背景场
$x_c$	雨量计的观测数据
$B$	背景误差协方差矩阵
$y_1, y_2$	左雷达和右雷达的观测数据
$H_1, H_2$	左雷达和右雷达的观测算子
$R_1, R_2$	左雷达和右雷达的观测误差协方差矩阵
$\sigma$	雷达的相对误差的均方根
$\sigma_{grad}$	雷达的相对梯度误差的均方根

### 3. 构建双雷达变分同化数学模型

#### 3.1. 构建背景误差协方差矩阵 $B$

为了精确估算背景误差协方差矩阵  $B$ , 可以利用背景 1 中两个独立雷达观测数据间的差异作为依据, 选取误差相对较小的雷达观测数据作为背景场, 构建背景误差协方差矩阵, 更精准地描述观测目标的背景状态不确定性。具体计算步骤如下公式所示:

误差:

$$\text{error} = \text{left}_{\text{rader}} - \text{right}_{\text{rader}} \quad (1)$$

误差均值:

$$\text{error}_{\text{mean}} = \frac{1}{n} \sum_{i=1}^n \text{error}_i \quad (2)$$

误差协方差:

$$C(\text{error}) = (\text{error} - \text{error}_{\text{mean}})(\text{error} - \text{error}_{\text{mean}})^T \quad (3)$$

添加正则化项:

$$B = C(\text{error}) + \lambda I \quad (4)$$

其中  $I$  表示单位矩阵,  $\lambda$  表示正则化因子, 添加正则化项的目的是确保  $B$  是正定矩阵, 避免在矩阵求逆时遇到问题[10]。

#### 3.2. 目标函数的确定

通过将两部具有不同探测精度雷达对同一区域的观测数据分别作为观测数据和背景场数据, 最小化一个目标函数, 从而得到最优的降水场估计[6], 目标函数表达式如下:

$$J(x) = \frac{1}{2}(x - x_b)^T B^{-1}(x - x_b) + \frac{1}{2c_1}(y_1 - H_1 x)^T (y_1 - H_1 x) + \frac{1}{2c_2}(y_2 - H_2 x)^T (y_2 - H_2 x) + \frac{1}{2c_3}(\nabla y_1 - \nabla x)^T (\nabla y_1 - \nabla x) + \frac{1}{2c_4}(\nabla y_2 - \nabla x)^T (\nabla y_2 - \nabla x) \quad (5)$$

其中,  $x$  是需要优化的降水场,  $x_b$  是背景场,  $B$  是背景误差协方差矩阵,  $y_1$  和  $y_2$  分别是观测场和背景场的观测数据,  $H_1$  和  $H_2$  是观测算子, 假设为单位矩阵,  $R_1$  和  $R_2$  是观测场和背景场的观测误差协方差矩阵,  $c_1$ 、 $c_2$ 、 $c_3$ 、 $c_4$  分别是观测场的观测误差和梯度误差。

目标函数的梯度函数定义为:

$$\begin{aligned} \nabla J(x) = & B^{-1}(x - x_b) - \frac{1}{c_1} H_1^T (y_1 - H_1 x) - \frac{1}{c_2} H_2^T (y_2 - H_2 x) \\ & - \frac{1}{c_3} (\nabla x - \nabla y_1) - \frac{1}{c_4} (\nabla x - \nabla y_2) \end{aligned} \quad (6)$$

### 3.3. 优化迭代算法 L-BFGS-B (Limited-Memory Broyden-Fletcher-Goldfarb-Shanno with Box Constraints)

L-BFGS-B [11]用来最小化目标函数, 从而找到最优的降水场估计。其基本思想是通过近似目标函数的二阶导数(即 Hessian 矩阵)来加速收敛过程。本文设置 L-BFGS-B 算法的最大迭代次数为 100、梯度容忍度为  $1e-5$ 。

### 3.4. 模型评价指标

1) 观测误差:

双雷达的相对误差的根方差  $\sigma$  :

$$\sigma = \sqrt{\frac{1}{N} \sum \frac{(x_{\text{real}} - x_{\text{observe}})^2}{x_{\text{real}}^2}} \quad (7)$$

其中  $N$  为观测区域观测点个数,  $x_{\text{real}}$  是观测区域的实际降水量,  $x_{\text{observe}}$  是观测区域的观测降水量。由于观测场范围很大, 降水量数据很多, 所以观测数据准确时计算出的相对误差的根方差的应当贴近  $\sigma$ , 因此可得:

$$P_1 = \frac{1}{2} \left( \sqrt{\frac{1}{N} \sum \frac{(x_{\text{left}} - y_1)^2}{x_{\text{left}}^2}} - \sigma_{\text{left}} \right)^2 + \frac{1}{2} \left( \sqrt{\frac{1}{N} \sum \frac{(x_{\text{right}} - y_2)^2}{x_{\text{right}}^2}} - \sigma_{\text{right}} \right)^2 \quad (8)$$

其中  $x_{\text{left}}$ ,  $x_{\text{right}}$  分别表示同化后降水场的左区域和右区域,  $\sigma_{\text{left}}$ ,  $\sigma_{\text{right}}$  分别表示左雷达和右雷达的观测误差。

2) 梯度误差:

梯度误差反映的是降水场的结构特征, 模型同化后的降水场的结构与观测场的结构的相对误差的根方差也应当尽可能地贴近雷达的梯度误差的根方差:

$$P_2 = \frac{1}{2} \left( \sqrt{\frac{1}{N} \sum \frac{(\nabla x_{\text{left}} - \nabla y_1)^2}{\nabla x_{\text{left}}^2}} - \sigma_{\text{grad, left}} \right)^2 + \frac{1}{2} \left( \sqrt{\frac{1}{N} \sum \frac{(\nabla x_{\text{right}} - \nabla y_2)^2}{\nabla x_{\text{right}}^2}} - \sigma_{\text{grad, right}} \right)^2 \quad (9)$$

其中  $\sigma_{\text{grad, left}}$ ,  $\sigma_{\text{grad, right}}$  分别为左雷达和右雷达的梯度误差。

3) 精度表达式:

$$P = 1 / \left( 1 + \frac{1}{2} P_1 + \frac{1}{2} P_2 \right) \quad (10)$$

综合观测误差和梯度误差得出最终的精度表达式，精度越高说明估测结果越准确。

## 4. 构建雷达 - 雨量计变分同化数学模型

### 4.1. 模型的建立

雨量计能够提供高精度的点雨量数据，但空间覆盖率有限，雷达能够实时探测一定区域内的降水情况，空间分辨率较高，但数据精度可能受到多种因素的影响。为了解决单一观测手段的缺陷，本文基于双雷达变分同化数学模型构建了一个能够高效整合雷达和雨量计多源观测数据的变分同化模型，构建过程如下：

1) 数据预处理，通过对需要优化的降水场  $x$  进行抽样，来同化雨量计和雷达的空间分辨率，提升降水估测的精度和可靠性；

2) 构建背景误差协方差矩阵  $B$ ，见双雷达变分同化数学模型构建过程；

3) 确定目标函数，假设目标函数中除雨量计外的项权重相同，限制权重之和为 1 且最多为两位小数，雷达 - 雨量计变分同化模型目标函数如公式(10)所示：

$$\begin{aligned}
 J(x) = & \omega_1 (x - x_b)^T B^{-1} (x - x_b) \\
 & + \omega_2 \left[ \frac{1}{2c_1} (y_1 - H_1 x)^T (y_1 - H_1 x) + \frac{1}{2c_2} (y_2 - H_2 x)^T (y_2 - H_2 x) \right] \\
 & + \omega_3 \left[ \frac{1}{2c_3} (\nabla y_1 - \nabla x)^T (\nabla y_1 - \nabla x) + \frac{1}{2c_4} (\nabla y_2 - \nabla x)^T (\nabla y_2 - \nabla x) \right] \\
 & + \omega_4 \sum_{i=1}^N (x_b - x_c)^2
 \end{aligned} \tag{11}$$

4) 自适应交替优化算法

自适应交替优化算法[12]通过交替更新背景场和权重系数来迭代逼近最优解，初始化降水场和权重，固定权重，优化降水场，比较不同权重下目标函数优化后的最终值，选择最小值的权重为最佳权重  $\omega^*$ ，计算公式如下所示：

$$x^* = \arg \min_x J(x, \omega^{(k)}) \tag{12}$$

其中  $x$  为降水场。

$$J_{min} = J(x, \omega^*) \tag{13}$$

5) 反演

使用 matlab 工具对最终数据进行反演，分析比较反演结果。

### 4.2. 模型的检验与分析

1) 双雷达检验

为了提高模型的泛化性，减少单次观测误差带来的偏差，需要进行观测误差调整，多次模拟观测，每次引入一个随机误差项，取这些模拟的平均值作为背景场。

$$D_{sim,i} = D \times (1 + \epsilon_i) \tag{14}$$

其中  $D$  为雷达观测数据， $\sigma$  为观测误差头  $\epsilon_i$  是服从正态分布  $\mathcal{N}(0, \sigma)$  的随机误差项。

为了更好地反映降水结构，需要进行梯度误差调整，让背景场很好地反映出降水的梯度特征。

$$G_m = \frac{\partial x_b}{\partial m}, \quad G_n = \frac{\partial x_b}{\partial n} \quad (15)$$

其中  $m, n$  分别为纬度索引和经度索引,  $G_m, G_n$  分别为观测误差调整后的背景场对  $m, n$  的偏导数。

$$G'_m = G_m \times (1 + \epsilon_{grad,i}), \quad G'_n = G_n \times (1 + \epsilon_{grad,i}) \quad (16)$$

其中  $\sigma_{grad}$  为降水场结构的估计误差,  $\epsilon_{grad,i}$  是服从正态分布  $\mathcal{N}(0, \sigma_{grad})$  的随机误差项, 使用积分的方式对  $m, n$  方向上做梯度的修正:

$$x'_{sim,i}(m, n) = \frac{1}{2} [B'(m-1, n) + G'_m(m, n) + B'(m, n-1) + G'_n(m, n)] \quad (17)$$

其中  $x'_{sim,i}(m, n)$  为背景场中各个点经过梯度误差调整的数值, 最后进行  $N$  次模拟取平均:

$$x_b = \frac{1}{N} \sum_{i=1}^N x'_{sim,i} \quad (18)$$

使用 matlab 画图得到对误差调整后的雷达观测数据进行反演, 结果如图 1 所示。

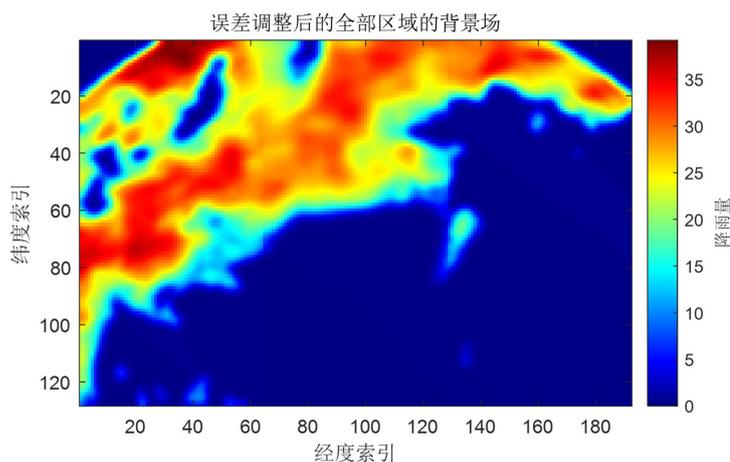


Figure 1. Error-adjusted radar background field  
图 1. 误差调整后的雷达背景场

将所得到的数据代入双雷达变分同化数学模型, 经变分同化模型优化后的降水场如图 2 所示:

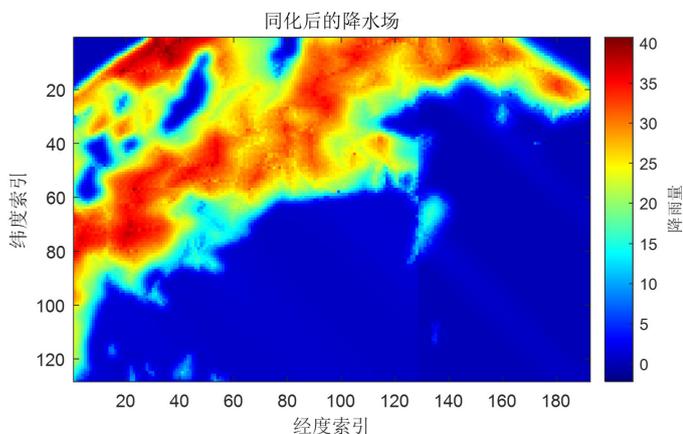


Figure 2. Precipitation field after variational assimilation  
图 2. 变分同化后的降水场

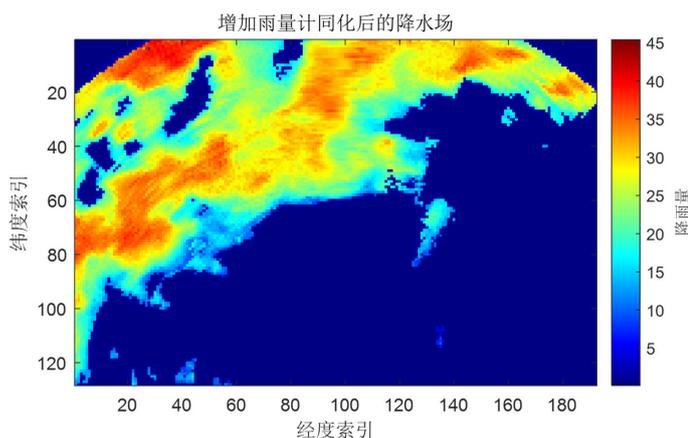
依据背景 1 的数据, 将同化后的降水场矩阵代入到精度表达式中, 得到该降水场的精度为 0.654007, 即降水场与观测场的相关误差的均方根极为贴近雷达的真实误差的均方根, 可以得出本文提出的变分同化模型可以很好地同化双雷达的数据, 反演出较为真实的降水场。由图 1 和图 2 对比可得, 图 1 边界相对较为粗糙, 存在着较为明显的跳跃与突变, 这在一定程度上反映了数据在空间连续性上的不足。然而, 在图 2 中, 经过同化技术处理后的降水场, 其边界展现出了平滑性, 这种平滑并非简单的模糊处理, 而是在保持降水场内部结构特征完整性的基础上, 实现了数据间的自然过渡与融合。

## 2) 雷达 - 雨量计检验

在双雷达数据的基础上增加雨量计观测数据, 通过联立方程:

$$\begin{cases} \omega_1 + \omega_2 + \omega_3 + \omega_4 = 1 \\ \omega_1 = \omega_2 = \omega_3 \\ \omega_4 > \omega_1 \\ \omega \text{ 最多有 2 位小数} \end{cases} \quad (19)$$

由公式(19)可得  $\omega$  的集合, 使用 L-BFGS-B 优化算法对不同的目标函数进行优化求解, 得到使得目标函数最小的权重, 将其代入雷达 - 雨量计变分同化数学模型, 最终得出最优权重下的反演场如图所示:



**Figure 3.** Precipitation field assimilated by increasing rain gauges  
**图 3.** 增加雨量计同化后的降水场

由图 2 和图 3 对比可得, 在增加了雨量计数据之后, 优化后的降水场通过平缓的过渡带展现了从高强度降水(红色区域)到低强度或无降水(蓝色区域)的自然过渡, 有效减少了预测中的不合理突变, 提升了预测结果的可信度和实用性。

在深入分析降水预测模型的性能提升过程中, 将经过同化处理的降水场数据代入精度评估表达式(10)中, 计算结果显示, 引入雨量计数据后的降水场精度提升至 0.759577, 相较于未采用雨量计数据时的降水场精度, 提升了 0.105570, 实现了 16.1% 的增长。这一数据直观地反映了雷达融合雨量计观测数据在提升降水预测精准度方面的重要作用。

综上所述, 将雨量计数据纳入降水预测模型的同化过程中, 是提升模型预测精度、优化降水场空间分布特征的有效途径。这一做法不仅增强了模型对复杂降水现象的捕捉能力, 也为相关领域的研究和应用提供了更加精准的数据支持。

## 5. 总结

为提升降水场估测精度, 降低灾害发生风险, 本文构建了双雷达变分同化模型以及雷达 - 雨量计变

分同化模型, 设计不同来源数据的权重系数选择模型, 并进行反演实验比较, 进一步验证了雷达雨量计变分同化模型对于降水估计准确率的优势。该模型能有效提高降水场估测精度, 满足提高气象预报的准确性和提升灾害预警能力的要求, 并为未来研究提供了参考价值。

## 参考文献

- [1] 赵坤, 葛文忠, 刘国庆, 李致家. 在雷达测雨和洪水预报中自适应卡尔曼滤波法的应用[J]. 高原气象, 2005, 24(6): 956-965.
- [2] Wei, M., Liu, G.-Q., Wang, C.-G., Ge, W.-Z. and Xu, Q. (2003) Optimal Selection for the Weighted Coefficients of the Constrained Variational Problems. *Applied Mathematics and Mechanics*, **24**, 936-944. <https://doi.org/10.1007/bf02446499>
- [3] Mu, X., Cheng, H., Xu, Q. and Liu, G. (2020) Stably Numerical Solving Inverse Boundary Value Problem for Data Assimilation. *Applied Mathematical Modelling*, **82**, 825-835. <https://doi.org/10.1016/j.apm.2020.01.071>
- [4] 张欣, 赵洪岩, 梁海雁, 等. 称重式与虹吸式雨量计观测雨量精度对比分析[J]. 中国防汛抗旱, 2020, 30(3): 32-36.
- [5] 雷梦佳, 易刚, 彭英杰, 等. 新型雨量计的设计和比测实验[J]. 中国防汛抗旱, 2023, 33(3): 67-71.
- [6] 张立凤. 暴雨数值预报若干关键技术发展的回顾与思考[J]. 暴雨灾害, 2024, 43(3): 243-254.
- [7] 王文丽. 雷达估测降雨在碧流河洪水预报中的应用研究[D]: [硕士学位论文]. 大连: 大连理工大学, 2020.
- [8] 黄静. 基于实时天气背景的雷达反射率因子同化及对暴雨预报的影响研究[D]: [硕士学位论文]. 南京: 南京信息工程大学, 2023.
- [9] 张新. 卡尔曼滤波算法在辽宁东部测雨雷达降水同化适用性分析[J]. 水利技术监督, 2024(6): 246-249.
- [10] 李梦迪, 戚友存, 张哲, 等. 基于雷达-雨量计降水融合方法提高极端降水监测能力[J]. 大气科学, 2022, 46(6): 1523-1542.
- [11] 汪卫国. 多波段雷达径向速度质量控制与循环同化研究[D]: [硕士学位论文]. 南京: 南京信息工程大学, 2024.
- [12] 赖锡军, 何国建. 基于四维变分同化法的巢湖流域南淝河水质模拟[J]. 湖泊科学, 2021, 33(5): 1458-1466.