0254-6124/2026/46(3)-0628-11 *Chin. J. Space Sci.* 空间科学学报

CHEN Qingyan, JI Yuanfa, SUN Xiyan, LIANG Wenbin, GHAZALI Kamarul Hawari Bin. Prediction of the Ionospheric Irregularities Based on Residual Compensation WHO-RF Model (in Chinese). *Chinese Journal of Space Science*, 2026, **46**(3): 628–638. CSTR:32142.14.cjss.2025–0117. DOI:10.11728/cjss2026.03.2025–0117

基于残差补偿 WHO-RF 模型的电离层不规则体预测*

陈清燕[1,2,4] 纪元法[1,2,3] 孙希延[1,2,3,4]
梁文斌[1,2] GHAZALI Kamarul Hawari Bin[3]

1(桂林电子科技大学 广西精密导航技术与应用重点实验室 桂林 541004)

2(桂林电子科技大学信息与通信学院 桂林 541004)

3(教育部时空信息与智能位置服务国际合作联合实验室 桂林 541004)

4(南宁桂电电子科技研究院有限公司 南宁 530031)

摘 要 针对电离层不规则体预测困难和单一随机森林 (RF) 模型在预测中存在精度不高且参数调整困难等问题, 在结合野马优化 (Wild Horse Optimizer, WHO) 算法的基础上进行残差补偿 (Residual Compensation, RC), 构建 RC-WHO-RF 电离层不规则体预测模型. 利用 2020 年 3 月 1 日至 2024 年 6 月 30 日香港 HKWS 站的观测数据, 计算电离层总电子含量变化率的标准差 (ROTI) 以及选择一系列与不规则体相关的背景电离层参数作为输入特征进行实验分析. 结果表明, ROTI、余弦相位日变化因子和地磁活动指数对电离层不规则体的预测至关重要; RC-WHO-RF 模型预测的均方根误差均小于 0.1 TECU·min^{-1}, 对突发性磁暴事件也具有优异的响应能力和预测精度; RC-WHO-RF 模型在提前 30 min 的短临预报中平均相对精度达 90.67%, 比 WHO-RF 提升 8.16%, 比单一 RF 模型提升 11.2%, 组合模型的预测效果要明显优于单一 RF 模型和 WHO-RF 模型.

关键词 电离层不规则体, 随机森林模型, 野马优化算法, 残差补偿

中图分类号 P352

Prediction of the Ionospheric Irregularities Based on Residual Compensation WHO-RF Model

CHEN Qingyan[1,2,4] JI Yuanfa[1,2,3] SUN Xiyan[1,2,3,4]
LIANG Wenbin[1,2] GHAZALI Kamarul Hawari Bin[3]

1(*Guangxi Key Laboratory of Precision Navigation Technology and Application, Guilin University of Electronic Technology, Guilin* 541004)

* 广西科技计划项目 (AA23062038, ZY23055048, AB23026120, 桂科 24206043), 国家自然科学基金项目 (U23A20280, 62161007, 62471153), 南宁市科学研究与技术开发计划项目 (20231029, 20231011) 和广西研究生教育创新计划项目 (YCSW2024329) 共同资助

2025-07-14 收到原稿, 2025-10-16 收到修定稿

©The Author(s) 2026. This is an open access article under the CC-BY 4.0 License (https://creativecommons.org/licenses/by/4.0/)

2(*School of Information and Communication, Guilin University of Electronic Technology, Guilin* 541004)

3(*International Joint Research Laboratory of Spatio-temporal Information and Intelligent Location Services, Guilin* 541004)

4(*GUET-Nanning E-Tech Research Institute Co., Ltd., Nanning* 530031)

Abstract In response to the difficulties in predicting ionospheric irregularities and the low accuracy and tendency to fall into local optima of a single Random Forest (RF) model in prediction, a RC-WO-RF ionospheric irregularities regression prediction model was constructed by combining the Wild Horse Optimizer (WHO) algorithm with Residual Compensation (RC). Using observation data from the Hong Kong HKWS station from 1 March 2020 to 30 June 2024, the Rate of Total Electron Content Index (ROTI) was calculated, and a series of background ionospheric parameters related to ionospheric irregularities were selected as input features. The results indicate that, ROTI, Cosine phase daily variation factor and geomagnetic activity are crucial for ionospheric irregularities; The root mean square error of the RC-WO-RF model is less than 0.1 TECU·min^{-1}, and it has excellent response capability and prediction accuracy for sudden geomagnetic storm events; The average relative accuracy of the RC-WO-RF model in short-term forecasting 30 min in advance is 90.67%, which is 8.16% higher than the WHO-RF model and 11.2% higher than the single RF model. The prediction performance of the combined model is significantly better than that of the single RF model and the WHO-RF model.

Key words Ionospheric irregularities, Random forest model, Wild Horse Optimizer (WHO) algorithm, Residual compensation

0 引言

在日落后的几小时, 低纬度地区电离层常出现瑞利–泰勒不稳定状态, 导致等离子体密度不规则体的形成. 电离层不规则体可使 100～4000 MHz 频率范围内的跨电离层无线电信号幅度、相位发生快速起伏变化, 引发强烈的电离层闪烁现象, 严重时甚至会导致卫星通信链路中断[1]. 随着现代社会对 L 波段无线电信号的依赖日益增加, 如何准确、及时预测或预警电离层不规则体, 已成为保障通信导航系统稳定运行的关键所在.

近几十年来, 有关电离层不规则体的气候学特征已有广泛的研究. 电离层不规则体的出现不仅与地方时强相关, 并且还依赖于季节和太阳活动等变化, 在中国低纬地区, 电离层不规则体在日落后出现, 其出现频率在 21:00－22:00 LT 附近达到峰值, 并持续至午夜[2]; 太阳活动高年春秋分前后电离层不规则体发生概率远高于其他季节[3]. 然而, 上述研究结果虽有助于预测电离层不规则体的长期变化趋势, 却难以实现其短期发生的可靠预报. 为此, 有研究探索利用各种机器学习算法预测电离层不规则体. Pi 等[4] 首次提出的 TEC (Total Electron Content) 变化率指数 ROTI 计算方法, 现已成为表征电离层不规则体的重要参数. 当前研究主要采用两种建模方式, 分类任务建模仅能判断不规则体是否发生, 而回归任务建模则可进一步预测其强度[5], Tete 等[6] 对比评估了极端梯度提升 (XGBoost)、随机森林 (RF)、长短期记忆 (LSTM) 等 5 种模型, 结果显示 RF 模型对 ROTI 的预测效果最佳; Habyarimana 等[7] 基于人工神经网络 (ANN) 建立的单站电离层不规则体预测模型的均方根误差 (RMSE) 为 0.148 TECU·min^{-1}; Atabati 等[8] 通过 ANN 模型与遗传算法 (GA) 相结合, 实现了决定系数为 80% 的 ROTI 日预测, Zhao 等[9] 提出的区域电离层不规则体预测模型在提前 1 h 预测中 RMSE 为 0.110 TECU·min^{-1}. 上述研究显示了人工智能技术在预测电离层不规则体方面具有巨大的潜力, 但是仍存在明显的局限性, 例如模型未充分考虑电离层不规则体与其驱动因素之间的关系重要性, 单一机器学习模型在超参数优化方面存在固有缺陷, 电离层不规则体预测精度还有待提升, 特别是在极端空间天气事件的预警性能方面表现显著不足.

针对上述问题, 本文提出一种改进的随机森林回

归预测模型 (RC-WHO-RF), 该模型通过野马优化算法 (Wild Horse Optimizer, WHO) 进行参数优化, 并引入残差补偿 (RC) 机制以提高预测精度. 使用 ROTI 作为目标变量, 同时选取包括太阳、地磁活动指数、时间变化因子等在内的多维背景电离层参数作为输入特征, 设计多组对比实验, 通过特定指标系统评估 RC-WHO-RF 模型在不同条件下的预测性能.

1 数据与算法

1.1 数据预处理

选择香港黄石 HKWS 监测站点 (22.436123°N, 114.327456°E), 时间采样率为 30 s 的全球导航卫星系统 (GNSS) 接收机采集的观测数据来计算电离层 ROTI 指数[10], 采用以下计算公式:

$$\delta_{\mathrm{STEC}} = \frac{1}{40.3} \times \frac{f_1^2 f_2^2}{f_1^2 - f_2^2}(L_1\lambda_1 - L_2\lambda_2), \tag{1}$$

$$\delta_{\mathrm{ROT}} = \frac{\delta_{\mathrm{STEC},i} - \delta_{\mathrm{STEC},i-1}}{\Delta t}, \tag{2}$$

$$\delta_{\mathrm{ROTI}} = \sqrt{\frac{1}{N}\sum_{j=i-N}^{i}\left(\delta_{\mathrm{ROT},j} - \delta_{\overline{\mathrm{ROT}}}\right)^2}. \tag{3}$$

式中, δ_{STEC}为卫星和接收机视线上的斜向 TEC, L_1 和 L_2 为载波相位, λ_1和λ_2为对应载波频率的波长; δ_{ROT}为δ_{STEC}随时间的变化率, Δt 为相邻两个历元间的时间差; δ_{ROT}和δ_{ROTI}单位均为 TECU·min^{-1}, δ_{ROTI}按 5 min 间隔δ_{ROT}的标准差计算; 为消除多径效应和信号遮挡的影响, 设置卫星截止高度角为 30°. 图 1 给出 HKWS 站点 2024 年 3 月 31 日发生中等强度电离层不规则体示例.

考虑到观测数据中存在测量噪声, ROTI 指数表现出较大的波动性. 因此, 采用移动平滑窗口对原始 ROTI 序列进行平滑处理[11], 图 2 为分别使用 6, 12 和 18 个点宽度的移动窗口对 ROTI 指数进行数据平滑处理的前后对比, 经对比选择 12 点宽度的窗口, 可以得到一条更好描述不规则体日变化的曲线, 而 6 点宽度的窗口曲线仍呈现明显抖动, 平滑效果不足, 18 点宽度的窗口过分平滑了曲线和相关的峰值.

电离层不规则体的单站预测不仅需要用 ROTI 指数, 还需要各种对电离层不规则体变化产生影响的环境参数. 基于对不规则体形成机制的考量, 选取表 1

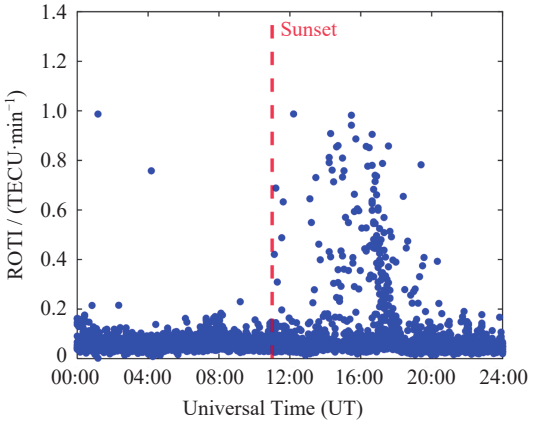

图 1 香港 HKWS 站点 2024 年 3 月 31 日日落后发生电离层不规则体时 ROTI 示例

Fig. 1 Example of ROTI when ionospheric irregularities occur at the HKWS station in Hong Kong after sunset on 31 March 2024

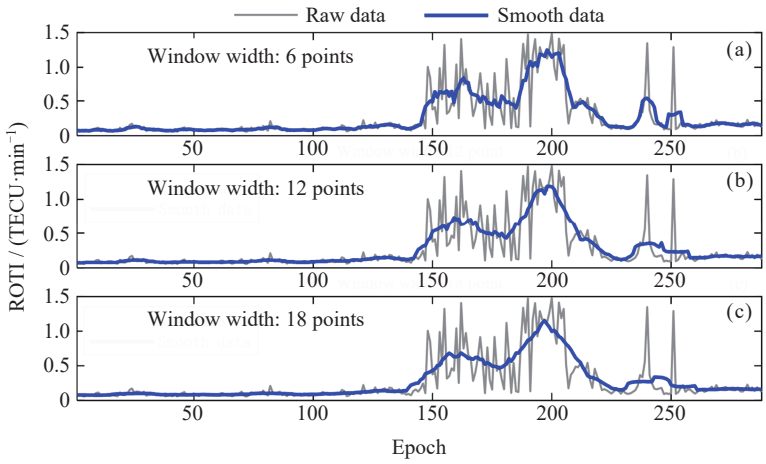

图 2 使用 6 点、12 点和 18 点移动窗口的平滑效果对比

Fig. 2 Comparison of smooth effects of moving windows using 6 points, 12 points, and 18 points

中所列多项电离层背景环境参数作为输入特征, 涵盖太阳辐射、地磁活动、赤道电急流变化以及 F_2 层等离子体条件等, 虽然各参数对不规则体形成的贡献度存在差异, 但是均会对电离层不规则体的发生、发展产生影响[7].

需要说明的是, 日变化因子 (HHC, T_{HHC}) 和 (HHS, T_{HHS}) 定义如下:

$$T_{\mathrm{HHC}}=\cos\left(\frac{2\pi H}{24}\right),T_{\mathrm{HHS}}=\sin\left(\frac{2\pi H}{24}\right). \tag{4}$$

式中, H 为一天中的协调世界时 (UTC).

由于香港地区测高仪设备与数据获取的限制, 此处使用的 F_2 层临界频率数据是由子午工程数据中心提供的广州天河站点的 f_0F_2. 广州天河站 (23.13°N, 113.35°E) 与香港 HKWS 站 (22.43°N, 114.32°E) 均位于中国低纬赤道异常北驼峰区域, 两地直线距离约 130 km, 在不规则体空间覆盖与漂移运动范围内, 具有高度空间相关性[12], 电离层 F 层峰值的日变化模式一致, 因此天河站 f_0F_2 数据能够用于表征香港区域的电离层背景状态.

1.2 模型构建

1.2.1 RF 模型

随机森林 (Random Forest, RF) 是由 Breiman 提出的一种集成决策树算法[13], 在决策树 (Decision Tree) 的基础上进行 Bagging 学习, 其主要思想是通过从原始训练集中随机抽取数据构造多个训练子集, 然后对每个训练子集采用节点的随机分裂技术构造多棵决策树, 最后将多棵决策树组合, 并采用平均法策略确定最终预测结果, 如图 3 所示.

表 1 电离层背景环境参数 (数据分辨率为 15 min)

Table 1 Ionospheric background environmental parameters (Data resolution is 15 min)

环境参数	数据描述
HHC	余弦相位日变化因子
HHS	正弦相位日变化因子
Dst	地磁指数
$F_{10.7}$	太阳射电流量指数
B_x	行星际磁场径向分量
B_y	行星际磁场东向分量
B_z	行星际磁场垂向分量
SYM/H	环电流的对称水平分量
Flow Pressure (FP)	磁流体压力
f_0F_2	F_2层临界频率

1.2.2 WHO 算法

野马优化算法 (WHO 算法是由 Naruei 和 Keynia 于 2021 年提出[14], 该算法主要通过模拟野马种群的生活行为 (觅食、追逐、支配、领导和交配) 进行寻优, 具有寻优能力强, 收敛速度快等特点.

(1) 种群初始化. 随机生成 N 个个体, 每个个体代表一组随机森林超参数, 种马占比为 P_s, 种群 $G=N\times P_s$, 将其余马驹平均分配在由种马领导的群体内, 之后根据式 (5) 的选拔机制并采用均方根误差 (RMSE) 作为适应度函数 $f(\cdot)$ 来重新选择种马个体, 有

$$N_{SG_i}=\begin{cases}X_{i,G}^{j}, & f\left(X_{i,G}^{j}\right)<f\left(N_{SG_i}\right);\\ N_{SG_i}, & f\left(X_{i,G}^{j}\right)\geqslant f\left(N_{SG_i}\right).\end{cases} \tag{5}$$

式中, N_{SG_i} 为第 i 个种群中种马的位置, $X_{i,G}^{j}$ 为第 i 个种群中成员 j 的位置, 如果其中一名成员的适应度值优于领导者, 这二者身份互换.

(2) 放牧行为. 种马位于放牧区中心带领着整个群体移动, 群成员围绕着中心进行 (放牧吃草) 搜索. 种群马驹位置更新公式为

$$\overline{X}_{i,G}^{j}=2Z\cos(2\pi RZ)\times\left(N_{SG_i}-X_{i,G}^{j}\right)+N_{SG_i}. \tag{6}$$

式中, R 为取值在 [–2, 2] 区间内的随机数, 主要用来控制个体与领导者放牧的角度. 自适应机制 Z 的计算公式为

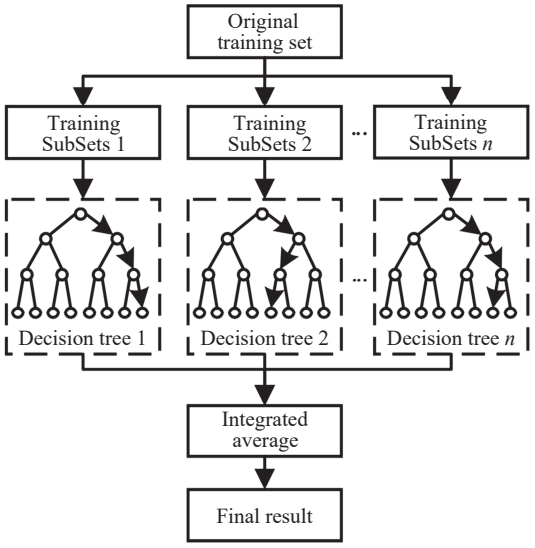

图 3 RF 模型结构

Fig. 3 Structure of RF model

$$Z = R_2 \odot P_{\mathrm{IDX}} + R_3 \odot (\sim P_{\mathrm{IDX}}),$$

$$P_{\mathrm{IDX}} = (P == 0),$$

$$\boldsymbol{P} = \boldsymbol{R}_1 < f_{\mathrm{TDR}}. \tag{7}$$

式中, $\boldsymbol{P}$为由 0 和 1 组成的向量; $\boldsymbol{R}_1$和R_3为 [0, 1] 空间内均匀分布的随机向量; R_2为取值在 [0, 1] 区间的随机数; P_{IDX}为当向量 P满足条件 ($P == 0$) 时, 取随机向量$\boldsymbol{R}_1$; "~"表示取反; f_{TDR} 为自适应因子, 从 1 线性递减为 0, 表达式如下:

$$f_{\mathrm{TDR}} = 1 - t \times (1/T_{\max}). \tag{8}$$

式中, t为当前迭代的次数, $T_{\max}$为最大迭代次数.

(3) 种群杂交行为. 种群内的马驹会在成熟后离开原种群, 与其他种群的年轻个体进行繁殖, 产生新的个体, 有

$$X_{G,k}^{p} = (X_{G,i}^{q}, X_{G,j}^{z})_{\mathrm{crossover}},\ i \neq j \neq k,\ q = z = S_{\mathrm{end}}. \tag{9}$$

式中, $X_{G,k}^{p}$为新繁殖的个体的位置, 通过对其父母个体位置取算术平均值来确定; $X_{G,i}^{q}$和$X_{G,j}^{z}$为对应该个体父母离开原种群后的位置信息; S_{end} 表示结束.

(4) 种群移动. 种群移动行为即是在寻找全局最优解, 种马领导种群前往更加合适的栖息地, 以种马位置信息为代表的种群位置更新公式为

$$\overline{N}_{SG_i} = \begin{cases} 2Z\cos(2\pi RZ) \times (P_{\mathrm{WH}} - N_{SG_i}) + P_{\mathrm{WH}}, \\ \qquad\qquad R_3 > 0.5; \\ 2Z\cos(2\pi RZ) \times (P_{\mathrm{WH}} - N_{SG_i}) - P_{\mathrm{WH}}, \\ \qquad\qquad R_3 \leqslant 0.5. \end{cases} \tag{10}$$

式中, $\overline{N}_{SG_i}$为第 i个种群中种马的候选位置, P_{WH} 表

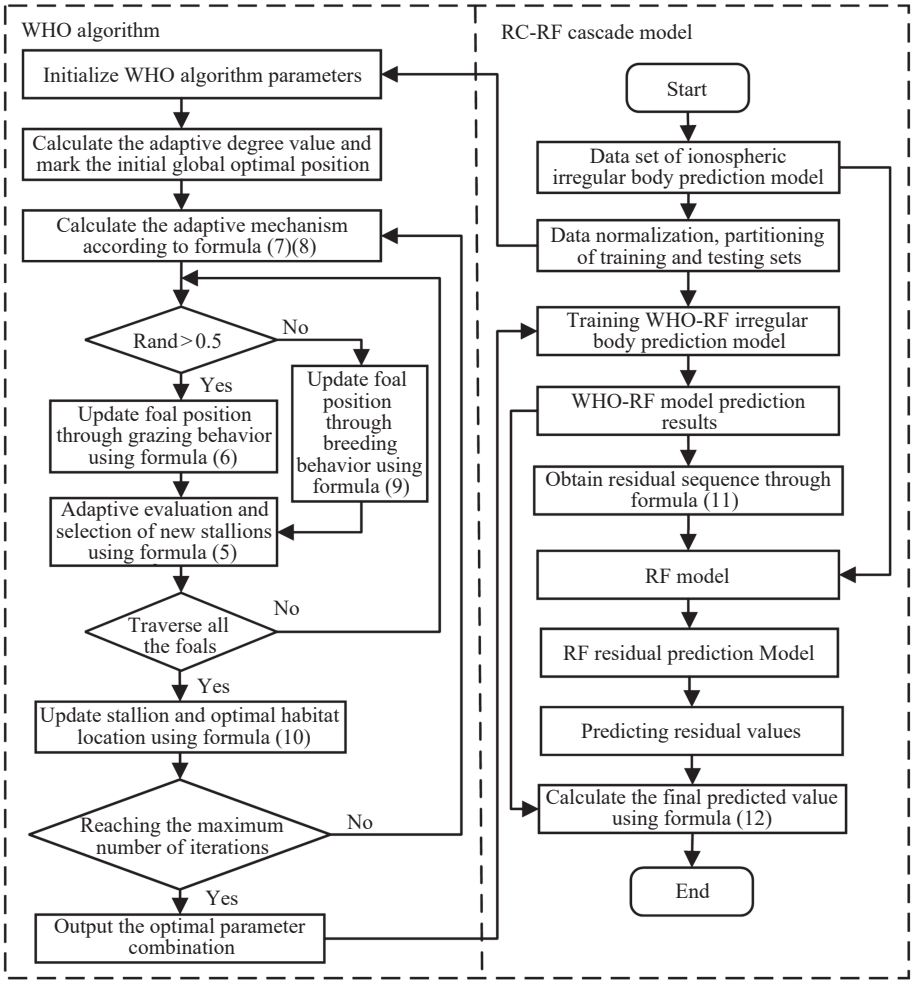

图 4 RC-WHO-RF 电离层不规则体预测模型流程

Fig. 4 Flow chart of RC-WO-RF ionospheric irregular body prediction model

示当前最优栖息地位置.

1.2.3 RC 算法

针对大多数时候 RF 模型存在预测值偏小, 尤其是在峰值区域误差显著的问题, 提出使用 RF 模型展开残差补偿 (RC) 级联预测的方法. 首先通过 RF 模型获得初始预测结果, 随后获取 ROTI 预测值向量$\hat{y}_{\mathrm{RF}}$与真实观测值向量y之间的残差序列向量ϵ, 公式为

$$\epsilon = y - \hat{y}_{\mathrm{RF}}. \tag{11}$$

接着, 利用该残差序列训练一个专门的 RF 残差预测模型; 最终, 系统的预测输出$\hat{y}_{\mathrm{final}}$由 RF 模型的初始预测值$\hat{y}_{\mathrm{RF}}$与残差模型的预测补偿值$\hat{\epsilon}$加权叠加构成, 公式为

$$\hat{y}_{\mathrm{final}} = \hat{y}_{\mathrm{RF}} + x\hat{\epsilon}. \tag{12}$$

1.2.4 RC-WHO-RF 预测模型

提出 RC-WHO-RF 电离层不规则体预测模型, 其流程如图 4 所示. 模型输入为 11 个与不规则体发生相关的参数, 将一个月的数据集分为训练集和测试集, 其中训练集占所有数据的 90%, 测试集占 10%, 输出为 1 个可表征不规则体发生与强度的 ROTI 指数; 利用 WHO 算法寻优能力强、收敛速度快的优点对 RF 模型的参数进行寻优; 并通过使用 RF 模型展开 RC 级联预测的方法, 获取最终预测结果. 该组合预测模型可以更好地适应电离层不规则体的变化和复杂性, 并提升峰值区域的预测精度.

1.3 评估指标

实验从不同的角度对 RC-WHO-RF 模型预测电离层不规则体的性能进行验证和分析, 选取决定系数 (R^2)、均方根误差 RMSE ($\varepsilon_{\mathrm{RMSE}}$)、平均绝对误差 MAE ($\varepsilon_{\mathrm{MAE}}$)、平均相对精度 MRA ($\varepsilon_{\mathrm{MRA}}$) 作为评价指标, 定义如下:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - y_{\mathrm{p}i})^2}{\sum_{i=1}^{n}(y_i - \overline{y}_i)^2}, \tag{13}$$

$$\varepsilon_{\mathrm{RMSE}} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_{\mathrm{p}i} - y_i)^2}, \tag{14}$$

$$\varepsilon_{\mathrm{MAE}} = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - y_{\mathrm{p}i}\right|^2, \tag{15}$$

$$\varepsilon_{\mathrm{MRA}} = \left(1 - \frac{1}{n}\sum_{i=1}^{n}\frac{|y_{\mathrm{p}i} - y_i|}{y_i}\right) \times 100\%. \tag{16}$$

式中, y_i和$y_{\mathrm{p}i}$分别为第 i 个 ROTI 真实值和 ROTI 预测值, n 为样本总数.

2 实验结果与分析

2.1 输入参数的相对重要性

为研究各个输入参数对预测电离层不规则体的影响, 使用 RF 算法, 通过测量模型精度的降低来评估每个特征的重要性. 图 5 给出不同输入特征的相对重要性排序结果: ROTI 凭借其强自相关性成为最重要的特征, 反映了历史 ROTI 序列对未来预测的主导作用; HHC 作为余弦相位的日变化指数, 其高排名可以解释为电离层不规则体的发生与日落后 (12:00 UTC) 电离层不稳定性增强相关, 与之相反的正弦相位日变化指数 HHS 因与关键高发时段匹配性较差排名最后. 此外, Dst, B_y 和 FP 也在前 5 个重要参数内, 而 B_x/B_z 和 SYM-H 影响相对有限, 可能因其效应已被其他特征间接涵盖, $F_{10.7}$ 则因主要反映长期太阳活动背景变化, 在短临预测中贡献表现较弱. 值得注意的是, f_0F_2 在本模型中的重要性排名较低, 这与文献 [15] 中强调的 f_0F_2 对预测电离层不规则体的重要作用矛盾, 导致该差异的原因可能是 f_0F_2 数据来源的空间局限性, 虽然与香港站同属赤道异常北驼峰区域, 但是站点间 130 km 的距离可能导致局部电离层状态的细微差异, 尤其是在不规则体的空间覆盖与漂移运动中, 这些差异可能降低了 f_0F_2 的贡献.

2.2 模型的参数与权重配置

2.2.1 WHO-RF 模型的参数设置

在提取特征之后, 设置寻优维度为 2, 寻优参数包括决策树数量和最大深度, 采用 WHO 算法对

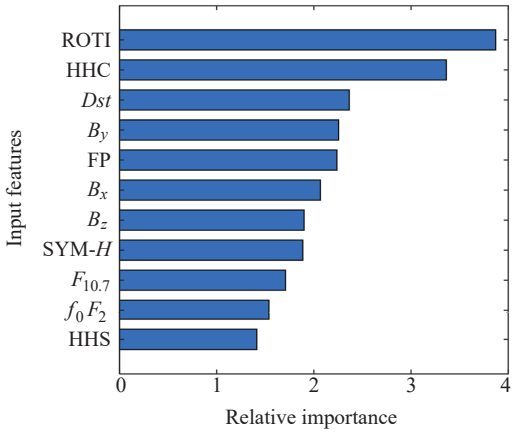

图 5 基于 RF 方法的输入参数重要性排序

Fig. 5 Importance ranking of input parameters based on random forest method

表 2 不同参数组合下的 WHO-RF 模型性能及寻优时间

Table 2 Performance and optimization time of WHO-RF model under different parameter combinations

Maxstep	10			20			30		
Pop	10	20	30	10	20	30	10	20	30
RMSE/(TECU·min^{-1})	0.110	0.103	0.098	0.105	0.098	0.099	0.100	0.097	0.092
MAE/(TECU·min^{-1})	0.068	0.059	0.058	0.061	0.057	0.058	0.058	0.056	0.054
R^2	0.780	0.800	0.803	0.782	0.803	0.799	0.798	0.805	0.810
Time/s	129	331	760	198	545	778	1013	1649	1797

RF 模型进行超参数寻优过程中, 需要人为确定种群数量和最大迭代次数两个参数. 为了确定最佳参数, 选取 2024 年 DOY 61−91 的不规则体数据输入到模型中进行寻优预测, 将种群规模 Pop 和最大迭代次数 Maxstep 均分别设置为 10, 20 和 30, 得到不同迭代次数下不同种群规模模型的性能指标值及寻优所需时间, 如表 2 和图 6 所示.

图 6 中 M10 P10 表示最大迭代次数 Maxstep=10、种群规模 Pop=10, 其他表示方法以此类推. 由表 2 和图 6 可知, 迭代次数为 20、种群规模为 20, 以及迭代次数为 30、种群规模为 20 和 30 的情况下, 模型的性能指标表现较佳, 虽然增大种群规模或迭代次数可以提高搜索能力, 但是相应的计算成本显著增加, 而预测精度并未获得显著性提高. 基于效率与性能的平衡考量, 这里最终采用 Maxstep=20, Pop=20 作为 WHO 的参数设置.

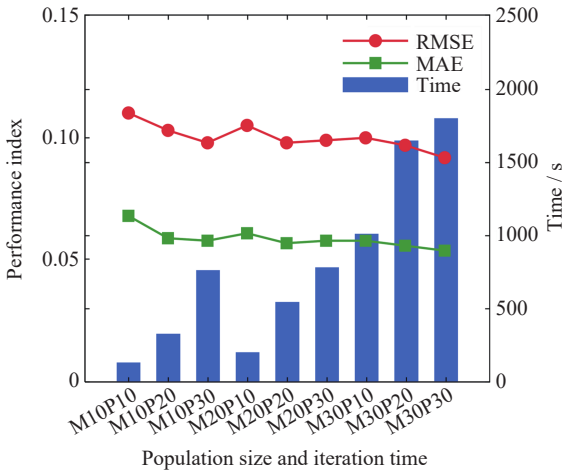

图 6 不同参数组合下的 WHO-RF 模型性能及寻优时间

Fig. 6 Performance and optimization time of WHO-RF model under different parameter combinations

2.2.2 RC-WHO-RF 模型权重值设置

为确定 RC-WHO-RF 模型的残差补偿值最优权重配置, 实验选取 2024 年 DOY 61−91 (3 月不规则体高发期) 和 DOY 153−182 (6 月不规则体低发期) 的数据进行建模分析. 如表 3 所示, 在高发期条件下, 当残差权重倍数取为 4 时, 模型的性能指标最优; 而在低发期条件下, 残差权重倍数取为 5 时展现出最佳补偿效果. 因此综合考虑, 在本文中当不规则体处于高发期 (春秋季) 时残差补偿倍数取 4, 在不规则体低发期 (夏冬季) 时, 残差补偿倍数取 5.

2.3 不同季节下三种模型效果对比

为评估不同不规则体发生率下 ROTI 预测模型的整体误差分布, 对三种模型在不规则体高发生率 3 月和低发生率 6 月进行比较分析[16]. 以 1 个月为建模周期, 将每个月的最后 3 天作为测试集, 使用 ROTI > 0.5 TECU·min^{-1} 的阈值来确定电离层不规

表 3 不同残差权重的 RC-WHO-RF 模型预测结果

Table 3 Prediction results of RC-WHO-RF models with different residual weights

不规则体发生率	高			低		
残差权重倍数	3	4	5	4	5	6
RMSE/(TECU·min^{-1})	0.059	0.056	0.074	0.082	0.075	0.093
MAE/(TECU·min^{-1})	0.042	0.037	0.048	0.044	0.038	0.063
R^2	0.902	0.935	0.891	0.843	0.864	0.797

则体的发生, 在保证数据具有代表性的同时, 使数据对比更加直观, 对 800 多个历元进行了步长为 30 的抽样. 图 7 显示了三种预测模型在 2024 年 3 月 29−31 日 ROTI 预测值与真实值的差异对比, 在闪烁高发期, 29−31 日 ROTI 的真实值皆大于 0.5 TECU·min^{-1}, 表明这 3 天都有不规则体的发生, 并且 RC-WHO-RF 模型的预测更接近真实值, 而 RF 和 WHO-RF 模型预测值在峰值处则存在较大的误差. 图 8 显示了各模型 ROTI 预测值的 R^2, RMSE 和 MAE 的分布, 可以看出, 单一 RF 模型预测的结果在三个性能指标中都是最差的; 使用优化算法的 WHO-RF 模型各性能指标都比 RF 有所提升, 但是提升效果有限; 相比之下, RC-MHO-RF 模型的预测性能实现了显著提升, 其R^2高达 0.935, 较 RF 提升 17.8%, 较 WHO-RF 提升 11.5%; RMSE 降低至 0.056; MAE 仅为 0.037. 这一结果表明, 经过残差补偿和算法优化的 RC-WHO-RF 模型具备更好的预测能力.

图 9 对比了 2024 年 6 月 28−30 日各种预测模型的 ROTI 预测值与真实值. 结果显示, 在通常的电离层闪烁低发期, 28 日却发生了中等强度的不规则体事件, 经分析是太阳上的暗条爆发伴随的日冕物质抛射到达地球引起的, 当天 19:54 LT, 地磁扰动达到极强地磁暴水平 (Kp = 8, Dst = –107 nT), 传统 RF

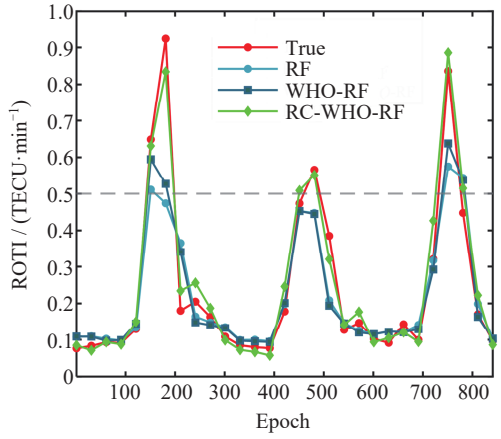

图 7 2024 年 3 月 29−31 日 ROTI 预测值与真实值的对比

Fig. 7 Comparison between ROTI forecast and actual values from 29 to 31 in March 2024

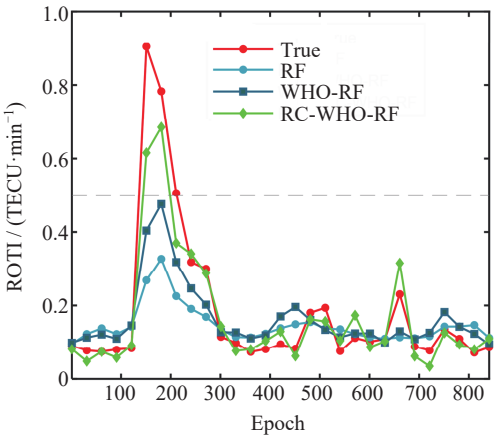

图 9 2024 年 6 月 28−30 日 ROTI 预测值与真实值对比

Fig. 9 Comparison between ROTI forecast and actual values from 28 to 30 in June 2024

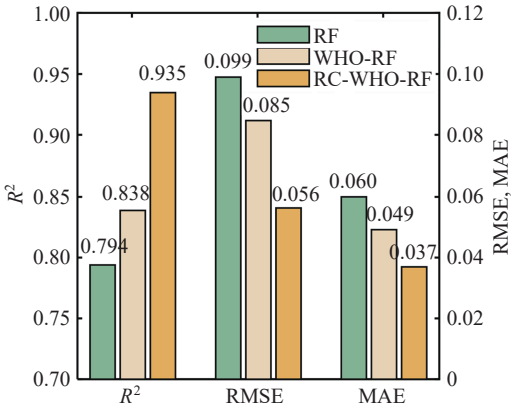

图 8 2024 年 3 月 29−31 日 ROTI 预测值的 R^2, RMSE 和 MAE 分布

Fig. 8 Distribution of R^2, RMSE and MAE for ROTI forecast values from 29 to 31 in March 2024

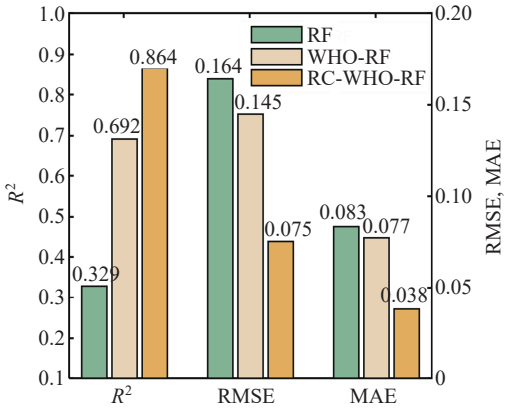

图 10 2024 年 6 月 28−30 日 ROTI 预测值的 R^2, RMSE 和 MAE 分布

Fig. 10 Distribution of R^2, RMSE and MAE for ROTI forecast values from 28 to 30 in June 2024

表 4 RC-WHO-RF 模型提前不同时间的预测性能

Table 4 RC-WHO-RF model predicts performance at different advance time

评价指标	提前时间/min						
	10	30	60	90	120	180	240
RMSE/(TECU·min^{-1})	0.026	0.039	0.043	0.048	0.053	0.059	0.066
R^2	0.987	0.972	0.965	0.962	0.951	0.949	0.942
MRA/(%)	91.94	90.67	89.15	87.40	86.97	85.54	82.79

表 5 太阳活动低/高年模型预测效果对比

Table 5 Comparison of prediction results between low/high solar activity year models

时段	模型	指标		
		RMSE/(TECU·min^{-1})	MAE/(TECU·min^{-1})	MRA/(%)
2020年3月	RF	0.036	0.016	85.40
	WHO-RF	0.028	0.013	87.59
	RC-WHO-RF	0.015	0.007	92.98
2024年3月	RF	0.099	0.060	79.47
	WHO-RF	0.085	0.049	82.51
	RC-WHO-RF	0.039	0.037	90.67

模型和 WHO-RF 对突发事件的响应明显不足, 其预测值显著低于真实观测值, 表明这些模型对极端空间天气条件下的电离层扰动不敏感. 相比之下, 改进后的 RC-WHO-RF 模型其预测曲线与真实值较为吻合, 成功捕捉到此次不规则体事件, 验证了该模型对于突发性电离层扰动事件具有优异的响应能力和预测精度. 在随后的 29 日和 30 日, 随着地磁活动逐渐恢复平静, 各模型之间的预测误差逐渐缩小. 图 10 显示各模型 ROTI 预测值的R^2, RMSE 和 MAE 的分布, 由于突发的磁暴事件, RC-WHO-RF 模型各项指标与其在 3 月份构建的模型性能指标有所下降, 但是仍然优于其他两个模型, 再次验证了 RC-WHO-RF 模型不仅预测精度高, 且对于电离层突发事件也具有优异的响应能力.

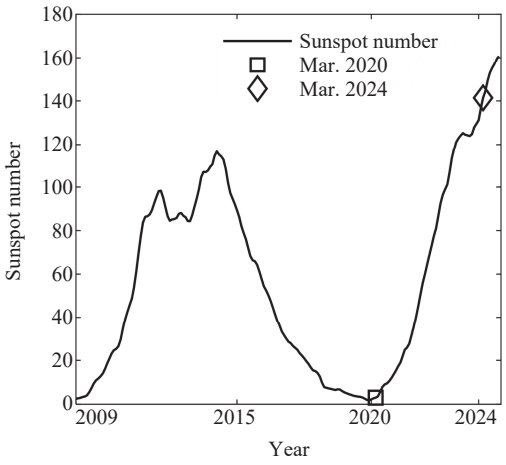

图 11 2009－2024 年的太阳黑子数变化

Fig. 11 Changes in sunspot numbers from 2009 to 2024

2.4 提前不同时间预测性能对比

为评估 RC-WHO-RF 模型在短临预报中提前不同时间的预测性能, 这里以一个月为建模周期, 将 3 月的最后一天作为测试集, 对 2024 年 3 月 31 日的 ROTI 值进行提前不同时间的预测验证, 时间跨度范围为 10～240 min (见表 4), 在 240 min 内的短临预报中模型均展现出优异的预测性能. 值得注意的是, 模型在 30 min 以内的短临预报中表现较为突出, 平均相对精度超过 90%; 在 30～180 min 保持较高精度, 平均相对精度大于 85%; 即使在提前时间为 240 min 的短临预报中仍维持良好性能, 平均相对精度在 80% 以上.

2.5 太阳活动高/低年模型效果对比

为评估提出的 RC-WHO-RF 电离层不规则体预测模型在不同太阳活动条件下的性能表现, 选取第

25 轮太阳活动周期中具有代表性的两个时间节点, 太阳活动极高年 2024 年 3 月和太阳活动极低年 2020 年 3 月的数据进行建模. 图 11 给出了 2009−2024 年采用 13 个月平滑的太阳黑子数变化.

以当月前 27 天的数据作为训练集训练模型, 对后 3 天内的电离层扰动指数 ROTI 值进行预测 (见表 5). 结果表明, 2020 年太阳活动低年不规则体预测结果的性能指标基本优于 2024 年太阳活动高年, 在太阳活动高年期间太阳释放的多种能量表现为电磁辐射、高能带电粒子流等, 先后到达地球空间, 与地球磁层发生复杂的相互作用, 显著增加了电离层环境的动态复杂性, 导致等离子体不规则体的时空演化特征更加难以预测[17]. 相比之下, 太阳活动低年期间电离层状态相对稳定, 不规则体指数变化较为平稳, 更易于模型捕捉其变化规律. 尽管如此, 无论在太阳活动高年还是低年, RC-WHO-RF 模型其各项评价指标显著优于单一 RF 和 WHO-RF 模型.

3 结论

为解决单一 RF 模型在预测电离层不规则体时存在参数调整困难且预测精度不高的问题, 利用第 25 个太阳活动周期香港 HKWS 站 2020 年 3 月 1 日至 2024 年 6 月 30 日的不规则体相关数据建立 RC-WHO-RF 模型, 并从不同的角度对该模型的性能进行评估, 得到主要结论如下.

(1) 历史 ROTI 序列对预测未来占主导作用, 表征日变化的 HHC 以及 Dst, B_y 等电离层磁场参数共同构成了电离层不规则体预测的重要驱动因素.

(2) RC-WHO-RF 模型对电离层不规则体预测的均方根误差小于 0.1 TECU·min^{-1}, 在不规则体高发生率的 3 月份, 其 R^2 高达 0.935, RMSE 降低至 0.056, MAE 仅为 0.037; 在不规则体低发生率的 6 月份, 其不仅预测精度高, 对于突发性磁暴事件也具有优异的响应能力和预测精度.

(3) RC-WHO-RF 模型在提前 30 min 的短临预报中平均相对精度达 90.67%, 比 WHO-RF 提升 8.16%, 比单一 RF 模型提升 11.2%, 组合模型的预测效果明显优于单一 RF 模型和 WHO-RF 模型, 且在提前 180 min 的短临预报中其平均相对精度仍达 85% 以上.

由于不规则体数据获取途径的限制, 对低纬区域性的预测建模研究仍有待进一步扩展; 后续还将深入分析更多可用的 $h_{\mathrm{m}}F_2$ (F_2 层峰值高度) 和 f_0F_2 数据, 探讨其特征重要性及如何用于对电离层不规则体进行预测, 并进一步探讨将预测得到的电离层不规则体信息应用于高精度 GNSS 载波相位定位技术, 以提高定位解算的准确度.

致谢 电离层科学数据由国家重大科技基础设施项目子午工程及 NASA/GSFC 空间物理数据设施的 OMNIWeb 服务提供, GNSS 观测数据由 International GNSS Service (IGS) 提供.

利益冲突 所有作者均声明不存在利益冲突.

参考文献

[1] BASU S, MACKENZIE E, BASU S. Ionospheric constraints on VHF/UHF communications links during solar maximum and minimum periods[J]. *Radio Science*, 1988, **23**(3): 363-378

[2] SHANG Sheping, SHI Jiankui, CHENG Zhengwei, *et al*. Event analysis of the influence of strong tropospheric convection in summer on low-latitude ionospheric irregularities in East Asia[J]. *Chinese Journal of Space Science*, 2022, **42**(6): 1111-1121 (尚社平, 史建魁, 程征伟, 等. 夏季强对流活动对东亚低纬电离层不规则体影响的事件分析 [J]. 空间科学学报, 2022, **42**(6): 1111-1121)

[3] ZHANG Hongbo, WANG Feifei, HAO Shuji, *et al*. Occurrence probability prediction model of ionospheric scintillation for the low-latitude region[J]. *Chinese Journal of Radio Science*, 2023, **38**(4): 691-697 (张红波, 王飞飞, 郝书吉, 等. 低纬地区电离层闪烁发生概率预报模型研究 [J]. 电波科学学报, 2023, **38**(4): 691-697)

[4] PI X, MANNUCCI A J, LINDQWISTER U J, *et al*. Monitoring of global ionospheric irregularities using the worldwide GPS network[J]. *Geophysical Research Letters*, 1997, **24**(18): 2283-2286

[5] LIU L, MORTON Y J, LIU Y X. Machine learning prediction of storm-time high-latitude ionospheric irregularities from GNSS-derived ROTI maps[J]. *Geophysical Research Letters*, 2021, **48**(20): e2021GL095561

[6] TETE S, OTSUKA Y, ZAHRA W K, *et al*. Leveraging machine learning techniques and GPS measurements for precise TEC rate predictions[J]. *GPS Solutions*, 2024, **28**(3): 115

[7] HABYARIMANA V, HABARULEMA J B, OKOH D, *et*

al. Single station modelling of ionospheric irregularities using artificial neural networks[J]. *Astrophysics and Space Science*, 2023, **368**(12): 105

[8] ATABATI A, ALIZADEH M, SCHUH H, *et al*. Ionospheric scintillation prediction on S4 and ROTI parameters using artificial neural network and genetic algorithm[J]. *Remote Sensing*, 2021, **13**(11): 2092

[9] ZHAO X K, LI G Z, XIE H Y, *et al*. A novel short-term prediction model for regional equatorial plasma bubble irregularities in East and Southeast Asia[J]. *Space Weather*, 2025, **23**(2): e2024SW004224

[10] WU Jiayan, ZHEN Weimin, OU Ming, *et al*. A characterization method of global ionospheric irregularities activities based on GNSS data[J]. *Chinese Journal of Space Science*, 2022, **42**(1): 82-90 (吴家燕, 甄卫民, 欧明, 等. 一种基于GNSS 数据的全球电离层不规则结构活动表征方法 [J]. 空间科学学报, 2022, **42**(1): 82-90)

[11] REZENDE L F C, DE PAULA E R, STEPHANY S, *et al*. Survey and prediction of the ionospheric scintillation using data mining techniques[J]. *Space Weather*, 2010, **8**(6): S06D09

[12] LIU Lei. Studying on Ionospheric Modeling Using Multi Source Data and Ionospheric Responses to Magnetic Storms[D]. Wuhan: Wuhan University, 2017 (刘磊. 基于多源数据的电离层建模及暴时电离层扰动分析 [D]. 武汉: 武汉大学, 2017)

[13] BREIMAN L. Random forests[J]. *Machine Learning*, 2001, **45**(1): 5-32

[14] NARUEI I, KEYNIA F. Wild horse optimizer: a new meta-heuristic algorithm for solving engineering optimization problems[J]. *Engineering with Computers*, 2022, **38**(4): 3025-3056

[15] LIU H, YANG P X, REN X D, *et al*. The short-term prediction of low-latitude ionospheric irregularities leveraging a hybrid ensemble model[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2024, **62**: 4100615

[16] YANG Yunzhen, LIU Lilong, MO Shilei, *et al*. Ionospheric TEC variation in Guangxi and surrounding during high and low solar activity years[J]. *Chinese Journal of Space Science*, 2024, **44**(6): 1021-1030 (杨芸珍, 刘立龙, 莫诗蕾, 等. 太阳活动高低年广西及周边地区电离层时空频域特性 [J]. 空间科学学报, 2024, **44**(6): 1021-1030)

[17] WANG Kai, LIU Ang, GAO Weiguang, *et al*. Analysis on open service performance of BDS during peak of solar cycle 25[J]. *Journal of Navigation and Positioning*, 2025, **13**(1): 61-68 (王凯, 刘昂, 高为广, 等. 太阳活动高峰期 BDS 服务性能分析 [J]. 导航定位学报, 2025, **13**(1): 61-68)

作者简介

陈清燕 女, 2001 年 6 月出生于广西梧州市, 现为桂林电子科技大学信息与通信学院硕士研究生, 主要研究方向为卫星导航电离层时空演变规律及预测.
E-mail: 2572642573@qq.com

纪元法（通信作者） 男, 1975 年 12 月出生于山东聊城, 现为桂林电子科技大学信息与通信学院教授, 博士生导师, 主要研究方向为卫星导航、信号处理.
E-mail: jiyuanfa@163.com