

文章编号 1004-924X(2026)19-2969-15

红外探测器背景辐射的数据驱动实时估算

梅 坚^{1,2}, 杨 林¹, 张荣国^{1,2}, 刘宇阳^{1,2}, 周晓萱¹,
延檬羽³, 陈巧妹^{1*}, 郭梦霏^{1*}

(1. 中国科学院 上海技术物理研究所 红外科学与技术全国重点实验室, 上海 200083;

2. 中国科学院大学, 北京 100049;

3. 西安北方庆华机电有限公司, 陕西 西安 710025)

摘要:高灵敏度红外探测器在轨运行时,其自身背景辐射受复杂时变温度场调制,辐射定标精度被严重制约,而传统基于热平衡假设、地面定标或经验系数的校正方法难以适应动态、非稳态的在轨热环境。为此,本文提出一种探测器背景辐射的高精度实时估算方法。首先,基于在轨温度遥测数据和深空定标观测数据,选取探测器多元耦合温度场作为模型输入,并通过时间匹配、异常值剔除和数据集划分构建模型数据库。然后,利用随机森林算法建立温度场与背景辐射之间的非线性映射模型,突破传统方法依赖静态假设和线性经验关系的局限。最后,面向不同卫星类型和不同红外谱段条件开展模型训练、验证和应用分析,实现复杂在轨热工况下探测器背景辐射亮度的实时估算。实验结果表明,实时估算的验证最高精度 R^2 可达0.985 4, RBias低至0.069%, RMSE低至0.009 6 W/(sr·m²),其在典型地物观测场景中的总平均相对误差仅为3.88%。该方法能够表征复杂温度场与背景辐射之间的非线性耦合关系,为提升红外探测器在轨背景辐射获取能力和辐射定标精度提供了方法途径。

关键词:红外辐射定标;仪器背景辐射;时变温度场;数据驱动;随机森林;实时估算

中图分类号:TP702 **文献标识码:**A

doi:10.3724/OPE.20263419.2969

CSTR:32169.14.OPE.20263419.2969

Real-time data-driven estimation of background radiation for infrared detectors

MEI Jian^{1,2}, YANG Lin¹, ZHANG Rongguo^{1,2}, LIU Yuyang^{1,2}, ZHOU Xiaoxuan¹,
YAN Mengyu³, CHEN Qiaomei^{1*}, GUO Mengfei^{1*}

(1. State Key Laboratory of Infrared Physics, Shanghai Institute of Technical Physics,
Chinese Academy of Sciences, Shanghai 200083, China;

2. University of Chinese Academy of Sciences, Beijing 100049, China;

3. Xi'an North Qinghua Electromechanical Co., Ltd., Xi'an 710025, China)

* Corresponding author, E-mail: Mavis6399@163.com; chommy_97@163.com

Abstract: During in-orbit operation, the background radiation of high-sensitivity infrared detectors is modulated by complex, time-varying temperature fields, limiting radiometric calibration accuracy. Conventional correction methods based on thermal equilibrium assumptions, ground calibration, or empirical coefficients

收稿日期:2026-03-23;修订日期:2026-04-25.

基金项目:上海市“科技创新行动计划”启明星资助项目(扬帆专项)(No. 24YF2753800)

cients are poorly suited to dynamic, nonequilibrium orbital thermal environments. To address this limitation, a high-precision method for real-time estimation of detector background radiance is developed. In-orbit temperature telemetry and deep-space calibration observations are used to construct a modeling database through temporal alignment, outlier removal, and dataset partitioning, with coupled temperature measurements from multiple detector-associated components selected as model inputs. A random forest model is then established to capture the nonlinear relationship between these temperature fields and background radiance, overcoming the reliance of conventional methods on static assumptions and linear empirical relationships. Model training, validation, and application assessment are conducted across different satellite types and infrared spectral bands, enabling real-time background radiance estimation under complex orbital thermal conditions. Validation results yield a maximum coefficient of determination (R^2) of 0.985 4, a relative bias (RBias) as low as 0.069%, and a root mean square error (RMSE) as low as 0.009 6 W/(sr·m²). An overall mean relative error of 3.88% is achieved in typical ground-target observation scenarios. These results demonstrate that the nonlinear coupling between complex temperature fields and detector background radiation can be effectively characterized by the proposed method, providing a new approach to in-orbit background radiance estimation and improved radiometric calibration accuracy for infrared detectors.

Key words: infrared radiometric calibration; instrument background radiation; time-varying temperature field; data-driven; random forest; real-time estimation

1 引言

当前,天基红外探测器在轨量化应用面临一个根本挑战,即探测器接收的信号中混有仪器自身热辐射产生的背景噪声^[1]。随着传感器辐射分辨率不断提升,其对自身背景辐射愈发敏感。若不能精确校正,将直接影响系统的最低可探测辐亮度与定标精度^[2]。

探测器背景辐射由卫星在轨的复杂热环境决定。太阳光照和地球红外辐射的周期性交替以及星上设备功耗波动共同导致卫星平台与载荷内部形成动态、非均匀的时变温度场^[3]。该温度场分布受材料热惯性和导热路径影响,在空间上存在梯度,在时间上则表现出明显的周期性和非稳态变化^[4]。仪器背景辐射主要来源于光学器件、支撑结构、探测器封装和电子元件等温度较高的部件,其背景辐射信号会叠加在探测器接收到的目标信号上^[5]。由于卫星温度场随时间呈现非稳态波动,仪器各部件的辐射通量也相应改变,从而引起仪器自身背景辐射的复杂变化^[6]。

当前背景辐射校正主要依赖三类方法:热平衡模型^[7]、地面实验室定标^[8]以及经验系数校正^[9-10]。其中,热平衡模型通常基于仪器结构、材

料热参数和辐射传输关系建立物理描述,理论上具有较强的可解释性,但实际应用中中对边界条件、表面发射率、视场因子及热传导路径等参数依赖较强,模型建立复杂,且难以准确适应在轨非稳态温度场。地面实验室定标能够在受控环境下获得较高精度的初始定标参数,但探测器及光机结构在长期在轨运行过程中会受到空间环境、材料老化和热控状态变化等因素影响,导致地面定标参数难以完全反映在轨真实状态^[11-12]。经验系数校正方法实现较为简单,但多依赖有限工况下获得的温度—响应关系或固定修正系数,对多部件耦合、轨道周期变化和长期漂移的适应能力不足。总体而言,传统方法在静态或准静态条件下具有一定有效性,但在复杂、动态及非均匀的在轨热环境中,仍存在适应性、实时性和泛化能力不足的问题。需要指出的是,完整热平衡模型通常依赖详细的结构参数、材料热物性参数、表面发射率、视场因子及在轨热边界条件,难以在仅有温度遥测和深空观测数据条件下直接复现。地面实验室定标获得的暗电流或背景响应多对应特定温度和工作状态,可作为固定修正量用于初始校正,但难以描述在轨动态温度场驱动下的背景辐射时变过程。因此,本文在后续实

验中采用多元线性回归作为经验系数校正的等效基线,并进一步引入简单机器学习模型开展统一对比。

近年来,随着在轨遥感数据规模的持续增长和机器学习方法的发展,数据驱动模型已逐渐被用于辐射定标、辐亮度偏差订正及传感器响应退化补偿。针对静止气象卫星红外辐亮度资料同化中的系统偏差问题,Zhang 等^[13]以 FY-4A/AGRI 红外通道为对象,引入随机森林方法建立观测辐亮度偏差的非线性订正模型。该研究表明,相比传统线性偏差订正方法,随机森林模型能够更有效表征红外亮温偏差与场景亮温、扫描几何及大气状态因子之间的非线性关系,从而提高红外辐亮度资料在数值天气预报同化中的适用性。Tang 等^[14]将深度学习方法引入风云三号 FY3A/B/C VIRR 传感器的相对辐射定标,提出基于时间信息的多任务卷积神经网络模型,用于估计传感器相对灵敏度系数,降低了传统相对定标中无变化像元筛选带来的不确定性,提高了多星长期观测数据之间的辐射一致性。Zhou 等^[15]以 HJ-2A/CCD3 为研究对象,采用 Sentinel-2/MSI 作为参考传感器,结合随机森林 BRDF 模型和光谱插值卷积方法开展交叉辐射定标,所得各波段总不确定度低于 5.16%,说明机器学习方法在跨传感器辐射传递和复杂观测几何修正中具有较好的适用性。

在红外载荷自身热辐射干扰校正方面,已有研究开始从温度场和仪器状态参数出发建立状态驱动的修正模型。Liang 等^[16]针对 FY-4B/GIIRS 扫描镜偏转引起的热辐射干扰,基于在轨深空观测数据分析了扫描镜角度变化对系统响应的影响,并建立了温度场驱动的扫描镜热辐射响应模型。经过扫描角差异补偿后,深空辐亮度偏差平均值降低至 $0.06 \text{ mW}/(\text{m}^2 \cdot \text{sr} \cdot \text{cm}^{-1})$,标准差降低至 $0.5 \text{ mW}/(\text{m}^2 \cdot \text{sr} \cdot \text{cm}^{-1})$ 以下。此外,针对 FY-4B/GIIRS 非标称光路杂散辐射与系统非线性响应耦合问题,相关研究基于发射前热真空定标实验建立修正模型,使 $200 \sim 320 \text{ K}$ 亮温内的定标偏差降低至 $0.3 \sim 0.4 \text{ K}$ ^[17]。这些结果表明,利用在轨温度状态、仪器响应参数和深空观测数据建立数据驱动或状态驱动模型,已成为提升红外载荷在轨辐射定标稳定性的重要方向。

总体而言,现有数据驱动辐射定标研究多围绕观测辐亮度偏差订正,传感器响应一致性补偿以及局部光机热辐射干扰修正展开,主要服务于系统偏差削弱、跨传感器辐射传递和特定部件响应偏移校正。相比之下,针对红外探测器在轨复杂时变温度场直接建立多元温度场与探测器自身背景辐亮度之间的映射关系,并用于仪器背景辐射实时估算的研究仍相对不足。

基于上述研究现状,本文提出一种基于数据驱动模型的探测器背景辐射实时估算方法。使用在轨获取的长时间序列温度场数据和对对应时刻的背景辐射观测数据构建模型训练与验证数据库,以卫星复杂耦合温度场数据作为输入,采用随机森林算法,突破了传统方法静态修正的局限,建立了探测器自身背景辐亮度估算模型,从而提高模型对在轨复杂时变热环境的表征能力,为实现背景噪声在轨实时估算与探测器精确定标提供了新路径。

2 模型数据库构建

2.1 卫星温度场选取

卫星探测器在轨温度场主要由外部热环境、内部热源以及仪器热传导特征与控温系统等多种因素共同决定。其中,外部热源主要包括太阳辐射、地球辐射(反射的太阳短波辐射与发射的长波辐射)以及宇宙杂散辐射,内部热源主要指的是光学元件发热、探测器自身发热及其他结构件发热等。仪器不同部件材料性质、几何布局、热源分布不同,最终导致温度在空间上存在明显的不均匀性,并且随卫星运行轨道周期而变化,形成了仪器温度场特征。在卫星设计时,为准确掌握其在轨状态,通常在关键元器件上装配多个测温点,如主镜、次镜、后光路等,并将测温信息实时传递到地面控制站,为实现背景辐射估算提供了长周期、多特征、准确的温度场数据支持。

由于不同卫星测温点的不同,在两类卫星类型中各选取了 8 个重要测温点作为卫星的代表性温度场,用于后续仪器背景估算模型输入,如表 1 所示。两类卫星虽然具有相同的光学口径,但其光学系统 F 数不同,对应的光机结构布局、热耦合关系以及背景辐射响应特性存在一定差异。因此,在相同轨道热环境条件下,两类卫星仪器

温度场的局部分布及其对背景辐射的影响规律并不完全一致。卫星特征温度场选择详情如表 1 所示。选取的温度测点基本全面覆盖卫星探测器重要元器件温度,可以较为准确地反映影响探测器在轨背景辐射变化的复杂温度场特征。研究中同时采用两类具有不同 F 数配置的卫星数据开展实验,进一步验证了本文所提背景辐射估算方法在不同光学系统条件下的适用性与普适性。

2.2 数据预处理与数据库构建

基于周期性深空观测辐射定标数据与下传的温度场数据进行时间匹配,分别建立卫星类型一与卫星类型二的短波红外与中波红外数据库,并对原始观测数据进行必要的预处理,据此构建满足后续模型训练与验证需求的标准化数据库。所构建的卫星类型一数据库观测时间跨度覆盖近一年周期,卫星类型二观测时间跨度覆盖近一

表 1 卫星特征温度场汇总

Tab.1 Summary of satellite characteristic temperature fields

编号	卫星类型一	卫星类型二
温度 1	主镜测温点 1	指向机构测温点 1
温度 2	主镜测温点 2	指向机构测温点 2
温度 3	主镜测温点 3	次镜测温点
温度 4	主镜测温点 4	指向镜测温点
温度 5	次镜测温点	主镜测温点
温度 6	后镜测温点	后光路测温点 1
温度 7	后光路测温点 1	后光路测温点 2
温度 8	后光路测温点 2	后光路测温点 3

季度周期。从原始数据的时间覆盖情况来看,两类观测数据均实现了完整的 24 小时覆盖,观测数据未出现明显的空缺时段或长时间连续缺测,体现了数据的完整性与连续性特征。数据库时序覆盖统计情况如图 1 所示。

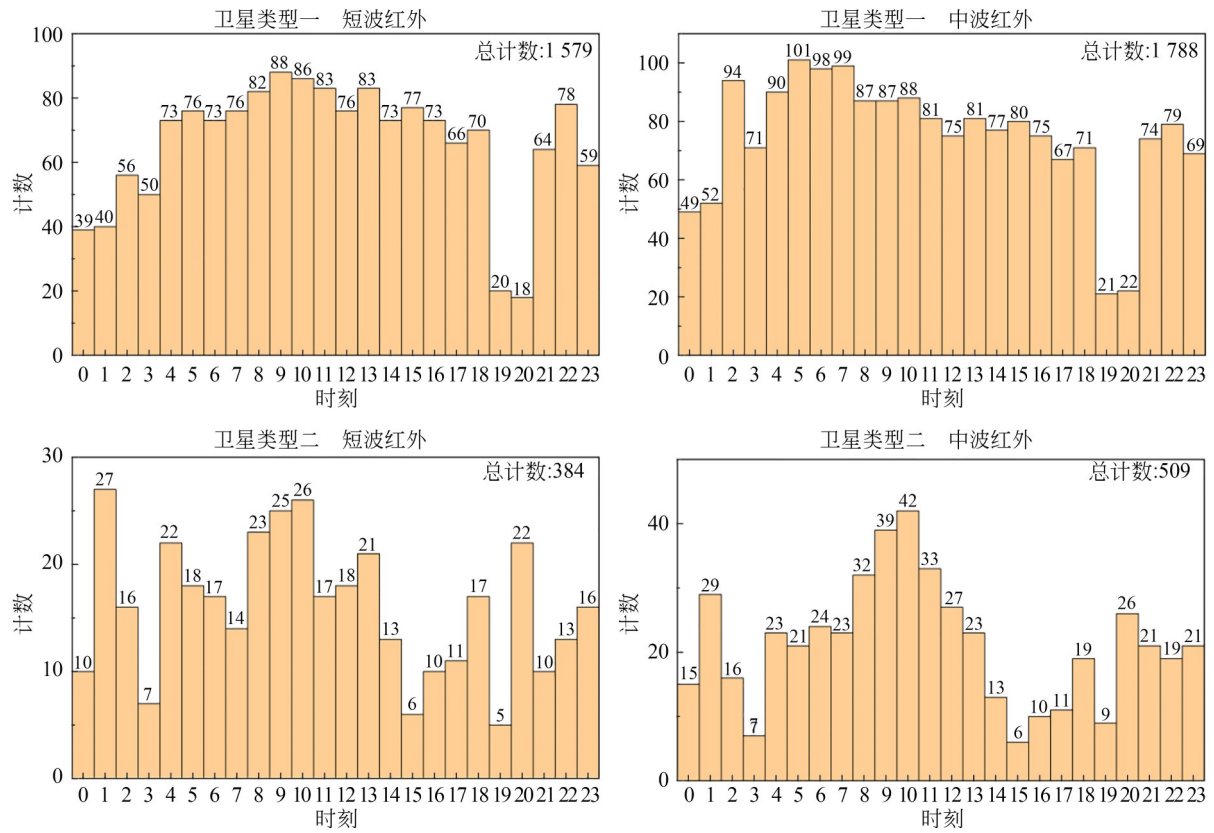


图 1 数据库时序覆盖统计图

Fig. 1 Statistical chart of database time series coverage

主要的预处理操作针对数据库中的背景辐射亮度进行。在各个数据库中,首先计算仪器背景

辐射亮度集合的均值和标准差,当某一数据点的观测值偏离集合均值超过 3 倍标准差时,即认为其

为异常点并予以剔除,也即 3σ 异常值处理。通过这一操作,可有效去除由短暂环境异常导致的极端数据点,大幅降低这些异常值对后续建模精度的不利影响。剔除异常点的总体比例低于总数据量的 2.5%,说明原始数据的整体质量较高,数据本身特性得到有效保留。

为进一步实现模型训练与验证的有效性与稳健性评估,对预处理完成后的数据集进行了随机分割处理,按 3:1 的比例划分为训练集和验证集,保证训练集与验证集之间的独立性。其中,训练集用于随机森林模型构建与超参数调整,而验证集则用于模型性能独立评估与泛化能力检验。

3 探测器背景辐射估算模型

3.1 温度场与背景辐射耦合机制

卫星探测器背景辐射是指由卫星各部分自身温度产生的热辐射信号。这些热辐射产生的信号叠加在目标观测信号上,形成辐射干扰或偏置,影响观测精度。依据热辐射基本原理,任何具有非零温度的物体都会向外发射热辐射,辐射强度与物体温度及材料发射率有关。单波段下,某器件表面单元产生的辐射可由普朗克公式描述:

$$B(\nu, T) = \frac{2h\nu^3}{c^2} \cdot \frac{1}{e^{\frac{h\nu}{kT}} - 1}, \quad (1)$$

其中: $B(\nu, T)$ 为辐射强度, ν 为频率, T 为温度, h 为普朗克常数, c 为光速, k 为玻尔兹曼常数。卫星仪器各器件依据其温度和材料发射率向探测器方向发射辐射信号,由之产生的整体背景辐射为各器件辐射累计值,即:

$$L_b = \sum_i \epsilon_i B(\nu, T) V_i, \quad (2)$$

其中: ϵ_i 为仪器第 i 个器件的表面发射率, $B(\nu, T)$ 为第 i 个器件在表面温度为 T 时产生的辐射强度, V_i 为第 i 个器件表面对探测器的视场因子。

卫星探测器温度场与自身背景辐射之间的关系具有显著的复杂性和非线性特征。温度场中任意区域的微小变化,都可能通过热辐射传输的方式对仪器整体的背景辐射产生显著影响。通过对普朗克公式进行一阶近似泰勒展开,可以

得到温度扰动对辐射亮度的影响:

$$\delta B(\nu, T) \approx \frac{2h^2\nu^4}{c^2 k T^2} \frac{e^{\frac{h\nu}{kT}}}{(e^{\frac{h\nu}{kT}} - 1)^2} \delta T. \quad (3)$$

这表明温度变化对辐射噪声具有高度非线性影响,当温度扰动较小时,辐射噪声的变化仍可能显著。此外,由于卫星仪器在实际轨道运行过程中受到复杂外部环境的动态影响,温度场在空间和时间维度上的分布高度不均匀且不断变化。温度场的空间非均匀性通过仪器内部的热传导、辐射传输以及结构件之间的相互作用,进一步强化了温度场和辐射噪声之间的复杂耦合关系。

温度场与辐射噪声之间的高度非线性和复杂耦合关系使得传统的线性回归方法难以有效地表征二者之间的内在联系。尽管基于有限元的热分析和辐射传输物理模型理论严谨,但实际应用中却面临仪器结构复杂、相关参量难获取与计算量过大等问题。由于卫星温度场与背景辐射之间存在显著的非线性、多变量耦合以及动态时变特征,传统基于解析表达式或线性假设的方法难以准确描述复杂热环境下背景辐射的变化规律。同时,卫星仪器内部不同部件之间存在热传导、热辐射以及热惯性等多重耦合作用,使得建立完整的物理机理模型具有较高难度。

因此,背景辐射估算模型不仅需要具备较强的非线性拟合能力,还需能够有效处理多维温度特征之间复杂的耦合关系,并保持良好的稳定性与泛化能力。基于上述特点,选择随机森林算法构建温度场与背景辐射之间的非线性映射模型。

3.2 随机森林模型介绍

基于前文对卫星温度场与背景辐射耦合机制的分析可以看出,两者之间存在明显的非线性、多变量耦合以及动态时变特征。针对传统解析模型难以准确描述复杂热环境下背景辐射变化规律的问题,采用随机森林算法构建背景辐射估算模型。随机森林是一种基于 Bagging 思想的集成学习回归算法,通过构建并集成多棵相互独立的回归树模型,实现对复杂非线性关系的有效拟合。它通过构建并集成多棵相互独立的回归树模型,以提升整体预测性能和模型稳健性。该方法由 Breiman 等^[18]在 2001 年提出,广泛应用于具有复杂变量关系或非线性特征的数据场

景^[19-22]。在本研究中,考虑到卫星仪器温度场的多特征性以及温度场与背景辐射的复杂非线性关系,旨在通过集成多回归树的预测,获得高精度、强泛化能力的探测器背景辐射估算模型。模型架构如图 2 所示。

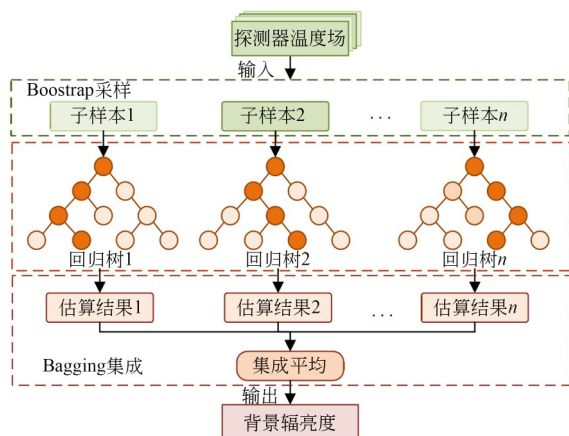


图 2 随机森林模型架构

Fig. 2 Architecture of random forest model

在训练过程中,基于 Bootstrap 采样生成训练子集,每棵回归树在独立的样本子集上进行建立。整体随机森林模型通过 Bagging 集成多棵树的估算结果,以均值作为最终输出,集成公式为:

$$y_e(\vec{x}) = \frac{1}{N} \sum_{i=1}^N T^i(\vec{x}), \quad (4)$$

其中: $y_e(\vec{x})$ 为整体模型估算结果, N 为回归树总数, $T^i(\vec{x})$ 为第 T 棵树的估算结果, $\vec{x}$ 为输入特征。本研究中随机森林参数调优是基于完整训练数据集的 MSE 进行的,用以获得稳定性最好、拟合效果最佳的结果。MSE 公式为:

$$MSE_{RF} = \frac{1}{M} \sum_{i=1}^M (y_r - y_e)^2, \quad (5)$$

其中: M 为训练数据集总样本量, y_r 为真实值, y_e 为估算值。在进行模型参数调优时,选择训练误差最小的模型作为最终方案,以确保在已知数据范围内获得最优拟合能力。

总体而言,随机森林模型通过非线性估算关系的建立以及对多回归树的集成平均具有稳定的预测能力、合理的误差控制机制及良好的可解释性,为基于复杂温度场的探测器背景辐射估算提供了可靠的建模基础。

需要指出的是,随机森林模型的计算过程可

分为离线训练和在线推理两个阶段。训练阶段需要根据历史温度场数据和背景辐射观测数据构建多棵回归树,计算量相对较大,但该过程可在地面离线完成,不参与实际在轨实时估算流程。模型训练完成后,在线推理阶段仅需将实时获取的温度遥测数据输入已建立的随机森林模型,通过各回归树的节点判断获得预测结果,并对多棵树的输出进行平均。该过程不涉及迭代求解、复杂物理仿真或大规模矩阵运算,计算复杂度较低。本文模型输入特征数量为 8,随机森林回归树数量为 200~1 200 棵,因此单次背景辐射估算所需计算量较小,具备满足基于在轨温度遥测数据开展快速背景辐射估算的可行性。

3.3 对比模型与实验设置

为进一步量化本文所提出随机森林模型相对于传统背景辐射校正方法和简单机器学习算法的性能差异,本文设置了若干对比模型。对比模型的选取主要考虑两方面:一方面,需要回应传统热平衡模型、地面实验室定标和经验系数校正等方法在探测器背景辐射估算中的适用性问题;另一方面,需要比较随机森林算法与其他常用回归算法在多元温度场背景辐射估算任务中的性能差异。

对于热平衡模型而言,其通常需要准确获得仪器结构参数、材料热物性参数、部件表面发射率、视场因子、热传导边界条件以及部件间热耦合关系,并需要结合轨道外热流、内部热源和热控状态建立完整的热分析模型。该方法在理论上具有较强的物理可解释性,但在实际在轨应用中,相关结构和热物性参数难以完全获取,且真实在轨边界条件随轨道周期和载荷工作状态动态变化,导致完整热平衡模型建模复杂、计算量较大,难以直接用于本文所关注的背景辐射实时估算任务。因此,本文不将完整热平衡模型作为直接对比模型,而是在后续分析中通过多温度输入的线性经验模型表征传统温度补偿方法的可实现估算能力。

地面实验室定标通常在受控温度、真空和稳定光学环境下获得探测器暗电流或背景响应,并将其作为后续辐射校正中的固定偏置项或初始定标参数。该方法能够为探测器提供必要的初始定标基准,但其本质上对应的是特定温度和特

定工作状态下的固定修正量。当天基红外探测器进入在轨运行后,太阳辐照、地球红外辐射、轨道周期变化以及星上设备功耗波动会共同引起载荷温度场的动态变化,使仪器背景辐射不再表现为固定值,而是随多部件耦合温度场发生时变化。因此,仅依赖地面实验室定标获得的固定暗电流或固定背景响应,难以适应复杂非稳态在轨热环境下的背景辐射实时修正需求。

对于经验系数校正方法,本文采用多元线性回归模型作为其等效基线模型。该模型以全部温度测点作为输入,将探测器背景辐亮度表示为多个温度变量的线性组合:

$$\hat{L} = \beta_0 + \sum_{i=1}^m \beta_i T_i, \quad (6)$$

其中: $\hat{L}$ 为背景辐亮度估算值, T_i 为第*i*个测温点温度, m 为温度测点数量, β_0 和 β_i 分别为截距项和回归系数。该模型能够反映多测温点经验系数校正方法的基本估算能力,但其本质上仍假设各温度变量对背景辐射的影响满足线性叠加关系,难以充分描述多部件在热耦合、热惯性及红外热辐射过程中的非线性响应。

除多元线性回归模型外,进一步选取K近邻回归模型和单棵决策树回归模型作为简单机器学习对比模型。K近邻回归是一种非参数局部回归方法,其预测结果主要由特征空间中相邻样本的背景辐亮度均值决定,能够在一定程度上刻画局部非线性关系,但对样本分布密度和特征尺度较为敏感。单棵决策树回归则通过递归划分特征空间建立分段预测模型,具有较好的非线性拟合能力和一定的可解释性,但容易受到训练样本扰动影响,存在过拟合和泛化稳定性不足的问题。

与上述模型相比,随机森林模型通过Bootstrap抽样构建多棵回归树,并对多棵树的输出结果进行集成平均,能够在保持决策树非线性拟合能力的同时降低单棵树模型的方差,提高模型对复杂温度场输入的稳定性及泛化能力。因此,本文选取多元线性回归、K近邻回归、单棵决策树和随机森林模型进行统一比较。这些模型分别对应线性经验修正、局部非参数回归、单树非线性建模和集成树模型等不同建模思路,可用于比较不同复杂度模型对探测器背景辐射估算任务的适用性。

4 模型训练与验证

4.1 模型训练与对比实验设置

考虑到不同类型卫星载荷结构、温度场分布及探测波段之间存在差异,分别针对卫星类型一短波红外、卫星类型一中波红外、卫星类型二短波红外和卫星类型二中波红外4类工况独立开展模型训练与验证。所有模型均采用第2.2节中相同的数据预处理结果和训练集/验证集划分方式,其中训练集用于模型参数拟合与超参数选择,验证集仅用于模型泛化性能评价。各模型输入均为对应卫星载荷的全部温度测点数据,输出均为对应时刻的探测器背景辐亮度,从而保证不同模型之间的对比具有一致性和可比性。

在随机森林模型训练过程中,所有模型均以完整的温度特征作为输入,即卫星载荷各关键部件所观测到的全部温度测量数据,以尽可能保留影响仪器背景辐射变化的多元温度场信息。为了优化随机森林模型的关键超参数,采用逐步循环寻优方式,对决策树数量(numTree)、叶节点最少样本数(minLeaf)以及决策树最大分裂次数(maxSplit)进行组合搜索。具体而言,在预先设定的参数范围内,以固定步长遍历不同超参数组合,并以训练集均方误差作为优化目标,选取综合拟合效果较优的参数组合作为最终模型参数。模型超参数调优范围、步长以及最优结果如表2所示。

依据上述调优流程,表2汇总了随机森林模型3个关键超参数的搜索范围、步长及其在4类工况下的最终取值。其中,numTree表示随机森林中回归树的数量,minLeaf表示叶节点最少样本数,maxSplit表示单棵树允许的最大分裂次数;各列最优参数均由训练集MSE最小原则确定,用于后续模型训练与验证。

为量化本文方法相对于传统经验修正方法和简单机器学习算法的性能差异,进一步设置多元线性回归、K近邻回归和单棵决策树回归作为对比模型。其中,多元线性回归用于表征经验系数校正方法在多温度输入条件下的等效估算能力;K近邻回归用于表征基于样本局部相似性的非参数回归能力;单棵决策树回归用于表征单树非线性划分模型的估算能力。上述对比模型与

随机森林模型采用相同的训练集、验证集、输入变量和输出变量,以避免由于数据划分或输入特征差异引入额外误差。

本研究采用三种评价指标对模型估算精度进行评价,分别为均方根误差(RMSE)、决定系数(R^2)以及相对误差(RBias)。公式为:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}, \quad (7)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \quad (8)$$

$$RBias = \frac{\sum_{i=1}^n (\hat{y}_i - y_i)}{\sum_{i=1}^n y_i} \times 100\%, \quad (9)$$

其中: n 为样本数量, y_i 为第*i*个样本的实际观测值, $\hat{y}_i$ 为第*i*个样本的模型估算值, $\bar{y}$ 为实际观测值的平均值。RMSE用于表征模型估算值与实际观测值之间的绝对误差水平, R^2 用于评价模型对观测值变化的解释能力,RBias用于衡量模型估算结果相对于实际观测值的整体偏差。

为进一步量化随机森林模型相对于对比模型的精度提升幅度,本文在对比分析中引入RMSE降低率作为补充量化指标:

$$\eta_{RMSE} = \frac{RMSE_{baseline} - RMSE_{RF}}{RMSE_{baseline}} \times 100\%, \quad (10)$$

其中: $RMSE_{baseline}$ 为对比模型在验证集上的均方根误差, $RMSE_{RF}$ 为随机森林模型在相同验证集上的均方根误差, η_{RMSE} 表示随机森林模型相对于对比模型的RMSE降低比例。

表 2 随机森林模型超参数调优详情

Tab. 2 Details of hyperparameter tuning for random forest model

超参数	调优范围	步长	卫星类型一 短波红外	卫星类型一 中波红外	卫星类型二 短波红外	卫星类型二 中波红外
numTree	200~2 000	200	200	1 200	200	800
minLeaf	1~10	1	1	3	1	3
maxSplit	200~2 000	100	1 200	1 800	1 600	1 200

4.2 随机森林模型训练与验证结果

在模型训练过程中,总计使用了3 126条有效数据。训练结果如图3所示,绝大多数数据点集中在1:1线上,训练结果表明随机森林模型无论在卫星类型一、卫星类型二还是短波红外、中波红外均能取得较高的训练精度。具体而言,在卫星类型一短波红外场景下模型RBias为0.836 7%,RMSE为0.000 1 W/(sr·m²), R^2 为0.989 7;在卫星类型一中波红外场景下模型RBias为0.076 7%,RMSE为0.001 3 W/(sr·m²), R^2 为0.987 4;在卫星类型二短波红外场景下模型RBias为5.263 7%,RMSE为0.000 3 W/(sr·m²), R^2 为0.981;在卫星类型二中波红外场景下模型RBias为0.207 5%,RMSE为0.005 1 W/(sr·m²), R^2 为0.994 1。

模型验证结果图4所示,共使用了1 045条数据,所使用验证数据集与训练集数据间相互独立。其中,在卫星类型一短波红外场景下模型

RBias为2.934 5%,RMSE为0.000 2 W/(sr·m²), R^2 为0.893 6;在卫星类型一中波红外场景下模型RBias为0.333 1%,RMSE为0.002 3 W/(sr·m²), R^2 为0.954 6;在卫星类型二短波红外场景下模型RBias为5.006 2%,RMSE为0.000 1 W/(sr·m²), R^2 为0.824 5;在卫星类型二中波红外场景下模型RBias为0.069%,RMSE为0.009 6 W/(sr·m²), R^2 为0.985 4。与训练结果相比,模型验证过程的准确性或性能有所下降,不过总体精度仍然较高,能够满足应用需求,表明模型在不同条件下均具有稳健的泛化能力与稳定性。

总体而言,尽管不同卫星平台的温度场特性及不同红外谱段的辐射响应存在差异,随机森林模型在各类工况下仍表现出较好的预测稳定性和估算精度,说明其能够较有效地表征多元温度场与探测器背景辐射之间的非线性映射关系,并适用于不同谱段探测器背景辐亮度的定量估算。

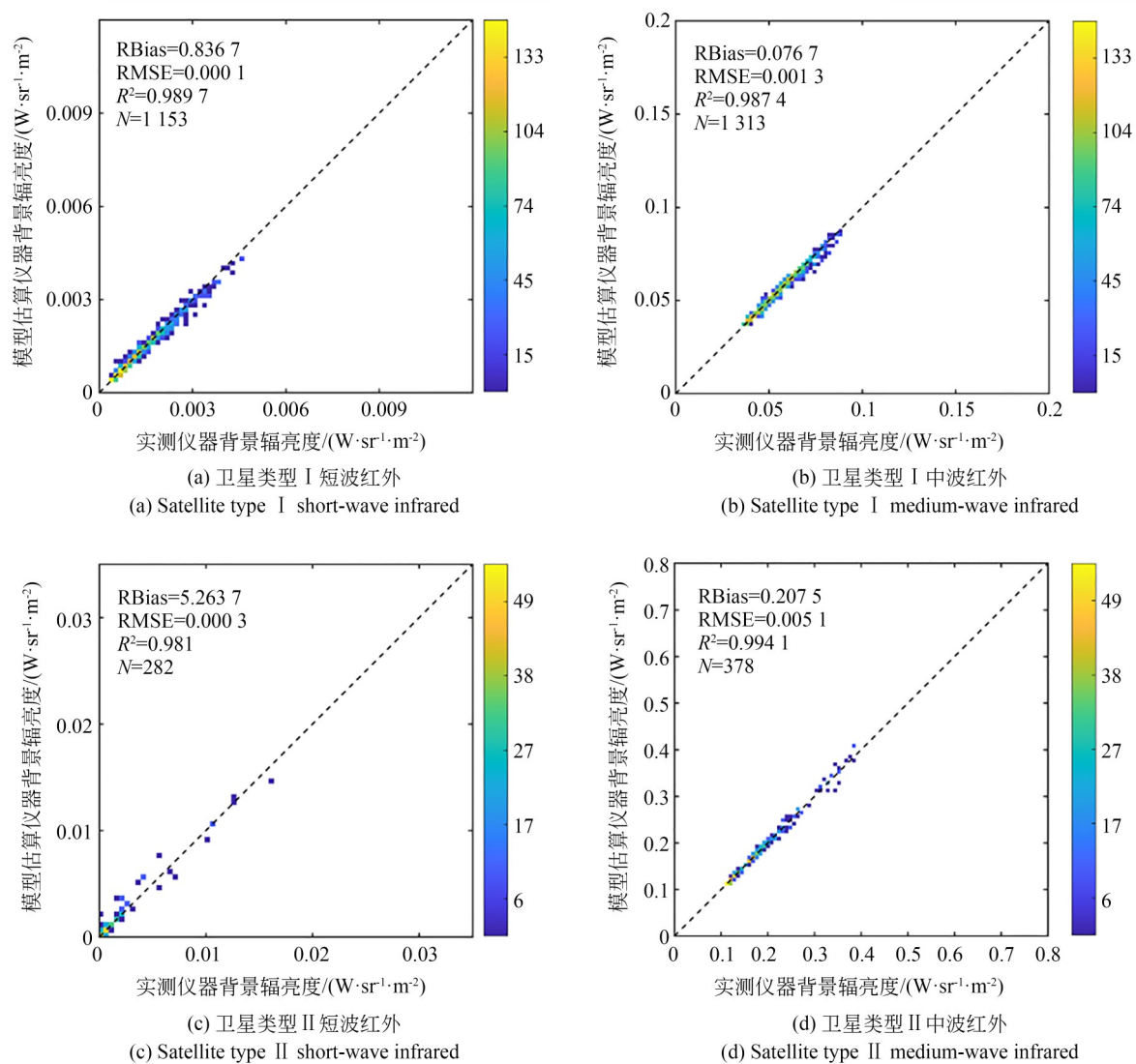
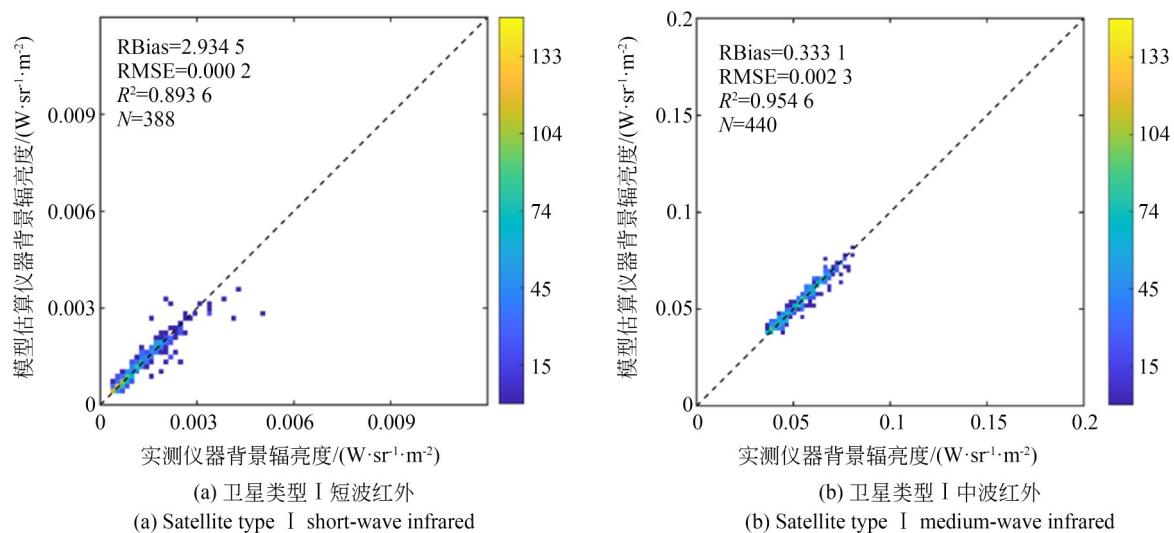


图 3 模型训练结果点密度

Fig. 3 Point density of model training results



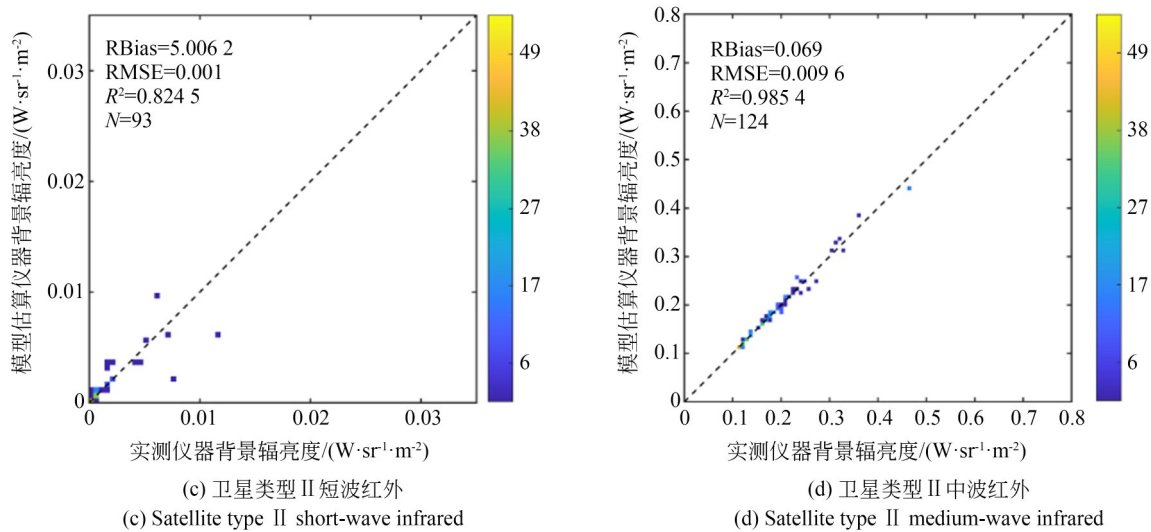


图 4 模型验证结果点密度图
Fig. 4 Point density map of model validation results

然而,仅给出随机森林模型自身的训练和验证结果,仍不足以量化其相对于传统经验修正方法的精度提升,也难以说明其相对于其他常用回归算法的适用优势。因此,本文进一步选取多元线性回归、KNN 回归和单棵决策树回归作为对比模型,在相同训练集、验证集和评价指标下开展对比分析。

4.3 不同估算模型对比分析

为进一步评价随机森林模型相对于传统经验修正方法及其他常用回归模型的性能优势,在相同训练集、验证集、输入变量和输出变量条件下,对多元线性回归、KNN 回归、单棵决策树和随机森林模型进行了统一验证。表 3 汇总了 4 种估算模型在验证集上的平均精度对比结果,用于评价不同建模思路在探测器背景辐亮度估算任务中的泛化能力。表中平均 R^2 、平均 RMSE 和平均 RBias 均由卫星类型一短波红外、卫星类型一

中波红外、卫星类型二短波红外和卫星类型二中波红外 4 类工况的独立验证结果计算后取平均得到。

由表 3 可以看出,在相同训练集、验证集和输入特征条件下,随机森林模型在 4 类对比模型中表现出最优的综合估算性能。其平均 R^2 为 0.914 5,高于多元线性回归、KNN 回归和单棵决策树模型;平均 RMSE 为 0.003 1 $W/(sr \cdot m^2)$,明显低于多元线性回归的 0.011 3 $W/(sr \cdot m^2)$ 、KNN 回归的 0.009 6 $W/(sr \cdot m^2)$ 和单棵决策树的 0.008 3 $W/(sr \cdot m^2)$;平均 RBias 仅为 2.085 7%,也显著低于其余对比模型。上述结果说明,随机森林模型不仅具有较好的拟合精度,而且在整体偏差控制方面也表现出更强的稳定性。

为进一步量化随机森林模型相对于传统经验修正等效基线的精度提升,本文计算了其相对于多元线性回归模型的 RMSE 降低率。结果表

表 3 不同估算模型验证精度对比

Tab. 3 Comparison of validation accuracy among different estimation models

模 型	方法类型	平均 R^2	平均 RMSE/ $(W \cdot sr^{-1} \cdot m^{-2})$	平均 RBias/%
多元线性回归	线性经验修正	0.653 6	0.011 3	56.464 6
KNN 回归	局部非参数回归	0.792 8	0.009 6	8.550 9
单棵决策树	单树非线性模型	0.674 1	0.008 3	15.967 6
随机森林	集成树模型	0.914 5	0.003 1	2.085 7

明,随机森林模型的平均RMSE较多元线性回归降低约72.57%,说明传统线性经验修正方法难以充分描述多元温度场与背景辐射之间的复杂非线性关系。与此同时,随机森林模型相对于KNN回归和单棵决策树模型的平均RMSE也分别降低约67.71%和62.65%,表明其并非仅相对于线性模型具有优势,而是在不同类型回归方法中均表现出更好的综合估算能力。

从模型机理上看,多元线性回归假设各温度变量对背景辐射的影响满足线性叠加关系,难以描述多部件热耦合、热惯性以及红外热辐射过程中的非线性响应;KNN回归主要依赖样本局部相似性,容易受到样本分布密度和特征尺度差异的影响;单棵决策树虽然能够刻画一定的非线性关系,但模型方差较大,泛化稳定性不足。相比之下,随机森林通过Bootstrap抽样构建多棵回归树,并对各树输出结果进行集成平均,在保持非线性拟合能力的同时降低了单树模型的随机波动和过拟合风险。因此,随机森林更适用于复杂多元温度场驱动下的红外探测器背景辐射实时估算任务。

4.4 特征重要性分析

为了进一步反映温度场特征在仪器背景辐射实际预测过程中的相对影响力,为模型的物

理解释与后续优化提供依据,对模型的输入特征开展了特征重要性分析。

采用随机森林模型中袋外预测误差提升量作为特征重要性度量指标。该指标基于变量扰动对袋外样本预测误差的影响量进行计算。其具体原理为,对于每棵随机森林中的决策树,首先计算其在袋外样本上的预测误差(MSE),随后将某一输入特征的取值打乱,重新计算袋外样本误差,两者误差之差即为该特征对预测性能的影响,越大表明该特征越关键,公式为:

$$Importance(x_j) = MSE_{OOB,j} - MSE_{OOB}, \quad (11)$$

其中: x_j 为第 j 个特征, $MSE_{OOB,j}$ 为第 j 个样本打乱后模型预测误差, MSE_{OOB} 为无扰动模型误差。在完成针对所有特征的重要性计算后,对上述指标进行归一化处理,以统一尺度并提升不同模型间的可比性。归一化方式采用单一特征的重要性占所有特征重要性总和的比例:

$$I_j = \frac{Importance(x_j)}{\sum_{j=1}^N Importance(x_j)}. \quad (12)$$

特征重要性分析结果如图5所示。结果表明,在不同卫星、不同红外谱段下,靠近红外光学路径或结构中热惯性较弱的部件对应的温度特征具有较高的重要性,例如,主镜与后光路温度

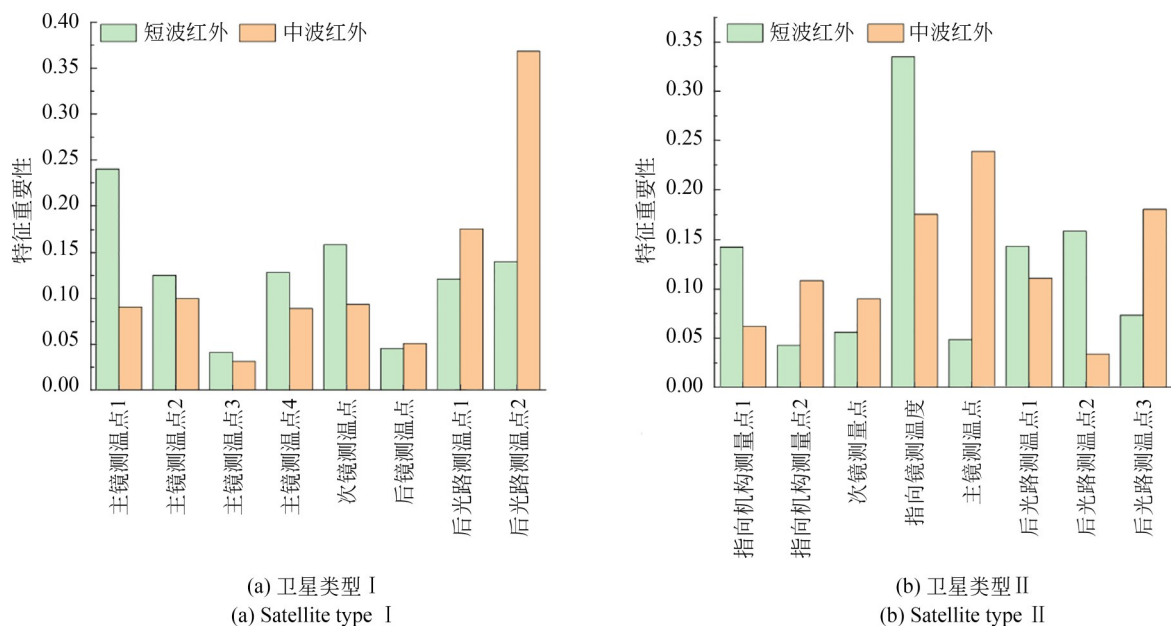


图5 不同卫星特征重要性分析结果

Fig. 5 Analysis result of feature importance for different satellites

都具有较高的特征重要性。这与仪器背景的物理来源高度一致,进一步验证了本研究所建立的数据驱动模型内部机制对实际物理过程的良好表征能力,模型对卫星温度场与仪器背景辐亮度间复杂耦合物理关系的有效把握也是其实现基于卫星温度场的探测器背景辐亮度高精度估算的主要原因。

5 应用与分析

5.1 多场景应用分析

在红外遥感观测中,探测器接收到的总辐射信号可以简单分为两类,一是接收到的地物目标辐射,二是仪器自身产生的背景辐射。仪器自身背景辐亮度作为红外探测系统的一个关键噪声源,其估算误差可以直接对有效目标辐射产生干扰。为准确量化这种干扰程度,定义了估算相对误差指标用于刻画探测器背景辐亮度估算误差在不同地物观测情境中的影响程度。估算相对误差公式为:

$$E_R = \frac{|\widehat{Radiance}_1 - Radiance_1|}{Radiance_B} \times 100\%, (13)$$

其中: $RError$ 为探测器背景辐亮度估算相对误差, $\widehat{Radiance}_1$ 代表模型估算的背景辐亮度, $Radiance_1$ 代表实测背景辐亮度, $Radiance_B$ 代表观测地物辐亮度。该指标描述了仪器背景辐亮度估算误差相对于有效地物信号的扰动幅度,可以有效反映仪器背景辐亮度估算误差在不同地物背景下可能造成的信噪干扰占比。

为系统评估该误差指标在真实观测任务中的表现,基于实测天基红外数据,选取了共计 101 组样本,涵盖了包括山地、平原、沙漠、海洋、湖泊、云层在内的 6 种典型地物类型。每组样本间相互独立,覆盖观测任务中常见且辐射特征差异明显的典型对象,保证了评估数据的真实性与广泛代表性。模型应用效果评估箱线图如图 6 所示。

从箱线图整体来看,随机森林模型在不同地物类型下均表现出较好的估算精度,仪器背景辐亮度估算误差对观测地物辐亮度的影响总体控制在较低水平。各类地物的平均相对误差分别为:海洋场景 4.44%,湖泊场景 6.80%,平原场景 4.58%,沙漠场景 3.13%,山地场景 1.81%,云层

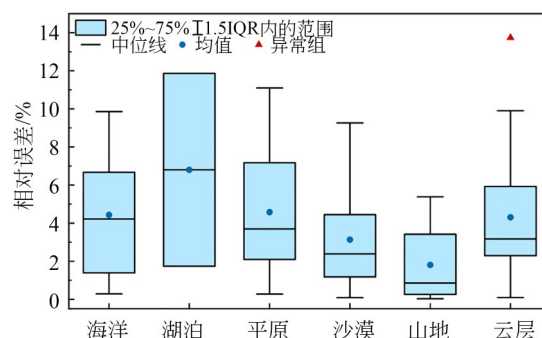


图 6 模型应用相对误差与地物成像估算精度箱线图

Fig. 6 Boxplot of relative error in model application and estimated accuracy of ground object imaging

场景 4.31%。可以看出,不同地物类型之间的误差水平并不完全按照地物辐亮度高低呈单调变化。例如,海洋与平原场景的平均相对误差较为接近,说明模型应用误差不仅受目标地物辐亮度大小影响,还与对应观测样本中的背景辐射估算残差、地物辐射空间均匀性、观测几何条件以及样本分布特征等因素有关。

其中,湖泊场景的平均相对误差相对较高,可能与其目标辐亮度较低、局部观测条件差异以及样本数量有限等因素共同有关;而山地和沙漠场景的平均相对误差相对较低,表明在部分辐射特征较稳定或目标辐亮度相对较高的场景下,背景辐射估算误差对地物辐射反演结果的相对影响更小。总体而言,模型在 6 类典型地物场景下的平均相对误差为 3.88%,不同类别之间未出现明显失效情形,说明本文建立的数据驱动模型在多种实际观测场景中具有较好的稳定性和应用适应性。

5.2 时序应用分析

将卫星类型一某星 11 月与卫星类型二某星 3 月平均卫星日温度场特征输入到建立的数据驱动模型当中,得到仪器背景辐亮度在 24 小时周期内随时间变化的连续估算序列。结果如图 7 所示,全面展现了仪器背景辐亮度日周期性变化规律,与模型在不同卫星类型估算效果与真值比对。

分析可知,模型估算结果与实测结果十分接近,尤其在短波红外探测器中偏差极小,中波红外探测器中虽然存在一定振荡但仍然吻合实际值的时序变化趋势。中波红外探测器背景辐射估算出现震荡可能是由于中波探测器受到热辐

射影响更明显,温度场的轻微震荡引起的背景辐射变化更大。总体而言,模型在不同卫星类型,

不同季节除太阳规避时间外的持续仪器背景辐射亮度估算任务当中具备良好的稳定性。

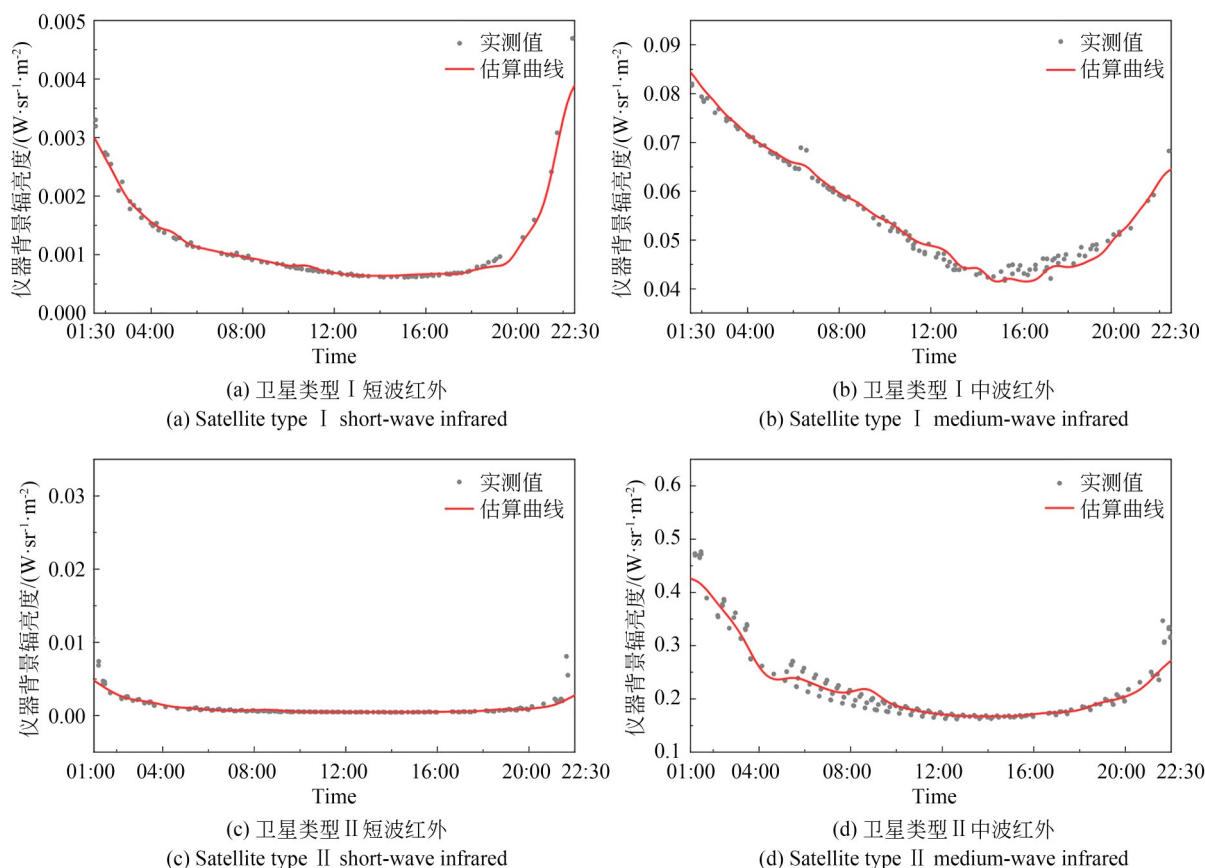


图 7 仪器背景辐射亮度时序估算效果曲线

Fig. 7 Curve of time-series estimation effect of instrument background radiance

6 结 论

卫星在轨的复杂热环境及其与背景辐射的非线性耦合关系对探测器背景辐射的高精度实时估算提出了极大的挑战。本研究从红外热辐射基本原理出发,提出了一种基于数据驱动模型的红外探测器自身辐射实时定量估算方法。研究首先选取了 8 个重要测温点用于后续仪器背景估算模型输入,对数据进行预处理后构建模型数据库。随后利用随机森林算法,建立了基于多温度特征输入的仪器背景辐射估算模型。模型精度评价结果表明,模型训练的 R^2 为 0.981 0~0.994 1,验证的 R^2 为 0.824 5~0.985 4。利用涵盖了包括山地、平原、沙漠、海洋、湖泊、云层在内的 6 种典型地物类型的实测天基红

外数据进一步分析模型应用效果,结果表明模型估算精度的多场景平均相对误差仅为 3.88%。总体而言,本研究所提出的数据驱动方法在不同条件下均具备良好的模型泛化能力与稳定性,能够有效应对卫星载荷温度场复杂多变的动态特征,为稳定地实现仪器背景辐射的高精度实时定量估算,提升红外探测器辐射定标水平提供了技术途径。

作者贡献声明:

梅 坚:实时估算方法提出,论文构思和撰写;

杨 林:论文审核与编辑写作;

张荣国:数据库搭建;

刘宇阳:实验数据分析;

周晓萱:模型建立;

延檬羽:实验结果验证;

陈巧妹:实验指导;

郭梦霏:估算方法策划与执行。

参考文献:

- [1] 陈凡胜, 胡琰悦, 李潇雁, 等. 宽幅高分辨热红外遥感成像技术研究[J]. 中国激光, 2021, 48(12): 1210002.
CHEN F SH, HU ZH Y, LI X Y, *et al.* Research on wide-range and high-resolution thermal infrared remote sensing imaging technology[J]. *Chinese Journal of Lasers*, 2021, 48(12): 1210002. (in Chinese)
- [2] 王浩, 晋利兵, 李岩, 等. 基于温度修正的 GF-5 全谱段光谱成像仪热红外谱段辐射定标[J]. 红外技术, 2020, 42(6): 547.
WANG H, JIN L B, LI Y, *et al.* Radiation calibration of thermal infrared bands of GF-5 VIMI based on temperature correction[J]. *Infrared Technology*, 2020, 42(6): 547. (in Chinese)
- [3] MASON J P, LAMPRECHT B, WOODS T N, *et al.* CubeSat on-orbit temperature comparison to thermal-balance-tuned-model predictions [J]. *Journal of Thermophysics and Heat Transfer*, 2018, 32(1): 237-255.
- [4] BOUSHON K E. *Thermal Analysis and Control of Small Satellites in Low Earth Orbit*[D]. Rolla: Missouri University of Science and Technology, 2018.
- [5] SMITH D, PETERS D, NIGHTINGALE T, *et al.* Challenges for in-flight calibration of thermal infrared instruments for earth observation[J]. *Remote Sensing*, 2020, 12(11): 1832.
- [6] PETERS D M, SMITH D, NIGHTINGALE T, *et al.* The next generation thermal infrared satellite in-flight radiometric calibration sources[J]. *Temperature: Its Measurement and Control in Science and Industry*, VOLUME 9: Proceedings of the Tenth International Temperature Symposium, 2024, 3230: 100005.
- [7] 刘栋, 戴聪明, 唐超礼, 等. 基于非局域热平衡模式的临近空间目标背景临边对比度模型[J]. 红外与毫米波学报, 2017, 36(5): 620.
LIU D, DAI C M, TANG CH L, *et al.* Model of near-space limb contrast based on non-local thermodynamics equilibrium model[J]. *Journal of Infrared and Millimeter Waves*, 2017, 36(5): 620. (in Chinese)
- [8] 金占雷, 晋利兵, 张九双, 等. 基于幂加权最小二乘拟合的红外相机辐射定标方法[J]. 光子学报, 2024, 53(4): 0428002.
JIN ZH L, JIN L B, ZHANG J SH, *et al.* Radiometric calibration method for infrared camera based on power-weighted least squares fitting[J]. *Acta Photonica Sinica*, 2024, 53(4): 0428002. (in Chinese)
- [9] 郭强, 陈博洋, 张勇, 等. 风云二号卫星在轨辐射定标技术进展[J]. 气象科技进展, 2013, 3(6): 6-12.
GUO Q, CHEN B Y, ZHANG Y, *et al.* Progress in on-orbit radiometric calibration technology for Fengyun-2 satellite[J]. *Advances in Meteorological Science and Technology*, 2013, 3(6): 6-12. (in Chinese)
- [10] ZHOU G Q, LI C Y, YUE T, *et al.* An overview of in-orbit radiometric calibration of typical satellite sensors[J]. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2015, XL-7/W4: 235-240.
- [11] GOTO A, UMEDA K, YUKUMATSU K, *et al.* Property changes in materials due to atomic oxygen in the low Earth orbit[J]. *CEAS Space Journal*, 2021, 13(3): 415-432.
- [12] SZNAJDER M, GEPPERT U, DUDEK M. Degradation of metallic surfaces under space conditions, with particular emphasis on Hydrogen recombination processes [J]. *Advances in Space Research*, 2015, 56(1): 71-84.
- [13] ZHANG X W, XU D M, LI X, *et al.* Nonlinear bias correction of the FY-4A AGRI infrared radiance data based on the random forest[J]. *Remote Sensing*, 2023, 15(7): 1809.
- [14] TANG L, ZHAO X G, HU X Q, *et al.* A multi-task convolutional neural network relative radiometric calibration based on temporal information[J]. *Remote Sensing*, 2024, 16(17): 3346.
- [15] ZHOU X, CHEN Y D, XIE Y, *et al.* Radiometric cross-calibration of HJ-2A/CCD3 using the random forest algorithm and a spectral interpolation convolution method with sentinel-2/MSI[J]. *Remote Sensing*, 2024, 16(22): 4337.
- [16] LIANG X, ZOU Y P, HAN C P, *et al.* Temperature-field driven adaptive radiometric calibration for scan mirror thermal radiation interference in FY-4B

- GIIRS[J]. *Remote Sensing*, 2025, 17(24): 3948.
- [17] LIANG X, ZOU Y P, HAN C P, *et al.* Radiometric correction of stray radiation induced by non-nominal optical paths in Fengyun-4B geostationary interferometric infrared sounder based on pre-launch thermal vacuum calibration [J]. *Remote Sensing*, 2025, 17(16): 2828.
- [18] BREIMAN L. Random forests [J]. *Machine Learning*, 2001, 45(1): 5-32.
- [19] BELGIU M, DRĂGUT L. Random forest in remote sensing: a review of applications and future directions [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2016, 114: 24-31.
- [20] SPEISER J L, MILLER M E, TOOZE J, *et al.* A comparison of random forest variable selection methods for classification prediction modeling [J]. *Expert Systems with Applications*, 2019, 134: 93-101.
- [21] SCHONLAU M, ZOU R Y. The random forest algorithm for statistical learning [J]. *Stata Journal*, 2020, 20(1): 3-29.
- [22] GRINSZTAJN L, OYALLON E, VAREQUAUX G. Why do tree-based models still outperform deep learning on tabular data? [EB/OL]. 2022: *arXiv*: 2207.08815. <https://arxiv.org/abs/2207.08815>.

作者简介:



梅 坚(1999—),男,浙江宁波人,博士研究生,2021年于大连理工大学获得学士学位,主要从事红外物理方面的研究。E-mail: meijian@mail. sitp. ac. cn

通讯作者:



陈巧妹(1984—),女,福建宁德人,硕士,工程师,2006年、2009年于西北工业大学分别获得学士和硕士学位,主要从事红外遥感相机的可靠性、安全性设计及国产自主可控技术等方面的研究。E-mail: chommy_97@163. com

通讯作者:



郭梦霏(1999—),女,山东烟台人,硕士,助理研究员,2021年于河海大学获得学士学位,2024年于北京师范大学获得硕士学位,主要从事天基红外遥感、辐射定量反演、智能化数据分析等方面的研究。E-mail: Mavis6399@163. com