

文章编号: 1672-8785(2025)09-0032-09

基于 HO-RF 算法的大气 透过率求解研究

傅 莉^{1*} 张 帅¹ 邓洪伟²

(1. 沈阳航空航天大学自动化学院, 辽宁 沈阳 110136;

2. 中国航发沈阳发动机研究所, 辽宁 沈阳 110015)

摘 要: 为提升红外辐射测试中大气透过率参数的获取效率与精度, 提出了一种基于 HO-RF 算法的大气透过率求解模型。以温度、湿度、作用距离和大气压等关键环境因素作为输入, 结合实测数据建立回归求解模型, 从而实现大气透过率的快速准确求解。仿真结果表明, 与传统反向传播(Back Propagation, BP)神经网络及随机森林(Random Forest, RF)模型相比, 所提模型在求解精度方面表现更佳, 均方根误差(Root Mean Square Error, RMSE)降低至 0.010745, R^2 值为 0.95877, 平均绝对误差(Mean Absolute Error, MAE)为 0.0080021。该模型能够有效降低实验操作的复杂性, 在精度、稳定性和可靠性方面均优于传统方法, 可提高战机的红外特性测试效率, 具有良好的实际应用价值。

关键词: 大气透过率求解; 河马优化算法; HO-RF 模型; 红外辐射测试

中图分类号: V218 **文献标志码:** A **DOI:** 10.3969/j.issn.1672-8785.2025.09.005

Study on Atmospheric Transmittance Calculation Based on HO-RF Algorithm

FU Li^{1*}, ZHANG Shuai¹, DENG Hong-wei²

(1. School of Automation, Shenyang Aerospace University, Shenyang 110136, China;

2. Shenyang Engine Research Institute of AECC, Shenyang 110015, China)

Abstract: To improve the efficiency and accuracy of atmospheric transmittance parameter acquisition in infrared radiation testing, a model for atmospheric transmittance calculation based on the HO-RF algorithm is proposed. Using key environmental factors such as temperature, humidity, operating distance, and atmospheric pressure as inputs, a regression model is established based on the measured data, enabling rapid and accurate atmospheric transmittance calculation. The simulation results show that the proposed model outperforms traditional back propagation (BP) neural network and random forest (RF) models in terms of accuracy, with a root mean square error (RMSE) reduced to 0.010745, an R^2 value of 0.95877, and a mean absolute error (MAE) of 0.0080021. This model effectively reduces experimental complexity and outperforms traditional methods in

收稿日期: 2025-03-30

基金项目: 国家自然科学基金项目(61602321)

作者简介: 傅莉(1968-), 女, 辽宁人, 教授, 主要研究方向为飞行器隐身测试与智能控制以及磁流变液应用技术。

*通讯作者: E-mail: ffulli@163.com

terms of accuracy, stability, and reliability. It can improve the efficiency of infrared characteristic testing for fighter aircraft and has excellent practical application value.

Key words: atmospheric transmittance calculation; Hippo optimization algorithm; HO-RF model; infrared radiation testing

0 引言

红外隐身技术的核心是通过各种措施尽量降低战机在威胁方向上的红外辐射特性,大幅减小敌方红外辐射探测系统发现、识别和锁定战机的可能性^[1]。

在红外辐射测试中,大气透过率作为关键影响因素,对于获取准确测量结果具有重要意义^[2]。由于大气介质的存在,几乎所有实验数据都会受到大气窗口特性的干扰。在红外辐射传播过程中,水蒸气、臭氧、二氧化碳、甲烷以及氮氧化物等温室气体会对特定波长的红外信号产生不同程度的吸收。同时,大气中还含有多种形态各异的粒子物质,包括冰晶、尘埃、云雾等。这些非均匀分布的粒子会引起辐射信号的散射效应。综上所述,大气对红外辐射呈现出波长选择性吸收与散射特征,最终导致辐射能在传输过程中不同程度地衰减^[3]。

目前有多种计算大气透过率的软件。自 20 世纪 80 年代以来,美国空军陆续开发了多种辐射传输仿真工具,如 LOWTRAN^[4]和 MODTRAN^[5]等。这类软件能够考虑多种微量气体的影响,具有较高的波段分辨率;在已知大气廓线信息的条件下,可实现较精确的计算。因此,有必要构建一种适用于工程应用、具备快速响应能力的大气透过率计算方法。

李飞^[6]对中长波红外辐射在大气传播过程中的特性进行了研究,重点分析了其与环境温度、相对湿度、大气高度及二氧化碳浓度等因素之间的关联性。吴辉阳等人^[7]针对大气中气溶胶对中长波红外辐射的散射效应进行了研究,推导了基于斜程等效近似的大气传输模型,并提出了适于工程应用的近似计算公式。陈芳芳等人^[8]提出了一种用于激光红外波段的大气透过率计算模型。与传统软件模拟方法相

比,该模型在保证计算精度的同时显著降低了计算量,且最小相对误差仅为 0.005%。张炳炎^[9]通过重新计算大气中的二氧化碳浓度,对 LBLRTM 辐射传输模型进行了改进,从而提升了模型的计算精度与预测可靠性。张骏等人^[10]指出,中段飞行阶段的导弹目标可近似看作点源。他们在此基础上提出一种基于空间辐射矩阵的算法来计算红外辐射特性。辛琪^[11]借助 MODTRAN 辐射传输模型,对中低层大气的红外辐射特性进行了系统分析;结合地表类型、高度变化以及季节差异等多个环境变量,探讨了在 3~5 μm 和 8~12 μm 两个典型红外波段内地球临边视角的大气辐射表现特征。

随着科技的发展,数据快速积累,计算机算力大幅提升。这直接促进了机器学习在各行各业的应用实践。目前,机器学习在大气透过率预测方面的应用较少。因此在现有研究的基础上,利用机器学习建立模型。本文提出一种基于 HO-RF 算法的大气透过率预测模型,并利用河马优化算法优化随机森林中的参数决策树数目和叶子数,从而提高预测模型的精度。

1 大气透过率数据获取

1.1 测量原理

当目标的红外辐射强度 I_λ 已知时,结合辐射计采集的电压信号数据,可计算出目标与辐射计之间传输路径的大气透过率:

$$\tau_\lambda = \frac{V \cdot R^2}{K \cdot I_\lambda} \quad (1)$$

式中, V 表示辐射计输出的电压信号; R 为待测目标中心与辐射计镜头中心之间的水平距离; K 为辐射计的响应度系数,用于表征其对辐射信号的灵敏度。

假设目标在镜头视场中的投影面积为 S ,

相应温度下的黑体辐射亮度为 $L_{bb}(T)$ ，则其辐射强度可表示为

$$I_{\lambda} = \frac{L_{bb}(T) \cdot S}{\tan^2(\theta/2)} \quad (2)$$

由于本实验中测量方向与黑体表面法线一致，即夹角 $\theta = 90^\circ$ ，对应 $\tan^2(\theta/2)$ 的值为 1。因此结合式(1)和式(2)，可进一步推导出大气透过率的表达式：

$$\tau_{\lambda} = \frac{V \cdot R^2}{L_{bb}(T) \cdot S \cdot K} \quad (3)$$

1.2 测量流程

在实验过程中，采用 MR170 型傅里叶光谱辐射计对不同距离的标准黑体进行观测，以获取相应的大气透过率。实验流程如下：

(1)利用温度计、湿度计与压力计记录实验时的大气环境参数。

(2)将标准黑体加热至 100°C ，并基于双温标定法对辐射计进行响应校准。

(3)使用完成标定的辐射计对黑体进行观测，获取对应的电压信号。

(4)确保辐射计与黑体中心位于同一水平线，通过激光测距仪调整黑体与辐射计的间距。在 $5 \sim 50\text{ m}$ 范围内每 5 m 设一个测距点，采集各间距下的红外辐射数据。图 1 展示了间距为 10 m 时的双通道电压测量结果。

(5)结合辐射计响应度及实测数据，利用式(3)计算各测距条件下的大气透过率。

要获得不同时刻的大气透过率，仅用距离作为唯一变量进行一维拟合难以满足精度要求，因此还需引入环境参数作为辅助变量，以提高模型的准确性。在使用辐射计测量大气透过率的过程中，通常需频繁进行设备标定与位置调整。这在一定程度上增加了实验的工作量。

针对上述问题，本实验在测量流程的基础上，在沈阳进行实地测试。选取夏季与冬季两个典型季节，每季采集 8 天的大气透过率数据。每天在早、中、晚三个时段各采集 10 组数据(共计 30 组/日)，最终累计获得 480 组观测样本。测量过程中同步记录温度、湿度、气压等环境条件，并将上述信息作为后续大气透过率预测模型的数据集。

本文利用所测得的 480 组大气透过率数据建立数据集。根据步骤(1)划分训练集与测试集。输入值为温度($^\circ\text{C}$)、湿度(%）、气压(atm)和作用距离(m)，输出值为 $3 \sim 5\text{ }\mu\text{m}$ 、 $8 \sim 14\text{ }\mu\text{m}$ 波段的大气透过率数值。表 1 列出了部分训练集数据。

受实验环境的限制，本文所建立预测模型中输入变量的取值范围限定如下：温度在 $-14.3 \sim 27.4^\circ\text{C}$ 之间，相对湿度在 $35\% \sim 97\%$ 之间，大气压强则分布在 $0.977 \sim 1.014\text{ atm}$ 范围内。

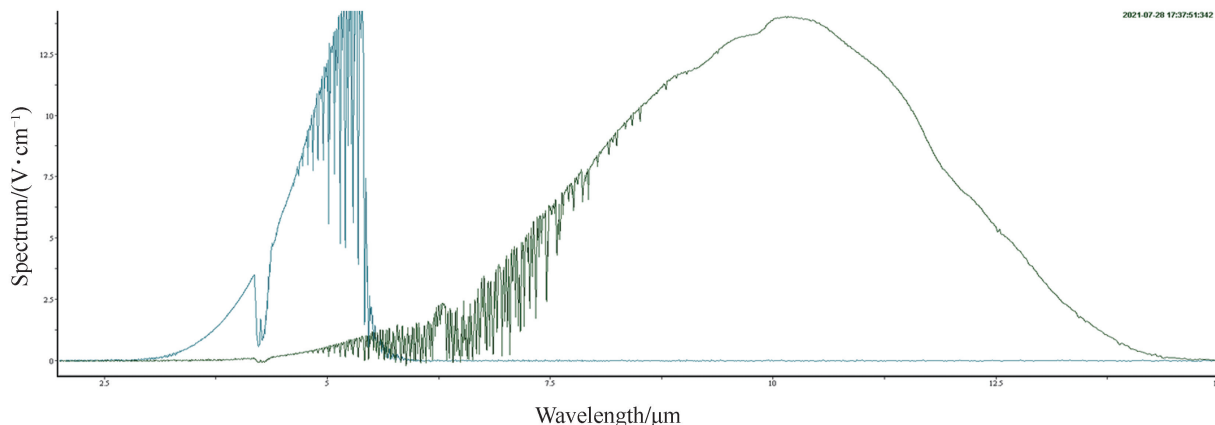


图 1 间距为 10 m 时的双通道电压测量结果

表 1 部分训练集数据

温度/℃	湿度/%	距离/m	大气压/atm	大气透过率
16.7	56	5	1.009	0.93645
22.6	76	10	0.99	0.92632
10.4	43	15	0.988	0.78621
29.7	88	35	0.975	0.83983
1.2	65	30	1.003	0.83825
-11.3	46	25	0.991	0.85011
31.8	87	35	0.989	0.86387
8.6	54	40	1.002	0.81374
7.5	42	25	1.034	0.86981
30.4	86	20	1.006	0.88767

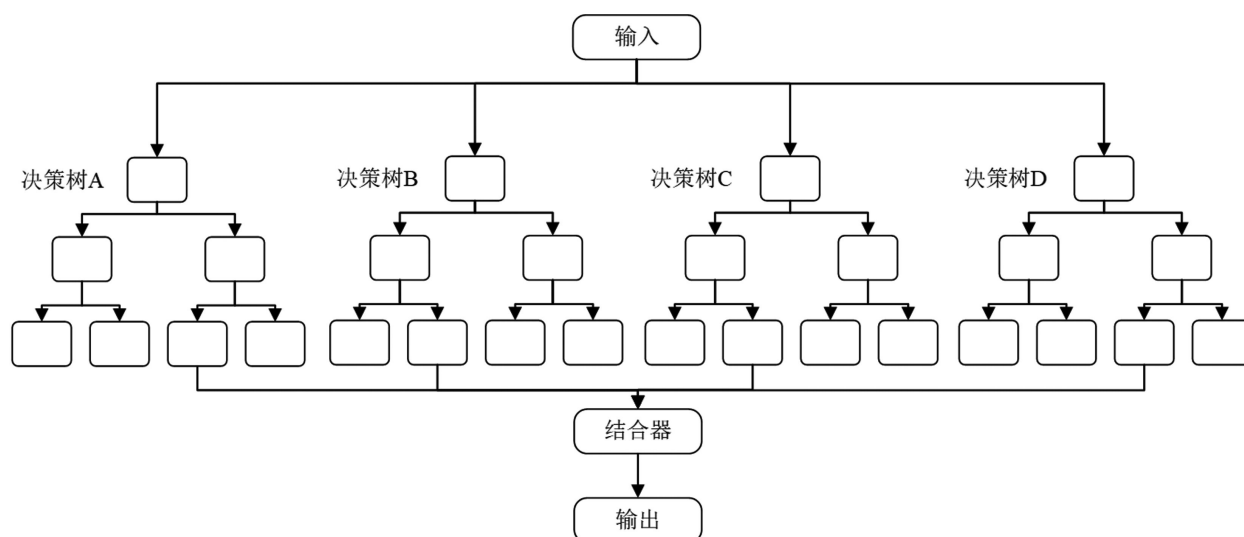


图 2 RF 算法的结构示意图

2 基于 HO-RF 算法的大气透过率预测模型

2.1 RF 算法

RF 算法是一种典型的集成学习方法, 由 Breiman L 于 2001 年提出^[12]。该方法并非独立的机器学习算法, 而是基于决策树构建的集成框架。其核心思想是通过集成多个弱学习器来降低模型误差。在具体实现中, 部分基学习器可能会出现判断偏差, 但这种误差可通过其他学习器的集体决策进行修正或平均, 从而有效缓解过拟合问题, 增强模型的泛化性能, 并提升整体预测精度。

RF 算法以决策树作为基本学习单元, 通过集成多个决策树来构建模型。其实现过程中

采用 Bagging 技术进行样本重采样, 生成多个训练子集; 随后, 在每棵决策树的构建过程中, 引入属性随机选择机制。该算法在构建决策树的过程中所体现的“随机性”主要来源于两个方面: 一为样本的随机抽样, 即通过对训练数据进行多次有放回的采样, 生成多个不同的训练子集, 用以构建结构各异的决策树; 二为特征的随机选取, 即在每个节点进行划分时, 仅从所有特征中随机选取部分子特征进行评估与分裂。上述机制有助于提升模型的多样性, 进而有效缓解单棵决策树容易出现的过拟合问题^[13]。

RF 算法通过对训练样本和特征同时进行随机采样, 有效增强了决策树之间的独立性, 从而提高了模型输出结果的稳定性与无偏性^[14]。

在该方法中,由多个决策树组成的集成体对输入样本分别进行判断,最终将各树的判别结果进行汇总,并采用投票机制确定模型输出。由于模型中包含大量相互独立的决策树,且每个样本均参与预测过程,只有在超过半数的决策树作出错误判断的情况下,整体模型才可能出现误判,因此该方法在准确率方面具有显著优势。图 2 为 RF 算法的结构示意图。

2.2 HO-RF 算法

HO 算法由 Amiri M H 等人^[15]在 2024 年提出,其灵感来源于河马的三种行为模式:(1)小河马由于好奇心而产生偏离群体的倾向;(2)河马的防御性行为(当受到捕食者攻击或其他生物侵入领地时触发);(3)河马逃离捕食者的行为。HO 算法采用三阶段策略建模,模拟了河马在水体中的位置更新行为、应对捕食威胁的防御机制以及逃逸过程。该方法通过自适应机制动态调节搜索空间的分辨率与搜索步长,从而在保证解的精度同时加快收敛速度,展现出良好的全局优化能力和求解精度。

在 HO 算法中,个体河马被视为优化问题的候选解,其在搜索空间中的位置对应于决策变量的具体取值。因此,每一头河马可用一个向量表示,整个种群在数学上则构成一个矩阵。类似于大多数智能优化算法,HO 算法在初始化阶段通过随机生成一组初始解来构建种群,为后续迭代过程提供起点。在此步骤中,利用式(4)生成决策变量的向量:

$$X_i: x_{ij} = lb_j + r \cdot (ub_j - lb_j), \quad (i = 1, 2, \dots, N, \quad j = 1, 2, \dots, m) \quad (4)$$

式中, X_i 表示第 i 头河马在每个决策变量下的所有位置信息,它是一个行向量; x_{ij} 表示第 i 头河马在第 j 个决策变量下的位置信息,即 x_{ij} 表示一个候选解; r 为取值范围在 $[0, 1]$ 之间的随机数; lb_j 与 ub_j 分别表示第 j 个决策变量的下界与上界; N 表示种群中河马个体的数量; m 为待优化问题的决策变量维数。

河马种群位置信息的矩阵为

$$X = \begin{bmatrix} X_1 \\ \vdots \\ X_i \\ \vdots \\ X_N \end{bmatrix}_{N \times M} = \begin{bmatrix} x_{1,1} & \cdots & x_{1,j} & \cdots & x_{1,m} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{i,1} & \cdots & x_{i,j} & \cdots & x_{i,m} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{N,1} & \cdots & x_{N,j} & \cdots & x_{N,m} \end{bmatrix}_{n \times m} \quad (5)$$

在第一阶段,水域环境中河马个体的空间更新行为:河马群体活动于河流或池塘等水域环境中。个体间表现出明显的聚集倾向,通常会趋向于彼此靠近以形成群体结构。通过对个体适应度(即目标函数值)的持续迭代,可识别出具有相对优势的河马个体。此阶段,群体中雄性河马的空间位置更新公式为

$$X_i^{Mhippo}: X_{ij}^{Mhippo} = x_{ij} + y_1 \cdot (D_{hippo} - I_1 \cdot x_{ij}), \quad (i = 1, 2, \dots, \frac{N}{2}, \quad j = 1, 2, \dots, m) \quad (6)$$

式中, X_i^{Mhippo} 代表雄性河马的位置; y_1 是 0 与 1 之间的随机数; D_{hippo} 代表具有支配地位河马(在当前迭代中具有最佳函数值的河马)的位置; I_1 等于 1 或 2。

用式(7)表示湖泊或池塘中雌性或未成熟河马的位置信息:

$$X_i^{FBhippo}: X_{ij}^{FBhippo} = \begin{cases} x_{ij} + h_1 \cdot (D_{hippo} - I_2 \cdot MG_i), & T > 0.6 \\ \Xi, & \text{else} \end{cases} \quad (7)$$

式中, h_1 和 h_2 是从 h 方程中随机选择的数或向量; I_2 为 1 与 2 之间的随机数; T 表示选择概率; MG_i 为一些从河马群中随机选择的河马的平均值。

在第二阶段,河马防御捕食者:当捕食者或其他异种个体入侵河马所占据的水域时,河马将启动本能的防御行为以捍卫其领地安全。此时,河马通常通过张开巨大的下颚并发出震慑性的低频咆哮来驱逐入侵者,从而构成其典型的威慑反应机制。在建模过程中,外来入侵者的位置如式(8)所示。

$$X_i^{HippoR} : X_{ij}^{HippoR} = \begin{cases} \overline{RL} \oplus Predator_j + (\frac{f}{(c-d \times \cos(2\pi g))}) \cdot (\frac{1}{\bar{D}}), F_{Predator_j} < F_i \\ \overline{RL} \oplus Predator_j + (\frac{f}{(c-d \times \cos(2\pi g))}) \cdot (\frac{1}{2 \times \bar{D} + r_{\bar{g}}}), \geq F_i \end{cases}$$

$$(i = [\frac{N}{2}] + 1, [\frac{N}{2}] + 2, \dots, N. \quad j = 1, 2, \dots, m)$$
(8)

式中, X_i^{HippoR} 是河马面对捕食者时的位置; f 是取值范围为 $[2, 4]$ 的均匀分布随机数; c 和 d 分别为 $[1, 1.5]$ 和 $[2, 3]$ 区间内的均匀随机数; g 是在 $[-1, 1]$ 区间内生成的均匀随机变量。变量 $\bar{R}_9$ 表示一个维度为 $1 \times m$ 的随机向量; 而 $\overline{RL}$ 则为服从 Levy 分布的随机向量, 主要用于模拟捕食者在对河马发起攻击时位置的突变行为。

在第三阶段, 河马躲避捕食者: 当警告行为未能有效阻止捕食者入侵时, 处于孤立、生病状态的河马或年幼的河马个体极易成为攻击目标。为规避风险, 这些个体将启动逃逸机制, 主动远离威胁源所在区域。在此阶段, 河马通过感知周围环境和执行局部搜索, 识别并移动至最近的相对安全位置。其所确定的目标安全位置可表示为

$$X_i^{HippoE} : X_{ij}^{HippoE} = x_{ij} + r_{10} \cdot (lb_j^{local} + s_1 \cdot (ub_j^{local} - lb_j^{local}))$$

$$(i = 1, 2, \dots, N. j = 1, 2, \dots, m)$$

$$lb_j^{local} = \frac{lb_j}{t}, ub_j^{local} = \frac{ub_j}{t}, t = 1, 2, \dots, \tau.$$

$$s = \begin{cases} 2 \times \overline{r_{11}} - 1 \\ r_{12} \\ r_{13} \end{cases}$$
(9)

式中, X_i^{HippoE} 是河马逃离捕食者时的位置, 通过搜索在最近处找到安全位置; s_1 是从 s 中随机选择的向量或数值; 集合 s 为算法提供了多样化的局部扰动方式, 有助于提升局部搜索的效率与精度; $\overline{r_{11}}$ 代表 0 到 1 之间的随机向量, r_{10} 和 r_{13} 表示 0~1 范围内产生的随机数, r_{12} 是正态分布的随机数; t 表示当前迭代次数; τ 表示最高迭代次数。

在 HO-RF 算法中, 采用河马优化算法优

化 RF 中的参数决策树数目和叶子数。

3 模型求解结果与分析

3.1 大气透过率求解

为了比较各模型在大气数据上的预测表现, 以 BP、RF 和 HO-RF 三种神经网络模型的预测值输出为基础, 绘制了预测值与真实值对比图(见图 3~图 5)。

3.2 误差分析

为了评估预测结果, 本文引入了当前常用的三类评价指标——MAE、RMSE 以及 R^2 , 以衡量模型的预测精度与拟合效果。三种误差评估模型公式如下:

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{x}_i - x_i|$$
(10)

$$RMSE(X, h) = \sqrt{\frac{1}{m} \sum_{i=1}^m (h(x_i) - y_i)^2}$$
(11)

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y}_i)^2}$$
(12)

利用以上三种评估方法对上述模型的预测精度进行了评估(结果见表 2)。从表 2 中可以看出, 基于 HO-RF 网络对大气透过率的预测结果提升显著, MAE 由 RF 算法的 0.012111 降低至 0.0080021, R^2 值由 RF 算法的 0.91481 逐步提升到 0.95877, RMSE 由 RF 算法的 0.02281 降低至 0.010745。

表 2 三种算法的误差评估结果

算法	MAE	RMSE	R^2
BP	0.014520	0.027397	0.84075
RF	0.012111	0.02281	0.91481
HO-RF	0.0080021	0.010745	0.95877

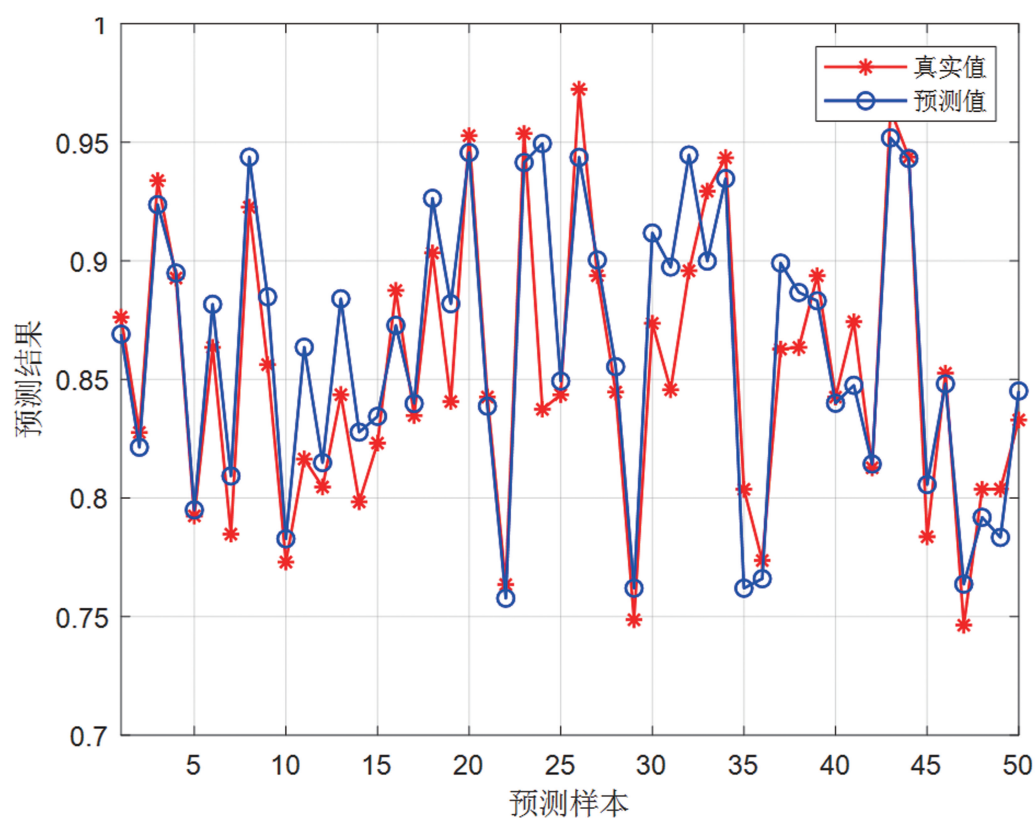


图 3 BP 模型的预测值与真实值对比

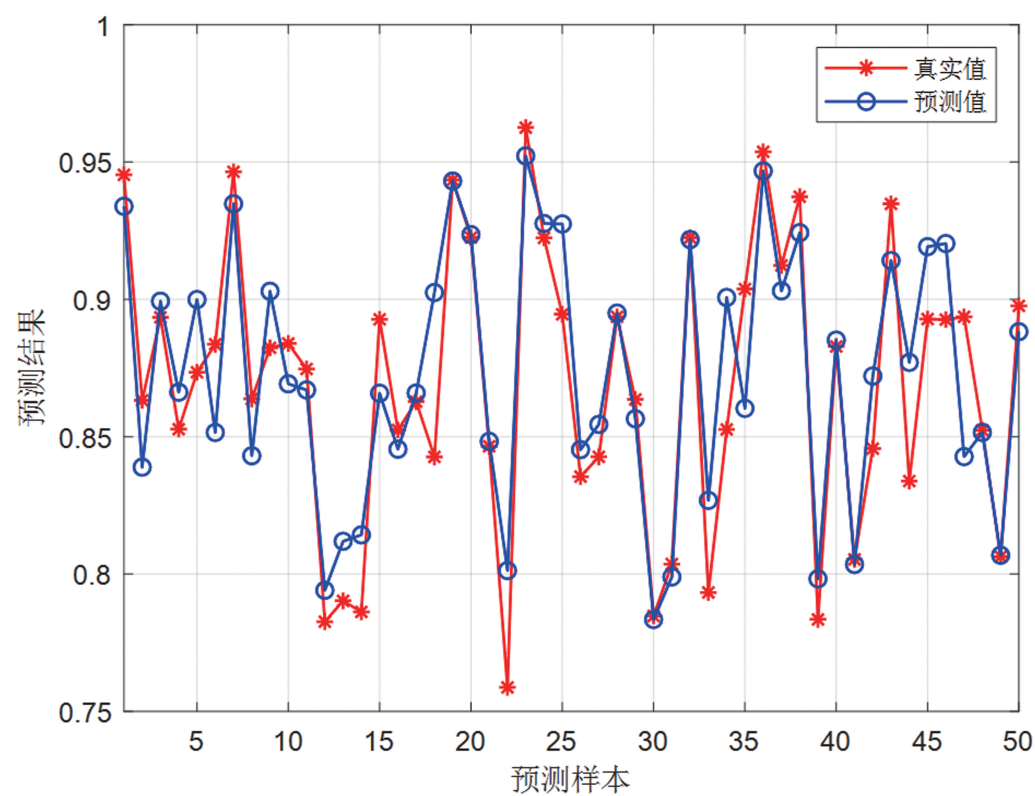


图 4 RF 模型的预测值与真实值对比

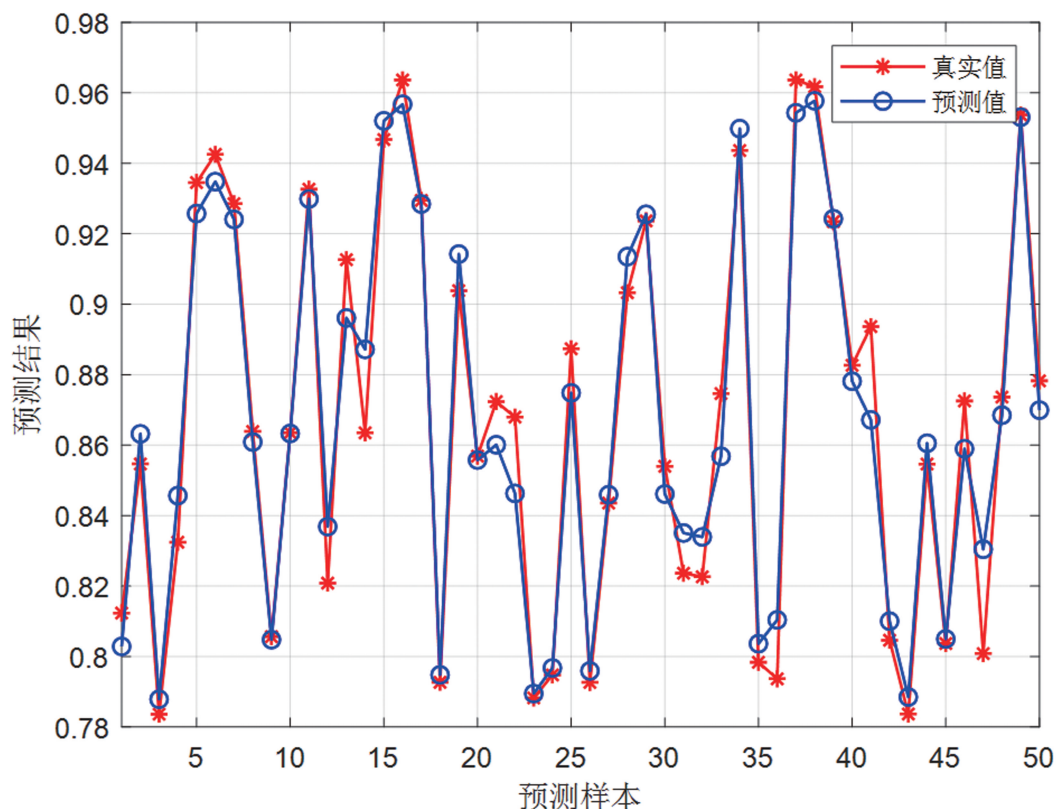


图 5 HO-RF 模型的预测值与真实值对比

综合求解 MAE、RMSE 和 R^2 等评价指标后可知, HO-RF 模型在大气透过率求解中表现最佳, 具有最小的 MSE 与 MAE 以及最大的 R^2 值, 充分体现了其优越的预测精度与拟合能力。

4 结束语

采用传统方法求解不同大气条件下的大气透过率时, 需对辐射计进行反复标定, 其过程相对繁琐。本文首先根据辐射计的测量数据建立大气透过率预测模型数据集, 然后根据温度、湿度、作用距离以及大气压等影响大气透过率的四种因素建立大气透过率求解模型。进一步地, 提出了一种基于 HO-RF 算法的大气透过率求解模型, 并将其分别与 BP 和 RF 两种方法进行对比。

仿真结果表明, 相较于 BP 与 RF 模型, 本文模型在大气透过率求解方面具有更高精度, MAE 降低至 0.0080021, R^2 值提升到 0.95877, RMSE 降低至 0.010745。可以看出, HO-RF 算法具有良好的大气透过率求解能力。

然而本文只根据温度、湿度、作用距离以及大气压四个参数对大气透过率进行求解, 后续可增加烟尘、风速、太阳辐射等进行深入分析。

参考文献

- [1] 彭瑾, 徐兴柱, 宋艳波, 等. 国外导弹隐身技术现状与发展趋势 [J]. 飞航导弹, 2009, 39(3): 23-27.
- [2] 王文慧, 王高, 张猛. 红外辐射大气透过率影响因素的研究 [J]. 激光与红外, 2017, 47(3): 308-312.
- [3] 李惠萍, 周起勃, 匡定波. 实时红外大气衰减计算仿真系统集成 [J]. 半导体光电, 2009, 30(4): 636-638.
- [4] Isaacs R G, Vogelmann A M. Multispectral sensor data simulation modeling based on the multiple scattering LOWTRAN code [J]. Remote Sensing of Environment, 1988, 26(1): 75-99.
- [5] Berk A, Anderson G P, Acharya P K, et al. Validation of MODTRAN® 6 and its line-by-line algorithm [J]. Journal of Quantitative Spectroscopy

- py & Radiative Transfer*, 2017, **203**: 542–556.
- [6] 李飞. 大气传输对中长波红外辐射衰减分析 [J]. *红外技术*, 2019, **41**(4): 311–316.
- [7] 吴辉阳, 王泽洋, 黄兴军, 等. 大气气溶胶红外散射透过率计算研究 [J]. *红外技术*, 2022, **44**(2): 151–156.
- [8] 陈芳芳, 耿蕊, 吕勇. 激光红外大气传输透过率模型研究 [J]. *红外技术*, 2015, **37**(6): 496–501.
- [9] 张炳炎. 大气辐射传输模型 LBLRTM 优化及临近空间大气温度反演初步研究 [D]. 北京: 中国科学院国家空间科学中心, 2020.
- [10] 张骏, 王一程, 吕相银, 等. 大气层外弹道式目标红外特征的数值分析 [J]. *红外与激光工程*, 2008, **37**(5): 765–769.
- [11] 辛琪. 临边大气辐射特性研究 [D]. 哈尔滨: 哈尔滨工业大学, 2014.
- [12] Breiman L. Random forests [J]. *Machine Learning*, 2001, **45**: 5–32.
- [13] 王晖. 基于云平台的信息丢失与突发性负载预测方法研究 [D]. 哈尔滨: 哈尔滨师范大学, 2023.
- [14] 王晓锋. 基于随机森林的跨膜螺旋接触预测与卷曲螺旋寡聚体状态预测 [D]. 北京: 中国农业大学, 2014.
- [15] Amiri M H, Mehrabi H N, Montazeri M, et al. Hippopotamus optimization algorithm: A novel nature-inspired optimization algorithm [J]. *Scientific Reports*, 2024, **14**(1): 5032.