

doi: 10.3969/j.issn.1000-8349.2025.02.XX

系外行星凌星信号检测的去趋势技术进展

陈雨^{1,2}, 葛健¹

(1. 中国科学院 上海天文台, 上海 200030; 2. 中国科学院大学, 北京 100049)

摘要:

凌星法已成为探测太阳系外行星最有效的方法之一,但其精度依赖于从高精度光度数据中可靠提取微弱的凌星信号。尤其是 Kepler 与 TESS 等空间望远镜获得的光变曲线中,常包含由仪器效应和恒星光变信号引起的复杂噪声,其幅度往往远大于类地行星的凌星深度,从而对行星探测与精确表征构成主要障碍。为此,去趋势(detrending)成为光度数据处理中至关重要的预处理步骤,其目标是在尽可能保留真实凌星信号的同时,去除非凌星成分并提升信噪比。

本文系统梳理了当前主流的去趋势技术,并将其概括为三类。第一类是应用最广泛的“先去趋势后搜索”方法,包括滑动中位数和平滑样条等平滑拟合技术、基于小波或余弦的信号分解方法,以及高斯过程和自回归模型等概率与时序建模方法。第二类是联合建模方法,通过同时拟合噪声与凌星信号,减少传统分步处理中可能引入的信号失真。第三类是数据驱动的“直接凌星探测”方法,利用深度学习模型直接从含噪光变曲线中识别凌星事件。作为相关案例,本文还介绍了基于 DeepLabv3+ 架构构建的 ProDeNet。虽然其测试样本量过小,但初步结果表明,分割式直接探测方法在低信噪比凌星事件识别中具有应用潜力。

综合不同方法的性能比较可见,目前尚不存在通用的最优去趋势算法。最适方案仍取决于科学目标、恒星类型与观测数据质量。本文总结了各类方法的适用条件与发展方向,可为未来高精度光变数据分析与凌星信号提取提供参考。

关键词: 系外行星; 凌星法; 去趋势; 光变曲线; 联合建模; 深度学习

中图分类号: P185 **文献标识码:** A

资助项目: 太空探源专项

通讯作者: 葛健, jge@shao.ac.cn

1 引言

凌星法是探测太阳系外行星最有效的方法之一^[1]。其基本原理是：当行星从地球与宿主恒星之间经过时，会短暂遮挡部分星光，从而使恒星亮度产生微弱且周期性的下降。尽管这一思想可追溯至早期凌日观测及相关讨论，但直到高精度光度测量技术发展成熟后，凌星法才真正成为可实施的系外行星探测手段^[2]。1999 年，HD209458 b 的凌星现象被成功观测到^[3,4]，标志着凌星法在系外行星探测中的实际突破，并推动了空间凌星巡天的发展。

此后，MOST (Microvariability and Oscillations of STars)、CoRoT、Kepler、TESS 和 CHEOPS 等任务相继开展或持续推进高精度测光观测^[5-9]。其中，CoRoT 和 Kepler 推动了空间高精度凌星巡天的发展，并系统揭示了系外行星系统的多样性^[10-13]；TESS 面向几乎全天空开展巡天，显著扩展了明亮近邻恒星样本^[14,15]；CHEOPS 则更侧重于已知行星系统的高精度后续测光表征^[16,17]。未来，PLATO、Roman 及 ET 等任务将进一步拓展系外行星研究的深度与广度^[18-25]。这些任务获取或将获取大量高质量光变曲线，为系外行星的发现与精确表征提供基础，同时也对噪声建模与去趋势方法提出了更高要求。

在实际观测中，光变曲线不可避免地受到多种噪声和系统误差的影响，这些干扰往往会掩盖或削弱微弱的凌星信号^[26]。对于地基观测，大气扰动长期是限制测光精度的主要因素^[27]；空间观测虽然避免了大气影响，却仍面临基本噪声、仪器噪声以及恒星光变信号等复杂干扰^[28]。以太阳型恒星为例，一颗地球大小行星产生的凌星深度仅约为 84 ppm，这要求观测与数据处理流程具备极高的精度和稳定性^[28]。因此，如何在保留真实凌星信号的同时有效去除非凌星成分，成为系外行星测光分析中的关键问题。

在相关文献中，常需区分 cotrending（共趋势）与 detrending（去趋势）两个概念。前者主要用于移除由望远镜或探测器引入的共性系统误差。其核心假设是这些仪器效应在所有或部分目标星的光变曲线中是共享的，或者与辅助工程数据高度相关，这种方法在保留目标特有的恒星光变信号方面更具优势^[29-31]。后者则通常指针对单颗恒星光变曲线，进一步去除低频变化成分，而不严格区分其来源是仪器效应还是恒星活动^[29]。在完成 cotrending 后，光变曲线中往往仍残留显著的恒星光变信号，其幅度通常远大于行星凌星信号。为了提高凌星搜索的信噪比，必须通过 detrending 去除这些低频变化，使数据更适合后续搜索与参数测量^[32]。因此，去趋势已成为凌星法数据分析流程中的核心预处理步骤。

本文聚焦于系外行星光变曲线中的去趋势问题，对当前主流方法进行系统综述。第二章介绍光变曲线中的主要噪声来源；第三章讨论应用最广泛的“先去趋势后搜索”范式；第四章评述联合建模方法；第五章介绍直接凌星探测这一新兴方向，并结合 ProDeNet 进行说明；第六章讨论面向凌星参数精测的去趋势策略；第七章总结全文并展望未来发展方向。

2 噪声

在系外行星探测中，深入理解光变曲线中的噪声来源是设计去趋势策略的基础。Kepler 任务以其前所未有的测光精度，为噪声分析提供了经典框架，其核心衡量指标为综合差分

表 1 Kepler 任务对典型 12 等星的噪声预算分解^[33]

噪声成分	噪声幅度 (ppm)	设计预算 (ppm)
恒星光变信号	19.5	10.0
基本噪声 (光子噪声 + 读出噪声)	16.8	14.1
仪器噪声 (探测器噪声 + 季度噪声)	13.3*	10.0
总计	29.0	20.0

注: 仪器噪声由内在探测器噪声 (10.8 ppm) 和季度相关噪声 (7.8 ppm) 以平方和的平方根形式 ($\sqrt{10.8^2 + 7.8^2}$) 合成。所有数据均来自 Gilliland 等人 (2011) 的研究。

测光精度 (combined differential photometric precision, CDPP)^[28]。总体而言, 影响凌星探测的噪声主要可分为三类: 基本噪声、仪器噪声和恒星光变信号。其中, 恒星光变信号通常是主导因素^[33]。表 1 和图 1 分别展示了 Kepler 对典型 12 等星的噪声预算分解, 以及不同噪声源在时域与频域中的典型表现。

各类噪声的相对重要性随恒星亮度而变化。对于作为 Kepler 设计基准的中等亮度恒星 ($K_p \approx 12$), 恒星光变信号是最大的噪声贡献项, 其次是由光子噪声和读出噪声组成的基本噪声^[33]。因此, 去趋势方法真正面临的挑战, 往往并非简单的随机噪声抑制, 而是在尽量保留凌星信号的同时, 处理具有时间相关性和物理复杂性的背景变化。下面分别对这三类噪声进行讨论。

2.1 基本噪声

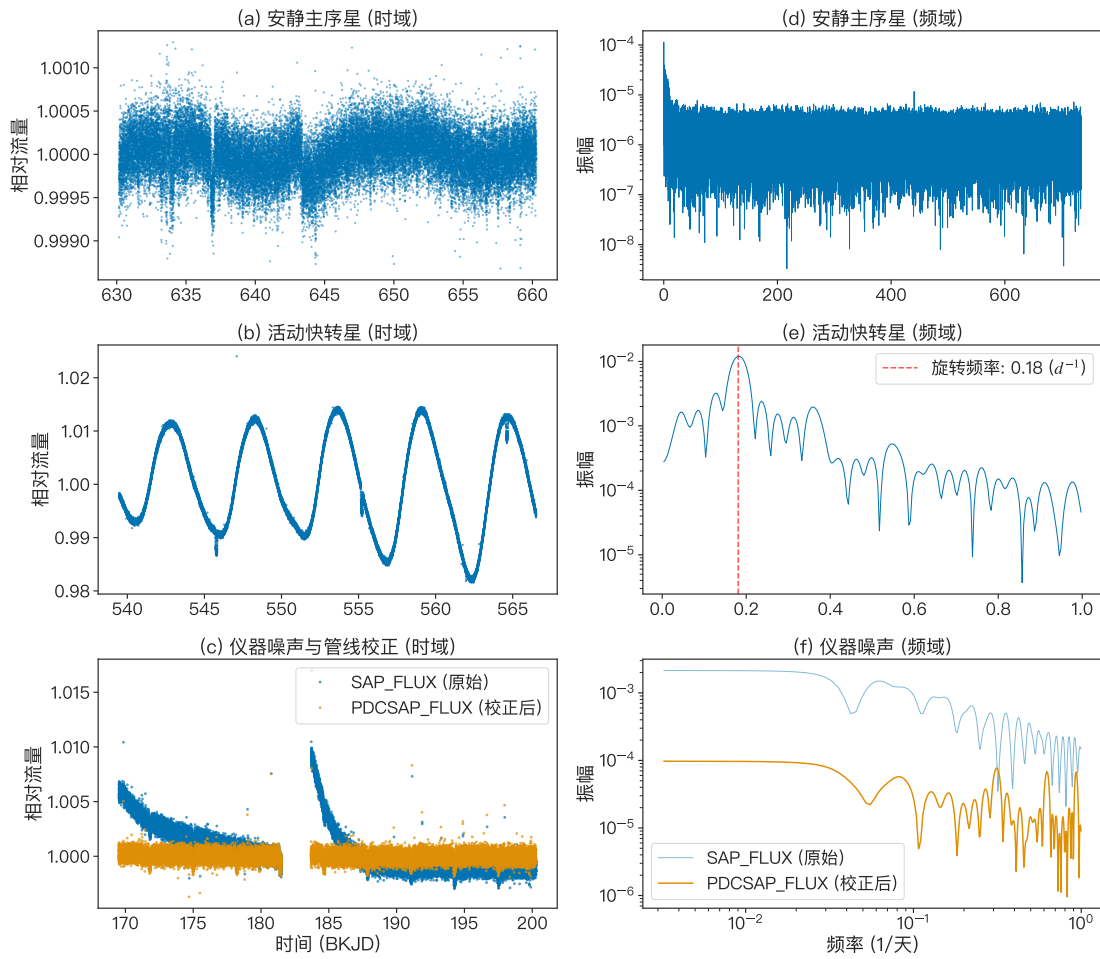
基本噪声是光度测量中无法通过去趋势算法消除的随机噪声背景, 它共同决定了探测灵敏度的理论上限。其主要来源包括目标恒星的光子噪声、天空背景以及探测器读出噪声^[33]。在 Kepler 的优异设计和低温工作环境下, 杂散光和暗电流等项通常相对较小^[34]。对于一颗典型的 12 等 G2V 型恒星, 仅恒星光子噪声在 6.5 小时积分尺度上便可达到约 15 ppm, 而读出噪声则构成次一级贡献^[33]。

尽管这部分噪声本身不具备明显的时间相关结构, 因而不属于去趋势算法的主要处理对象, 但它为凌星探测设定了不可逾越的物理下限。尤其对于类地行星这类浅凌星信号, 基本噪声往往已占据相当比例, 因此任何进一步的数据处理都必须建立在这一噪声底限之上。

2.2 仪器噪声

仪器噪声是由航天器平台、光学系统及其运行环境引入的系统误差, 通常首先由 cotrending 处理, 但其残留仍会进入后续 detrending 阶段。本节将对 Kepler 任务中的仪器噪声进行分类讨论, 重点关注航天器指向不稳定性与热效应引起的焦点变化。

(1) 航天器指向的不稳定性包含高频的抖动 (jitter) 和低频的漂移 (drift)^[34]。抖动通常幅度很小, 对光度精度的影响有限^[34]。更具挑战性的是低频漂移, 它指望望远镜视线的缓慢、系统性移动。这种漂移的来源多样, 包括任务早期的导星系统问题、由航天器轨道运动引起的差分速度像差 (differential velocity aberration, DVA)、定期的姿态微调, 望远镜温度变化导致的像移, 以及因进入安全模式或数据下行导致的指向中断^[33-35]。这些效应共同在光变曲线中引入了缓慢变化的低频趋势和突然的不连续跳变, 是去趋势算法必须有效识别和移除的



注：左列为时域光变曲线，右列为基于 **Lomb-Scargle** 方法计算的频域振幅谱。(a) 和 (d)：展示了一颗安静主序星 (Kepler-10)。其时域噪声近似随机，对应的振幅谱在频域上相对平坦，呈现近似白噪声特性。(b) 和 (e)：展示了一颗活跃恒星 (Kepler-63)。其时域上的准周期性波动源于星斑引起的旋转调制，这在频域上表现为自转频率处的一个高信噪比尖峰。(c) 和 (f)：对比了 Kepler-4 未经校正的原始通量 (SAP_FLUX) 与经过官方管线校正后的通量 (PDCSAP_FLUX)。结果显示，管线校正在很大程度上抑制了仪器趋势，这在频域上体现为低频振幅减弱。

图 1 开普勒 (Kepler) 光变曲线中不同噪声源的时域表现与频域特征。

关键特征。

(2) 焦点变化是 Kepler 数据中另一个重要的仪器效应，通常由望远镜组件的热效应引起^[36]。温度的波动，无论是来自反作用轮加热器的短周期循环（在早期任务数据中尤为明显），还是航天器绕日运行带来的季节性长周期变化，都会导致仪器的点扩散函数 (point spread function, PSF) 发生改变，从而引起实测光通量的变化^[36]。由于这些热效应在焦平面上具有空间相关性（即会同时影响多颗恒星），因此它们非常适合用共同趋势来建模。在 Kepler 官方处理的 PDCSAP_FLUX 数据中，这类系统误差已在很大程度上通过共同趋势基

向量 (cotrending basis vectors, CBVs) 进行了校正^[36]。

2.3 恒星光变信号

在探测系外行星时, 恒星光变信号是最重要的天体物理噪声来源, 其振幅范围极广, 从星斑、脉动等引起的百分之几变化, 到光球层对流引起的微弱波动, 都可能出现在观测中^[37]。与基本噪声和多数仪器噪声不同, 这类信号直接来源于恒星本身, 因此往往更难通过统一模型加以处理。理解其时间尺度、周期性和形态, 是选择有效去趋势策略的前提。

(1) 恒星脉动。恒星脉动由恒星内部声波和重力波的传播引起, 导致表面半径、温度和亮度发生周期性变化^[38]。从去趋势角度看, 可大致将其分为规则脉动和复杂脉动两类。前者通常由一个或少数几个稳定频率主导, 典型代表包括 Cepheids、RR Lyrae 和 β Cephei 等^[39-41]。后者则表现为多频叠加、短周期且往往非正弦的复杂变化, 典型如 δ Scuti 与 γ Doradus^[38, 42, 43]。总体而言, 脉动信号虽然形态多样, 但通常仍具有较强规律性, 因此在去趋势处理中往往可以作为确定性的周期趋势建模, 其对最终残余噪声的贡献相对有限。

(2) 恒星磁活动是低质量主序星 (如太阳) 表面产生磁场的结果, 表现为光球上的亮区 (如光斑) 和暗区 (如星斑)^[44, 45]。星斑会随恒星自转在视线中出现与消失, 引起准周期性、非正弦调制^[46]。星斑寿命依赖恒星类型与自转, 冷星通常更持久。耀斑是由磁重联触发的爆发事件, 表现为快速上升和指数衰减^[47], 有时叠加成复杂多峰形态^[48]。它们在光变曲线上表现为短时高振幅的尖锐信号, 需在去趋势处理中单独识别与处理。综合来看, 与脉动不同, 这类信号往往随时间演化, 其局部形态和幅度难以用简单解析模型准确描述, 因此是掩盖微弱凌星信号的主要困难来源之一。

(3) 双星相关变光。双星系统, 尤其是食双星, 是凌星搜索中的重要假阳性来源^[49-51]。从证认角度看, 这类信号需要与真实行星凌星仔细区分。但从去趋势角度看, 双星引起的变光通常具有较强周期性, 因此往往可以视作可建模、可剔除的周期性趋势, 其对基线噪声抑制本身的挑战通常小于恒星活动。

(4) 对流噪声与类太阳振荡。在具有对流包层的恒星中, 米粒组织会引入典型的红噪声, 其特征时间尺度在太阳上为数十分钟, 在红巨星上则可延长至数小时甚至数周^[52, 53]。由于其时间尺度有时与凌星持续时间相近, 这种局域扰动会严重干扰凌星检测。更重要的是, 米粒组织本质上属于随机现象, 缺乏简单的确定性函数形式, 因此可被视为高精度凌星测光中的噪声下限^[54-56]。针对这一问题, 已有研究利用高斯过程回归结合星震学功率谱特征对恒星光变信号进行建模, 并提高了行星与恒星半径比等参数测量的精度与准确性^[53]。相比之下, 类太阳振荡周期较短、振幅较低, 通常可通过 15–30 分钟分箱显著减弱其影响^[39, 44, 57]。

3 先去趋势后搜索

在系外行星凌星探测流程中, 去趋势—搜索是最常见的两阶段策略。其核心思路是: 首先利用去趋势算法对光变曲线进行预处理, 以移除与行星信号无关的光变成分, 从而提升信噪比。随后在平滑化的光变曲线上执行凌星搜索, 例如采用箱型最小二乘算法 (BLS)^[58]



图 2 系外行星光变曲线去趋势方法分类旭日图

或凌星最小二乘算法 (TLS)^[59] 进行周期信号检测。这类方法的优势在于结构清晰、计算高效,并能针对不同噪声类型采用相应的去趋势模型。然而,其主要挑战在于:若去趋势模型过于灵活,可能将微弱的凌星信号一并拟合掉,造成信号削弱或漏检。因此,如何在去除长期趋势与保留短时凌星之间取得平衡,是所有“先去趋势后搜索”策略的核心问题。本章将系统回顾这一范式下的主要技术路线,包括平滑与拟合、信号分解、概率建模、时序建模以及基于机器学习的方法。

3.1 平滑拟合

平滑与拟合方法在处理天文光变曲线时,通常侧重于通过局部或全局的函数逼近来消除噪声和系统效应。这些方法通常包括滑动滤波器、局部回归、多项式拟合和样条拟合等,它们各有优缺点,适用于不同特征的噪声和数据结构。其核心在于通过数学手段,在滤除基线趋势的同时,最大限度地保护真实的凌星信号。

3.1.1 滑动滤波器

滑动滤波器通过在数据序列上移动一个固定大小的窗口,并对窗口内的数据执行某种统计操作来平滑时间序列。这类方法因其概念简单、计算效率高而广泛应用于光变曲线的初步去趋势。

(1) 滑动均值和滑动中位数。最基础的滑动滤波方法是滑动均值和滑动中位数。滑动均值通过计算窗口内数据的平均值来拟合基线,但其对异常值极为敏感。尽管如此,一些研究仍会出于简洁性的考虑采用均值滤波,但通常需要配合严格的迭代剔除来保证数据质量^[60]。

相比之下, 滑动中位数对离群点具有天然的鲁棒性, 因此在系外行星探测中更为常用。例如, Handberg & Lund (2014)^[61] 在 KASOC 滤波器中便采用滑动中位数来移除长期趋势和恒星活动。然而, 滑动中位数的性能高度依赖于窗口长度的选择。若窗口过短, 滤波器会将凌星信号本身误当作趋势拟合掉, 导致信号深度变浅。若窗口过长, 则难以及时跟踪恒星活动的变化^[62]。Rizzuto 等人 (2017) 进一步指出, 在处理存在自转调制的恒星时, 如果不对凌星区间进行特殊保护, 例如采用凹槽滤波器 (notch filter), 传统的滑动中位数方法可能会削弱甚至完全掩盖较弱的凌星信号^[63]。此外, 该方法虽然在处理缓慢变化时表现良好, 但对于时间尺度较快或存在严重数据缺失的情况, 其效果会明显下降^[64]。除了作为主要去趋势工具外, 滑动中位数也常被单独用于识别和剔除离群点^[65-69]。

(2) Savitzky-Golay 滤波器。Savitzky-Golay 滤波器通过在数据序列的滑动窗口上拟合低阶多项式 (通常是 2 阶或 3 阶) 来进行平滑处理^[32]。与简单均值滤波相比, 它在保留局部曲线形态方面具有一定优势, 但在凌星探测中需谨慎使用。Hippke 等人 (2019) 指出, Savitzky-Golay 滤波器可能会过拟合并削弱凌星信号, 使凌星深度降低约 10%–20%^[32]。其根本原因在于, 这类局部最小二乘拟合会尝试“拉平”光变曲线中的突变结构, 从而错误地将凌星或耀斑的一部分吸收到趋势模型中。

尽管如此, Savitzky-Golay 滤波器在某些场景下仍具有实用价值, 尤其是作为预处理降噪步骤时。例如, 在 SHERLOCK 流程中, 该方法并不直接作为主去趋势器, 而是先用于初步平滑, 以降低局部高频白噪声, 再结合 biweight 等更鲁棒的方法进行后续处理^[70]。相关研究表明, 这种做法在个别低信噪比候选体的恢复中可以带来明显增益, 但也可能伴随更多假阳性, 因此需要谨慎审查^[70-72]。

(3) 基于 M-估计器的滑动滤波器。为了兼顾高斯噪声下的效率与对异常值的鲁棒性, 研究者引入了基于 M-估计器的滑动滤波器^[73]。其基本思想是在局部窗口内, 不再简单采用均值或中位数, 而是根据数据点偏离中心的程度动态调整其权重, 从而在平滑性与抗干扰性之间取得平衡。

其中, Huber 估计器在残差较小时采用类似均值的平方损失, 以保留较高效率, 在残差较大时转为类似绝对值损失, 以降低异常值影响。Tukey biweight 则进一步增强了鲁棒性, 当残差超过一定阈值后, 极端异常值会被赋予近零权重, 因此在处理含有离散大幅污染的数据时更有优势。在实际应用中, 基于 biweight 的滑动滤波器已被证明是一类非常有效的去趋势工具。Hippke 等人开发的 wotan 是一个面向光变曲线去趋势的 Python 工具包, 系统实现并比较了多种常见方法。其中, biweight 在 Kepler 和 K2 数据上的恢复率表现较好^[32] (见图 3)。不过, 与所有滑动方法一样, 其效果依然高度依赖于窗口大小。窗口必须足够长以保留完整凌星, 同时又足够短以跟踪恒星活动的变化, 经验上通常建议窗口长度约为目标凌星持续时间的三倍^[32]。此外, 该方法对目标类型也较为敏感: 对于安静恒星表现通常较好, 而在活动较强的恒星上, 样条类方法有时更具优势^[62]。在一些更复杂的流程中, biweight 也常作为二级滤波器, 用于进一步清理残余噪声^[74]。

3.1.2 局部回归

局部回归方法通过在每个数据点邻域内拟合一个简单的局部函数来构建平滑基线，因此能够更灵活地适应光变曲线中非线性、时变的趋势。其中最常见的是 **LOWESS**，它是一种非参数的局部加权回归技术，通过在每个数据点的邻域内拟合低阶多项式来完成平滑^[32]。

与 **Savitzky–Golay** 滤波器相比，**LOWESS** 引入了距离加权，离中心点越近的数据权重越大，因此在处理快速演化、非周期性的恒星活动时通常更具优势^[75]。但与所有局部方法一样，其表现高度依赖于窗口尺度的选择：窗口过小会残留恒星噪声，窗口过大则可能抹去真实信号。具体参数通常需要根据数据特性来设定。例如，在处理 **TESS 30 分钟** 采样的年轻恒星时，15 小时（约 30 个数据点）的窗口被证明是较好的折中选择。而在处理 **TRAPPIST-1** 的 **K2** 数据时，0.15 天左右的窗口则能够较有效地去除恒星旋转调制^[75, 76]。

由于 **LOWESS** 对复杂、演化性的恒星活动适应性较强，它在处理年轻、活跃恒星时尤其受到重视。**Battley** 等人（2020）开发的 **YSD**（**Young Stars Detrending**）流程便以 **LOWESS** 为核心，而在针对 **PLATO** 任务模拟数据的算法比较中，**YSD** 及其与 **Huber spline** 的组合也取得了较好的恢复率^[62, 75]（见图 3）。

3.1.3 样条拟合

样条拟合通过分段多项式函数来逼近数据，并在各分段连接点保持一定阶数的连续与光滑，因此能够较灵活地拟合复杂的非线性趋势。为了提高对异常值和凌星信号的鲁棒性，样条方法通常会结合迭代式的 σ 剔除策略，或引入 **Huber**、**biweight** 等鲁棒损失函数^[32]。

目前在天文光变曲线处理中，较常见的鲁棒样条主要包括以下几类。

hspline（**Huber spline**）：**hspline** 在样条拟合中引入 **Huber** 损失函数，使模型在小残差区域保持平滑拟合能力，在大残差区域降低异常点的影响。因此，它既能像普通样条一样描述基线趋势，又能在一定程度上避免耀斑或凌星信号对拟合结果的污染。对于由快速自转等引起的复杂光变，**hspline** 被认为是处理年轻、活跃恒星的一类有效方法^[32]。在 **PLATO** 任务模拟数据的比较中，**hspline** 在地球尺寸小行星的恢复上表现较好，尤其适合长基线数据^[62]（见图 3）。

rspline（**Robust spline**）：**rspline** 主要依赖迭代的 **sigma-clipping** 增强鲁棒性。算法在每轮拟合后自动剔除偏离趋势过大的点，再基于剩余数据重新拟合，直至收敛。已有研究指出，与中位数滤波和 **biweight** 滤波相比，**rspline** 对凌星形状和深度的影响较小，并且在处理光变曲线中存在的数据间断时也能提供更优的趋势拟合效果^[77, 78]。

pspline（**Penalized spline**）：惩罚样条通过在拟合中加入惩罚项自动控制平滑度，从而缓解传统样条需要手动调节节点间距的问题^[78]。这使其在处理光变剧烈的恒星时具有一定优势，并可提高注入行星的回收率。然而，**pspline** 也可能引入周期性残差，特别是在恒星自转调制振幅较高时，迭代裁剪可能误移除旋转峰值，从而在残差中留下可能与凌星混淆的伪影^[79]。

尽管样条方法灵活，但它们也面临一些共有的挑战。所有这类算法都始终存在改变凌星深度甚至完全抹除浅凌星信号的风险，因此需要谨慎设置节点间的最小间距，以避免算

法干扰真实的物理信号^[77, 78]。

3.2 信号分解

信号分解类方法的核心思想是将复杂光变曲线视为目标凌星信号、[恒星光变信号](#)以及仪器噪声等多种成分的叠加, 并通过特定的数学变换将其拆解为在频率、尺度或周期性上具有不同特征的基础分量。去趋势的关键便是在分解后的空间中识别并抑制与噪声相关的成分, 同时尽可能保留短时标凌星事件的关键特征。常见方法主要包括小波变换与余弦滤波。

3.2.1 小波变换

小波变换能够同时在时间和频率两个维度上分析信号, 因此在处理非平稳、具有瞬态特征或多尺度相关噪声的光变曲线时, 相较于传统傅里叶方法具有优势。

在去趋势应用中, 一个代表性例子是 Grziwa 等人 (2016) 开发的 VARLET 小波滤波器^[80]。作为一种无需先验模型的去噪方法, VARLET 能够有效移除恒星变化和仪器噪声, 从而凸显微弱凌星信号, 尤其适用于存在间断或非等距采样的数据。然而, 该方法也会系统性削弱凌星深度, 因此更适合作为探测工具, 而不适用于后续精确参数表征。尽管如此, 比较研究表明, 在一些困难场景下, 例如行星轨道周期与恒星自转周期接近时, VARLET 对小行星的探测成功率在该比较中最高^[62] (见图 3)。

3.2.2 余弦滤波

余弦滤波是一类频域方法, 其基本思想是利用一组不同频率的余弦或正弦函数来拟合并移除光变曲线中的周期性或准周期性趋势, 例如仪器热漂移、姿态变化以及恒星自转或脉动引起的变化。该方法的关键在于如何确定谐波数量: 若模型过于简单, 则噪声残留较多。若模型过于复杂, 则容易过拟合并削弱真实凌星信号。

现有方法大致可分为固定参数法和自适应优化法。前者通过人为设定谐波数量或频率上限实现滤波, 但参数选择不当时容易引入吉布斯现象、削弱凌星信号, 甚至产生较多假阳性^[81-83]。后者则依据残差特性自动调节滤波强度, 例如 CoFiAM (Cosine Filtering with Autocorrelation Minimization) 通过最小化残差自相关来选择最佳谐波数^[65]。类似地, STANLEY 会根据残差周期图是否仍存在低频结构来逐步增加谐波数量^[74]。这类方法在保留信号和抑制噪声之间提供了更灵活的平衡, 但也对参数和流程设置较为敏感, 有时仍可能产生伪信号^[84, 85]。此外, Kepler 官方管线中的谐波移除也可视为自适应思想的一种应用, 但其角色更具针对性。它并非通用的去趋势工具, 而是一种预处理步骤, 主要用于迭代识别并剔除由恒星脉动等物理原因产生的离散高信噪比周期信号, 以便后续通用去趋势方法在更干净的数据上工作^[51]。

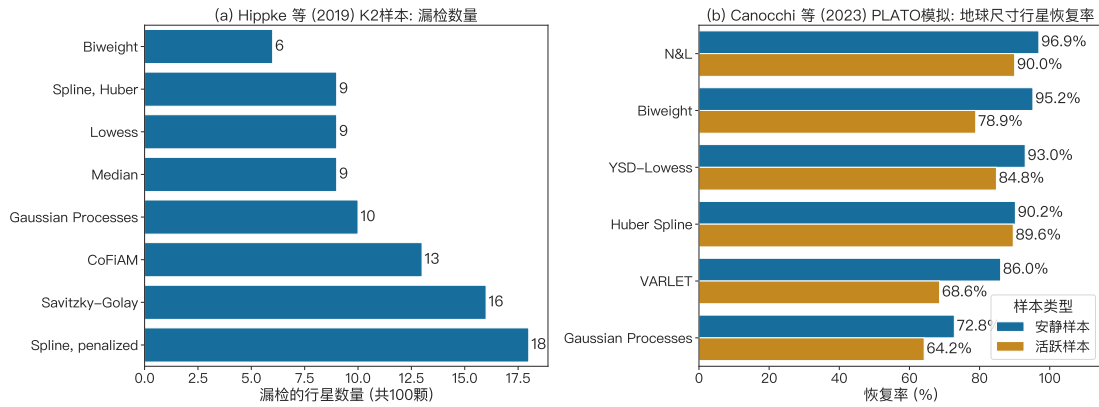
3.3 概率建模

在处理复杂、非平稳的时间相关噪声时, 以高斯过程 (Gaussian Processes, GP) 为代表的概率建模方法提供了灵活的非参数化框架^[86, 87]。其核心在于通过协方差函数刻画噪声的相关结构, 并将关于物理过程的先验知识融入模型。例如, 平方指数核适合描述平滑变化, 准周期核常用于表征由星斑引起的准周期活动, 而 SHO (stochastically-driven damped harmonic

oscillator) 核则可用于建模米粒组织等现象^[46, 88]。借助 *celerite* 等现代计算框架, GP 也能够较高效地处理大规模光变曲线。

GP 的灵活性使其在实际应用中取得了一些成功。典型例子是 K2SC 流程, 它通过多输入 GP 同时建模仪器系统误差和天体物理变化, 并已被用于 K2 数据中的候选体发现^[87, 89]。但与此同时, GP 在先去趋势后搜索范式下也存在明显风险: 由于模型足够灵活, 它可能将真实的凌星信号误吸收到噪声模型中, 从而削弱信号。为缓解这一问题, 许多流程会结合异常点剔除或对可拟合时间尺度进行限制, 以在去除噪声的同时保护潜在凌星事件^[32, 87, 90]。

尽管如此, GP 在该范式下的整体表现仍存在争议。一些基准测试指出, 标准 GP 去趋势后的行星恢复率并不突出, 且可能系统性影响参数测量^[32]。进一步的比较研究也表明, 在处理活跃恒星或探测地球尺寸小行星时, GP 的表现并不总是理想^[62, 82] (见图 3)。因此, GP 更适合在具有明确噪声结构假设、且配合适当保护机制的条件下使用。



注: (a) 算法在 K2 任务数据中的表现, 数据改编自 Hippke 等人 (2019) 的图 5。该指标为在 100 个包含强仪器噪声的 K2 光变曲线样本中, 算法未能检测到的行星数量。(数值越低代表算法在强噪声下的鲁棒性越好。)(b) 算法在 PLATO 任务模拟中的表现, 数据源自 Canocchi 等人 (2023)。该指标展示了算法对地球尺寸行星的信号恢复率。(数值越高代表算法对微弱信号的保留能力越强。)

图 3 不同去趋势方法在基准测试中的性能对比^[32, 62]

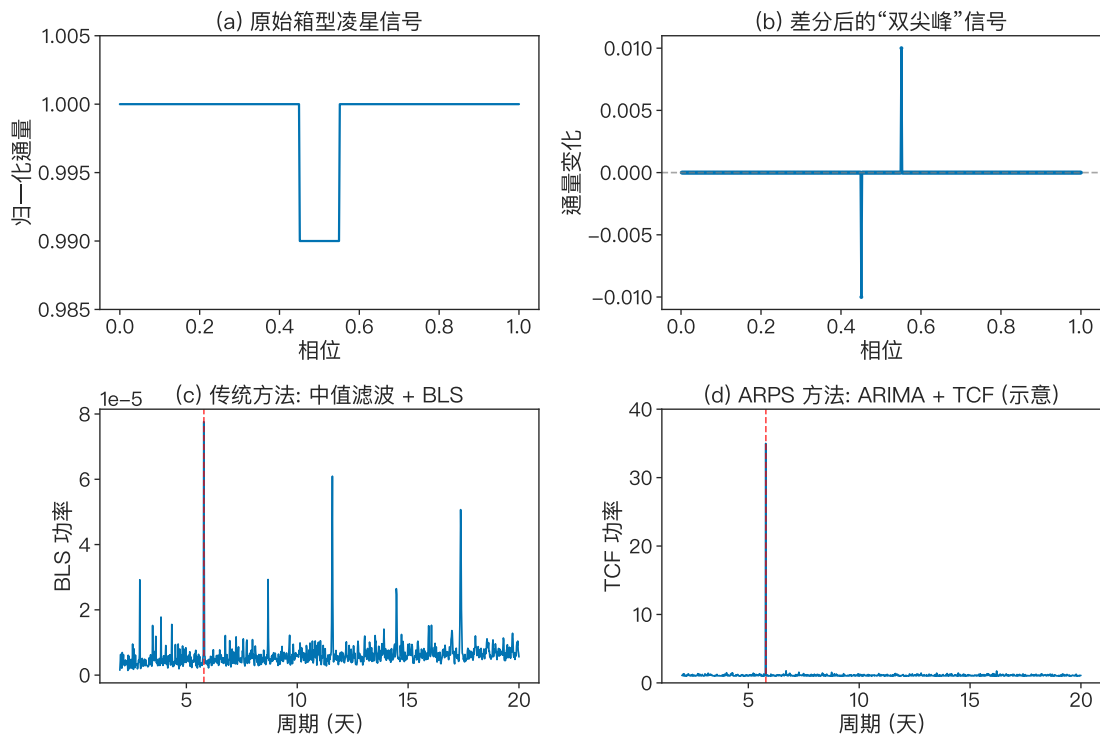
3.4 时序建模

自回归积分移动平均 (autoregressive integrated moving average, ARIMA) 模型是一类经典的参数化时间序列方法, 可用于刻画序列的自相关结构、长期趋势和随机扰动^[91, 92]。在系外行星探测中, Caceres 等人将其引入为自回归行星搜索 (AutoRegressive Planet Search, ARPS) 框架, 用以建模并移除恒星光变信号中的自相关噪声, 从而得到更接近白噪声的残差序列, 并凸显其中的微弱周期性凌星信号。

ARIMA 特别适用于存在强自相关噪声的光变曲线, 例如受星斑活动或耀斑影响较大的目标^[91, 92]。该方法的一个特点是差分操作会将传统箱型凌星信号转化为双尖峰结构, 因此通常需要与专门设计的 TCF (Transit Comb Filter) 周期图配合使用。相关研究表明, 在自相关

噪声显著场景下, ARIMA+TCF 组合在某些测试中对微弱凌星信号的探测灵敏度优于“高斯过程+BLS”等传统流程^[93]。

不过, ARIMA 也存在明显局限。首先, 其标准形式通常要求数据近似等间距, 这限制了它在不规则采样地面数据中的直接应用^[91, 92]。其次, 由于 TCF 不显式依赖凌星时长, ARPS 对长周期、长持续时间凌星的灵敏度下降快于 BLS^[91, 92]。此外, 作为线性模型, ARIMA 对某些非线性、剧烈变化的恒星行为拟合能力有限。尽管如此, ARPS 项目的实际应用表明, 该方法在处理大规模 Kepler 数据时仍具有实际价值, 并成功发现了一批此前未被识别的浅凌星候选体^[91, 92]。



注: 本图展示了 ARPS 方法的核心原理及其优势。上排 (a, b): 展示了 ARIMA 流程中的差分操作如何将一个传统的箱型凌星信号 (a) 转变为一个独特的双尖峰信号 (b), 分别对应凌星的进入和离开。下排 (c, d): 对比了两种方法的周期图探测效果。传统方法 (c, 中值滤波 +BLS) 虽然能探测到信号 (红色虚线), 但背景噪声和伪峰较多。相比之下, ARPS 流程 (d, ARIMA+TCF) 通过匹配新信号形态, 能产生一个信噪比较高的探测峰值和非常平坦的噪声基底。

图 4 ARPS 方法原理示意: 差分后的双尖峰信号与 TCF 周期图^[91, 92]

3.5 机器学习

机器学习在光变曲线去趋势中的应用仍处于探索阶段, 目前主要扮演两类角色。其一是作为传统去趋势方法的辅助工具, 用于自动优化滤波窗口、正则化强度等超参数, 从而减少人工调参, 提高处理流程的自动化程度^[32]。其二是尝试以深度学习模型直接替代传统

算法,实现端到端的数据驱动去趋势。例如, Morvan 等人分别探索了基于 LSTM^[94] 和去噪 Transformer^[64] 的方法来学习光变曲线基线特征。

尽管这类方法在处理复杂时序模式方面具有潜力,但其在真实天文数据上的稳定性、泛化性以及微弱凌星信号的保真度仍有待进一步验证。特别是在恒星活动复杂、仪器噪声多样的情况下,如何避免模型在去噪过程中同时抑制真实信号,仍是当前研究中的关键问题。因此,机器学习目前更适合作为有前景的新方向,而非已经成熟的通用解决方案。

4 联合建模

传统的先去趋势,后搜索策略虽然直观,但面临一个核心挑战:去趋势过程可能会因过度拟合而削弱甚至消除真实的凌星信号。为解决这一问题,研究者逐渐发展出一种思路,即在信号搜索的同时对噪声进行建模,以期提升探测的灵敏度和鲁棒性。为便于讨论,本文将这类方法统称为“联合建模”(joint modeling)范式。其实现路径主要可分为两类:将噪声建模与周期搜索深度整合的一体化框架,以及借鉴其思想但仍分步执行的混合式框架。

4.1 端到端搜索

这类方法将噪声建模与周期搜索整合在一个统一的算法体系内,实现了从输入数据到候选体列表的端到端处理。

(1) 线性-周期两步法

为避免在完整的(周期、历元、时长)三维参数空间中进行穷举式盲搜所带来的巨大计算成本,该路径将问题分解为线性搜索和周期性搜索两个高效步骤。

Foreman-Mackey 等人 (2015)^[95] 的开创性工作为此路径奠定了基础。他们提出,由于凌星信号与系统误差(噪声)永远不会完全正交,任何独立的去趋势操作都会不可避免地扭曲信号。为此,他们开发了一个同时拟合系统误差模型和凌星模型的框架,并证明了其在理论上的优越性。然而, Kovács 等人 (2016) 很快提出了反论,他们认为在初始探测阶段,联合建模过高的模型自由度反而更容易将噪声误认为信号,从而降低真实信号的信噪比。其基于 TFA (trend filtering algorithm) 算法的实验表明,传统的序贯方法(先去趋势,后搜索)在信号恢复率上反而更高。这场辩论揭示了联合建模在实践中的复杂性,其成败高度依赖于噪声模型的准确性。

Garcia 等人 (2024)^[82] 开发的 Nuance 算法可被视为这一问题的一个较新实现。它采用计算成本较高但极为灵活的高斯过程作为噪声模型,在联合建模框架下对活跃恒星进行搜索。在大规模的注入-恢复实验中, Nuance 的表现在 93% 的案例中都优于所有参与对比的传统序贯方法,成功恢复了 76% 的可探测信号。这表明,在噪声模型较为准确且计算资源允许的前提下,联合建模在实际应用中可以表现出相对优势。

(2) 整体式贝叶斯假设检验

与分解式的两步法不同, Taaki 等人 (2020, 2025)^[96, 97] 提出的整体式贝叶斯探测,是在一个统一的贝叶斯框架下,对一个完整的、周期性的信号模型进行一次性的整体假设检验。

该方法在统计学上更为严谨, 理论上能在任何给定的假警报率下最大化探测概率。然而, 这种严谨性也带来了极端的计算成本(需要千万亿次级的超级计算机)和巨大的数据处理量, 使其成为一个高性能计算挑战。此外, 测试表明其对特定的、未被完全建模的仪器误差较为敏感。

4.2 局部联合拟合

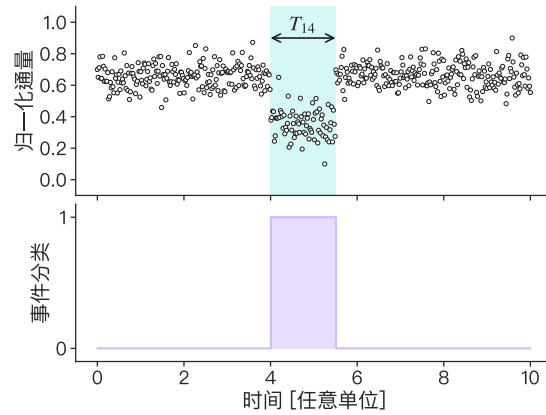
这类方法并未构建完全一体化的搜索框架, 但它们借鉴了联合建模的同时拟合信号与噪声的核心思想, 以更智能的方式进行去趋势。一个典型的例子是 Rizzuto 等人 (2017)^[63] 提出的 **notch filter**。在处理快速旋转的年轻恒星时, 该方法并非先抹平整个基线, 而是在一个滑动窗口内, 同时用多项式拟合一个局部的噪声模型和一个代表凌星的凹槽模型。这使得它能在移除复杂恒星活动的同时, 有助于减少凌星信号在去趋势过程中被削弱的程度^[63]。

5 直接凌星探测

传统分析流程通常包含一个独立的去趋势步骤, 例如使用中值滤波、样条拟合或高斯过程来滤除仪器噪声和恒星活动信号。然而, 这类预处理可能在一定程度上削弱甚至移除真实的凌星信号。为应对此问题, 研究者提出了直接凌星探测方法, 将信号识别与噪声处理置于同一框架内, 由模型从原始含噪数据中学习凌星相关特征。在此背景下, 深度学习方法主要沿着两条技术路径发展: 早期的分类任务, 以及近年来提出的分割任务。

最初, 研究者们将凌星探测构建为一个分类任务, 其核心目标是判断一段完整的光变曲线是否包含凌星信号。这些模型的训练数据均没有经过传统意义上的去趋势处理, 而是直接包含了模拟的恒星活动和噪声。Zucker 和 Giryes (2018)^[98] 的开创性研究证明, 一个卷积神经网络 (CNN) 能够直接从含有高斯过程噪声的原始模拟数据中学习, 在固定的 0.01 误报率下实现了 0.94 的真实正例率。同样, Pearson 等人^[99] (2018) 的 CNN 模型也在未去趋势的模拟数据上进行训练, 在一个独立的测试集上达到了 88.45% 的总体准确率。近期, Malik 等人 (2025) 更是通过利用公民科学家的标记数据进行训练, 使得其分类模型在恢复已知行星候选体 (TOIs) 方面的性能比人类志愿者还高出 10%。然而, 纯粹的分类方法存在一个固有局限性: 它仅回答了有没有凌星的问题, 却未告知在哪里。这在实际应用中可能导致问题: 由于模型对一段或整条曲线给出一个宏观判断, 它无法排除由噪声引起的局部假阳性特征, 如果敏感度设置过高, 可能会导致大量不含真实信号的光变曲线也被标记, 使得大规模检测的筛选效果大打折扣。

为了克服这一局限性, 近年来提出的研究将凌星探测定义为一个分割任务, 其目标不仅是判断, 更是定位。Vivien 等人 (2025)^[100] 提出的 PANOPTICON 模型将凌星探测任务重构为基于一维 U-Net3+ 架构^[101] (该架构源于经典的 U-Net 分割网络^[102]) 的逐点分类问题, 旨在直接从未经去趋势的光变曲线中定位单次凌星事件。该模型利用编码器的迭代下采样操作提取上下文信息, 使卷积核能够覆盖更大的时间窗口以评估多尺度数据, 并通过全尺度跳跃连接将解码器上采样的输出与不同层级提取的特征相结合, 从而生成高分辨率的分割



注：上方面板展示了作为模型输入的归一化光变曲线。图中包含了模拟的噪声，淡蓝色的背景区域高亮了持续时间为 T_{14} 的行星凌星事件。下方面板展示了模型的理想输出——一个事件分类图。模型成功地在对应的时间段内输出高信号（值为 1），并在其他基线部分输出低信号（值为 0），从而准确识别出凌星事件的位置与持续时间。

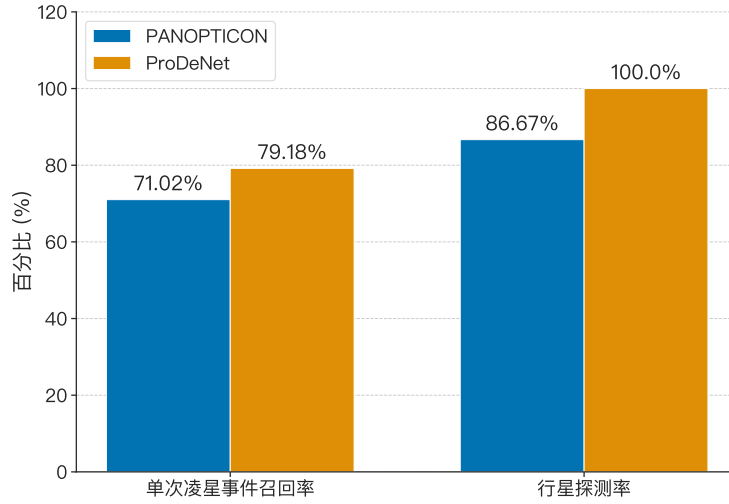
图 5 PANOPTICON 模型的理论输入与输出示意图^[100]

割图。Vivien 等人^[100]认为这种架构设计使得深度学习模型能够构建一个内部噪声模型，在不需要预先去趋势的情况下，通过内部学习辨别不同类型的信号差异。在针对 PLATO 任务的模拟数据分析中，这种端到端方法取得了较好的结果：在未经去趋势的光变曲线中，模型以 1% 的虚警率（FAR）实现了 90% 的总体恢复率，即便对于信噪比极低的地球大小行星，其恢复率也超过了 25%。此外，测试结果显示该模型的恢复率与恒星活动指标之间不存在强相关性，证明了其在强噪声下的鲁棒性。同时其推理速度极快（0.2 秒/条），在推理速度方面具有优势，也减少了对人工调参与监督流程的依赖。尽管目前的模型仍无法仅凭光度数据有效区分行星与食双星，因此其主要仍作为高灵敏度事件探测器运行，但未来通过将质心数据纳入网络输入，有望实现对这两类天体物理信号的分类^[100]。

尽管 PANOPTICON (U-Net3+) 利用全尺度跳跃连接优化了特征融合，但其编码器固有的固定尺度卷积核限制了对复杂光变曲线的解析能力。在处理叠加于长周期恒星活动之上的短时凌星信号时，U-Net 类架构常面临下采样带来的分辨率损失与浅层特征视野受限之间的矛盾。针对这一尺度失配问题，我们构建了一个基于 DeepLabv3+ 架构^[103]的模型，记为 ProDeNet。

不同于 U-Net3+ 主要依赖复杂跳跃连接恢复信息的路径，ProDeNet 以空洞空间金字塔池化（ASPP）为核心模块，通过多个膨胀率（ $r = \{6, 12, 18\}$ ）的空洞卷积并行构建多尺度感受野。较小膨胀率更适于表征入凌和出凌阶段的局部高频特征，而较大膨胀率能够覆盖更长时间尺度，用于描述恒星自转等低频变化。因此，该模型能够在保持时间分辨率的同时兼顾局部细节与全局趋势。

为了验证上述设计的有效性，我们构建了一个高难度测试子集。该数据集包含从 NASA 系外行星档案馆精选的 30 条 Kepler 已确认行星宿主恒星的光变曲线，所有样本均来自已确



注: 评价指标包括: (1) 单次事件召回率, 即被模型正确识别的真实凌星事件数占所有可观测真实凌星事件总数的比例。(2) 行星探测率, 即在单条光变曲线中成功恢复至少 3 次真实凌星事件的样本比例。相较 PANOPTICON, ProDeNet 的单次事件召回率和行星探测率分别提升了 8 个百分点和 13 个百分点。

图 6 PANOPTICON 与 ProDeNet 在测试集上的性能对比。

认行星系统, 且每条光变曲线均包含至少 3 次已知真实凌星事件。这些样本的观测基线均超过 100 天, 单次凌星事件信噪比最低可至 3.9, 旨在模拟长周期、低信噪比条件下的困难探测场景。真实值由 NASA Exoplanet Archive 已发布的行星参数 (t_0 、周期 P 和持续时间 T_{14}) 确定。若模型检测到的事件中心时间与真实凌星中心时间的偏差不超过 $\max(T_{14}/2, \Delta t_{\text{cadence}})$, 则判定为一次正确匹配, 其中 $\Delta t_{\text{cadence}}$ 为该光变曲线的中值采样间隔。单次事件召回率定义为被正确匹配的真实凌星事件数占所有可观测真实凌星事件总数的比例。行星探测率定义为成功恢复至少 3 次真实凌星事件的样本占全部 30 条光变曲线的比例。

对比测试结果 (如图 6 所示) 表明, ProDeNet 在该测试集中表现出更优的微弱信号识别能力: 其单次事件召回率较 PANOPTICON 提升了 8 个百分点, 而行星探测率提升了 13 个百分点。虽然测试样本量过小, 但是这一初步结果表明, 多尺度感知机制在处理低信噪比、长周期时域数据时具有潜力, 值得进一步在更大规模数据集上验证。

6 凌星参数精测去趋势: 凌星发现后的精细调整

在系外行星研究中, 凌星信号的发现只是第一步, 其后的精确表征对于测量行星半径、轨道参数、TTVs 以及大气特征等至关重要。因此, 这一阶段的目标已不再是“发现信号”, 而是尽可能准确地恢复已知凌星的真实形状与物理参数。

常规去趋势方法在去除恒星活动和仪器噪声时, 往往会不同程度地扭曲凌星的深度与

形状。为此,精测阶段通常利用已知的凌星信息(如历元和持续时间)在去趋势过程中主动保护信号完整性,以减少信号失真和深度变浅,从而提高参数测量的精度与可靠性。

最基础且应用最广的策略是凌星掩膜,即在拟合趋势模型时将凌星区间的数据排除,仅利用凌星外的数据估计背景趋势与**恒星光变信号**。围绕这一思想,不同研究采用了高斯过程^[104–107]、滑动滤波器^[107–110]、样条^[111]、CoFiAM^[65, 84, 108, 112, 113]以及集成学习^[69]等不同趋势建模器。尽管实现方式不同,其共同核心都是通过掩膜避免趋势模型削弱真实凌星信号。

相比之下,更完整的思路是同时对凌星信号与背景噪声进行联合建模。相关研究指出,在信号发现之后,最终参数拟合应采用完整模型,以避免分步处理中可能出现的信号挤压(signal squeezing)效应^[114]。其中,基于高斯过程的方法是一类常见实现:它以凌星模型作为均值函数,以协方差函数描述仪器系统误差和恒星活动等相关噪声,从而在统一框架下实现高精度参数推断^[115, 116]。

此外,一些新方法也被用于精测阶段的基线恢复。例如,一种思路是将深度学习模型应用于基线重建。尽管 Morvan 等人提出的 LSTM^[94]和去噪 Transformer^[64]原本旨在实现通用去趋势,但在凌星参数精测阶段,这类模型可被进一步用于一个新的任务:在掩膜已知凌星区间后,利用剩余数据学习恒星活动的非线性特征,并对掩膜区域进行基线重建。对于多波段观测,还可以利用恒星活动具有颜色依赖而凌星信号近似消色的性质,在时间和波长两个维度上联合建模噪声,从而进一步提高凌星参数测量精度^[117]。

7 总 结

从高精度光变曲线数据中可靠地提取微弱的系外行星凌星信号,是现代天文学面临的核心挑战之一。本综述系统性地梳理了为应对这一挑战而发展的各类去趋势技术。这些技术的核心目标是在有效移除由仪器效应和**恒星光变信号**引起的复杂噪声的同时,尽可能保留真实的凌星信号,从而为行星的发现和精确表征奠定基础。

我们的探讨揭示了去趋势方法论主要沿着三大范式演进。第一种是应用最广泛的先去趋势后搜索范式,其优势在于流程清晰、计算高效。在该范式下,从经典的滑动滤波器、样条拟合到现代的概率建模,各类技术都得到了深入发展。多项基准测试表明,基于鲁棒统计的滑动滤波器在盲搜任务中表现出较高的行星恢复率,而 Savitzky-Golay 滤波器和标准高斯过程则因其固有的过拟合风险,可能系统性地削弱甚至完全移除凌星信号,需谨慎使用。对于具有强自相关噪声的活跃恒星,ARIMA 时序建模则提供了一条独特的、通过白化噪声来凸显信号的有效路径。第二种是理论上更一致的联合建模范式。它通过在一个统一框架内同步拟合噪声与凌星信号,从方法上减少了先去趋势方法可能引入的信号失真问题。尽管计算成本高昂,但以 Nuance 算法为代表的现代实现表明,在噪声模型较为准确的前提下,该范式在处理传统方法难以应对的活跃恒星时往往更有优势。第三种是近年来兴起的、由数据驱动的直接凌星探测范式。以深度学习为核心,该方法不再依赖传统的独立去趋势步骤,直接从原始含噪数据中学习并识别凌星特征。其技术路径已从早期的分类任务(判断

有无凌星) 演进到近年来提出的分割任务(定位何处凌星), 在抑制假阳性的同时, 在部分测试中表现出与传统方法相竞争、甚至更优的探测性能。

综上所述, 不存在任何一种万能的去趋势算法。最佳方法的选择高度依赖于具体的科学目标、恒星的物理特性以及数据的质量。在行星被发现后, 去趋势的目标从探测转向精确, 此时采用凌星掩膜或同步拟合等策略来主动保护信号完整性, 是获取高精度物理参数的必要步骤。

展望未来, 该领域的发展可能集中在以下几个方向。首先, 随着计算能力的提升, 联合建模与直接凌星探测方法有望在未来的大规模巡天任务中发挥更重要的作用。初步测试结果表明, 引入多尺度感知机制可能是提升这类方法性能的一条可行路径。其次, 混合式策略, 例如利用机器学习自动为每条光变曲线选择更合适的去趋势算法及其超参数, 可能有助于提高处理流程的自动化程度和效率。最后, 利用恒星活动与凌星信号在不同波段上的差异, 发展多波段联合建模等技术, 有望为从物理层面分离天体物理噪声提供新的思路。总体而言, 开发和完善更智能、更鲁棒、更高效的去趋势方法, 仍将是未来从海量数据中发掘类地行星乃至地球 2.0 的重要方向。

参考文献:

- [1] Briot D, Schneider J. Handbook of Exoplanets. 2018: 35
- [2] Borucki W J, Summers A L. Icarus, 1984, 58: 121
- [3] Henry G W, Marcy G W, Butler R P, et al. The Astrophysical Journal, 2000, 529: L41
- [4] Charbonneau D, Brown T M, Latham D W, et al. The Astrophysical Journal, 2000, 529: L45
- [5] Walker G, Matthews J, Kuschnig R, et al. Publications of the Astronomical Society of the Pacific, 2003, 115: 1023
- [6] Auvergne M, Bodin P, Boisnard L, et al. Astronomy & Astrophysics, 2009, 506: 411
- [7] Borucki W J, Koch D, Basri G, et al. Science, 2010, 327: 977
- [8] Ricker G R, Winn J N, Vanderspek R, et al. Journal of Astronomical Telescopes, Instruments, and Systems, 2015, 1: 014003
- [9] Benz W, Broeg C, Fortier A, et al. Experimental Astronomy, 2021, 51: 109
- [10] Léger A, Rouan D, Schneider J, et al. Astronomy & Astrophysics, 2009, 506: 287
- [11] Deleuil M, Aigrain S, Moutou C, et al. Astronomy & Astrophysics, 2018, 619: A97
- [12] Batalha N M. Proceedings of the National Academy of Sciences, 2014, 111: 12647
- [13] Quintana E V, Barclay T, Raymond S N, et al. Science, 2014, 344: 277
- [14] Gandolfi D, Barragán O, Livingston J H, et al. Astronomy & Astrophysics, 2018, 619: L10
- [15] Gilbert E A, Barclay T, Schlieder J E, et al. The Astronomical Journal, 2020, 160: 116
- [16] Lendl M, Csizmadia S, Deline A, et al. Astronomy & Astrophysics, 2020, 643: A94
- [17] Fortier A, Simon A E, Broeg C, et al. Astronomy & Astrophysics, 2024, 687: A302
- [18] Rauer H, Catala C, Aerts C, et al. Experimental Astronomy, 2014, 38: 249
- [19] Rauer H, Aerts C, Cabrera J, et al. Experimental Astronomy, 2025, 59: 26
- [20] Johnson S A, Penny M, Gaudi B S, et al. The Astronomical Journal, 2020, 160: 123
- [21] Wilson R F, Barclay T, Powell B P, et al. The Astrophysical Journal Supplement Series, 2023, 269: 5
- [22] Ge J, Zhang H, Zang W, et al. arXiv preprint arXiv:220606693, 2022
- [23] Ge J, Zhang H, Deng H, et al. The Innovation, 2022, 3: 100271
- [24] Ge J, Zhang H, Zhang Y, et al. Space Telescopes and Instrumentation 2024: Optical, Infrared, and Millimeter Wave: Vol. 13092. SPIE, 2024: 393
- [25] Ge J, Chen W, Chen Y, et al. Chinese Journal of Space Science, 2024, 44: 400
- [26] Kim D W, Protopapas P, Alcock C, et al. Monthly Notices of the Royal Astronomical Society, 2009, 397: 558

- [27] Pont F, Aigrain S, Bakos G, et al. *Astronomical Society of the Pacific Conference Series: Vol. 366 Transiting Extrasolar Planets Workshop*. Astronomical Society of the Pacific, 2006: 3
- [28] Koch D G, Borucki W J, Basri G, et al. *The Astrophysical Journal*, 2010, 713: L79
- [29] Smith J C, Stumpe M C, Van Cleve J E, et al. *Publications of the Astronomical Society of the Pacific*, 2012, 124: 1000
- [30] Stumpe M C, Smith J C, Van Cleve J E, et al. *Publications of the Astronomical Society of the Pacific*, 2012, 124: 985
- [31] Smith J C, Stumpe M C, Jenkins J M, et al. No. KSCI-19081-003, NASA Ames Research Center, 2020: 8
- [32] Hippke M, David T J, Mulders G D, et al. *The Astronomical Journal*, 2019, 158: 143
- [33] Gilliland R L, Chaplin W J, Dunham E W, et al. *The Astrophysical Journal Supplement Series*, 2011, 197: 6
- [34] Van Cleve J E, Caldwell D A. No. KSCI-19033-002, NASA Ames Research Center, 2016: 1
- [35] Jenkins J M, Caldwell D A, Chandrasekaran H, et al. *The Astrophysical Journal*, 2010, 713: L120
- [36] Van Cleve J E, Christiansen J L, Jenkins J M, et al. No. KSCI-19040-005, 2016: 2
- [37] McQuillan A, Aigrain S, Roberts S. *Astronomy & Astrophysics*, 2012, 539: A137
- [38] Aerts C, Christensen-Dalsgaard J, Kurtz D W. *Asteroseismology*. 2010
- [39] Basu S, Chaplin W J. *Asteroseismic Data Analysis. Foundations and Techniques*. 2018
- [40] Kurtz D W. *Annual Review of Astronomy and Astrophysics*, 2022, 60: 31
- [41] Clementini G, Ripepi V, Garofalo A, et al. *Astronomy & Astrophysics*, 2023, 674: A18
- [42] Hey D R, Montet B T, Pope B J S, et al. *The Astronomical Journal*, 2021, 162: 204
- [43] Li G, Van Reeth T, Bedding T R, et al. *Monthly Notices of the Royal Astronomical Society*, 2020, 491: 3586
- [44] Oshagh M. *Handbook of Exoplanets*. 2018: 239
- [45] Costes J C, Watson C A, de Mooij E, et al. *Monthly Notices of the Royal Astronomical Society*, 2021, 505: 830
- [46] Angus R, Morton T, Aigrain S, et al. *Monthly Notices of the Royal Astronomical Society*, 2018, 474: 2094
- [47] Moffett T J. *The Astrophysical Journal Supplement Series*, 1974, 29: 1
- [48] Davenport J R A, Hawley S L, Hebb L, et al. *The Astrophysical Journal*, 2014, 797: 122
- [49] Santerne A, Díaz R F, Moutou C, et al. *Astronomy & Astrophysics*, 2012, 545: A76
- [50] Fressin F, Torres G, Charbonneau D, et al. *The Astrophysical Journal*, 2013, 766: 81
- [51] Thompson S E, Coughlin J L, Hoffman K, et al. *The Astrophysical Journal Supplement Series*, 2018, 235: 38
- [52] Grunblatt S K, Huber D, Gaidos E, et al. *The Astronomical Journal*, 2017, 154: 254
- [53] Pereira F, Campante T L, Cunha M S, et al. *Monthly Notices of the Royal Astronomical Society*, 2019, 489: 5764
- [54] Chiavassa A, Bigot L. *European Astronomical Society Publications Series*, 2014, 70: 151
- [55] Chiavassa A, Caldas A, Selsis F, et al. *Astronomy & Astrophysics*, 2017, 597: A94
- [56] Morris B M, Bobra M G, Agol E, et al. *Monthly Notices of the Royal Astronomical Society*, 2020, 493: 5489
- [57] Christensen-Dalsgaard J. *Reviews of Modern Physics*, 2002, 74: 1073
- [58] Kovács G, Zucker S, Mazeh T. *Astronomy & Astrophysics*, 2002, 391: 369
- [59] Hippke M, Heller R. *Astrophysics Source Code Library*, 2019: ascl:1910.007
- [60] Stefanov A K, González Hernández J I, Suárez Mascareño A, et al. *Astronomy & Astrophysics*, 2025, 693: L3
- [61] Handberg R, Lund M N. *Monthly Notices of the Royal Astronomical Society*, 2014, 445: 2698
- [62] Canocchi G, Malavolta L, Pagano I, et al. *Astronomy & Astrophysics*, 2023, 672: A144
- [63] Rizzuto A C, Mann A W, Vanderburg A, et al. *The Astronomical Journal*, 2017, 154: 224
- [64] Morvan M, Nikolaou N, Yip K H, et al. *arXiv preprint arXiv:220702777*, 2022
- [65] Kipping D M, Hartman J, Buchhave L A, et al. *The Astrophysical Journal*, 2013, 770: 101
- [66] Aigrain S, Hodgkin S T, Irwin M J, et al. *Monthly Notices of the Royal Astronomical Society*, 2015, 447: 2880
- [67] Hippke M. *The Astrophysical Journal*, 2015, 806: 51
- [68] Wittenmyer R A, Clark J T, Trifonov T, et al. *The Astronomical Journal*, 2022, 163: 82
- [69] Yahalomi D A, Kipping D, Solano-Oropeza D, et al. *The Astrophysical Journal Supplement Series*, 2026, 283: 51
- [70] Dévora-Pajares M, Pozuelos F J, Thuillier A, et al. *Monthly Notices of the Royal Astronomical Society*, 2024, 532: 4752
- [71] Delrez L, Murray C A, Pozuelos F J, et al. *Astronomy & Astrophysics*, 2022, 667: A59
- [72] Dransfield G, Triaud A H M J, Guillot T, et al. *Monthly Notices of the Royal Astronomical Society*, 2022, 515: 1328
- [73] Huber P J. *The Annals of Mathematical Statistics*, 1964, 35: 73
- [74] Martin D V, Fabrycky D C. *The Astronomical Journal*, 2021, 162: 84

- [75] Battley M P, Pollacco D, Armstrong D J. *Monthly Notices of the Royal Astronomical Society*, 2020, 496: 1197
- [76] Luger R, Sestovic M, Kruse E, et al. *Nature Astronomy*, 2017, 1: 0129
- [77] Heidari N, Boisse I, Orell-Miquel J, et al. *Astronomy & Astrophysics*, 2022, 658: A176
- [78] Giacalone S, Dressing C D. *The Astronomical Journal*, 2025, 169: 45
- [79] Fernandes R B, Mulders G D, Pascucci I, et al. *The Astronomical Journal*, 2022, 164: 78
- [80] Grziwa S, Pätzold M. *arXiv preprint arXiv:160708417*, 2016
- [81] Bonomo A S, Aigrain S, Bordé P, et al. *Astronomy & Astrophysics*, 2009, 495: 647
- [82] Garcia L J, Foreman-Mackey D, Murray C A, et al. *The Astronomical Journal*, 2024, 167: 284
- [83] Mazeh T, Faigler S. *Astronomy & Astrophysics*, 2010, 521: L59
- [84] Rodenbeck K, Heller R, Hippke M, et al. *Astronomy & Astrophysics*, 2018, 617: A49
- [85] Castro-González A, Díez Alonso E, Menéndez Blanco J, et al. *Monthly Notices of the Royal Astronomical Society*, 2022, 509: 1075
- [86] Danielski C, Kacprzak T, Tinetti G. *European Planetary Science Congress: Vol. 8. 2013*
- [87] Aigrain S, Parviainen H, Pope B J S. *Monthly Notices of the Royal Astronomical Society*, 2016, 459: 2408
- [88] Rasmussen C E, Williams C K I. *Gaussian Processes for Machine Learning*. The MIT Press, 2005
- [89] Pope B J S, Parviainen H, Aigrain S. *Monthly Notices of the Royal Astronomical Society*, 2016, 461: 3399
- [90] Zink J K, Hardegree-Ullman K K, Christiansen J L, et al. *The Astronomical Journal*, 2020, 159: 154
- [91] Caceres G A, Feigelson E D, Jogesh Babu G, et al. *The Astronomical Journal*, 2019, 158: 57
- [92] Caceres G A, Feigelson E D, Jogesh Babu G, et al. *The Astronomical Journal*, 2019, 158: 58
- [93] Gondhalekar Y, Feigelson E D, Caceres G A, et al. *The Astrophysical Journal*, 2023, 959: L16
- [94] Morvan M, Nikolaou N, Tsiaras A, et al. *The Astronomical Journal*, 2020, 159: 109
- [95] Foreman-Mackey D, Montet B T, Hogg D W, et al. *The Astrophysical Journal*, 2015, 806: 215
- [96] Taaki J S, Kamalabadi F, Kembell A J. *The Astronomical Journal*, 2020, 159: 283
- [97] Taaki J S, Kembell A J, Kamalabadi F. *The Astronomical Journal*, 2025, 170: 14
- [98] Zucker S, Giryes R. *The Astronomical Journal*, 2018, 155: 147
- [99] Pearson K A, Palafox L, Griffith C A. *Monthly Notices of the Royal Astronomical Society*, 2018, 474: 478
- [100] Vivien H G, Deleuil M, Jannsen N, et al. *Astronomy & Astrophysics*, 2025, 694: A293
- [101] Huang H, Lin L, Tong R, et al. *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2020: 1055
- [102] Ronneberger O, Fischer P, Brox T, Navab N, Hornegger J, Wells W M, et al., eds. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Cham: Springer International Publishing, 2015: 234
- [103] Chen L C, Zhu Y, Papandreou G, et al. Ferrari V, Hebert M, Sminchisescu C, et al., eds. *Computer Vision – ECCV 2018*. Cham: Springer International Publishing, 2018: 833
- [104] Luger R, Agol E, Kruse E, et al. *The Astronomical Journal*, 2016, 152: 100
- [105] David T J, Cody A M, Hedges C L, et al. *The Astronomical Journal*, 2019, 158: 79
- [106] Wang S, Jones M, Shporer A, et al. *The Astronomical Journal*, 2019, 157: 51
- [107] Brahm R, Espinoza N, Jordán A, et al. *The Astronomical Journal*, 2019, 158: 45
- [108] Heller R, Hippke M, Freudenthal J, et al. *Astronomy & Astrophysics*, 2020, 638: A10
- [109] Kunitomo M, Matthews J M, Ngo H. *The Astronomical Journal*, 2020, 159: 124
- [110] Schanche N, Pozuelos F J, Günther M N, et al. *Astronomy & Astrophysics*, 2022, 657: A45
- [111] Schwarz P I, Dreizler S, Heller R. *Astronomische Nachrichten*, 2026, 347: e70059
- [112] Yana Galarza J, Ferreira T, Lorenzo-Oliveira D, et al. *The Astronomical Journal*, 2024, 168: 91
- [113] Heller R, Hippke M. *Nature Astronomy*, 2024, 8: 193
- [114] Kovács G, Hartman J D, Bakos G Á. *Astronomy & Astrophysics*, 2016, 585: A57
- [115] Gibson N P, Aigrain S, Roberts S, et al. *Monthly Notices of the Royal Astronomical Society*, 2012, 419: 2683
- [116] Günther M N, Daylan T. *The Astrophysical Journal Supplement Series*, 2021, 254: 13
- [117] Gordon T A, Agol E, Foreman-Mackey D. *The Astronomical Journal*, 2020, 160: 240

A Review of Detrending Techniques for Exoplanet Transit Signals Detection

CHEN YU^{1,2}, GE JIAN¹

(1. Shanghai Astronomical Observatory, Chinese Academy of Sciences, Shanghai 200030, China; 2. University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract:

The transit method has become one of the most effective techniques for detecting exoplanets, but its precision relies on the reliable extraction of weak transit signals from high-precision photometric data. In light curves obtained by space missions such as *Kepler* and *TESS*, complex noise arising from instrumental systematics and stellar variability often exceeds the depth of Earth-like planetary transits, posing a major obstacle to both planet detection and accurate characterization. Detrending is therefore a crucial preprocessing step in photometric data analysis, aiming to remove non-transit components while preserving true transit signals as much as possible and improving the signal-to-noise ratio.

In this review, we systematically summarize the major detrending techniques currently used in exoplanet transit studies and group them into three categories. The first category is the widely adopted “detrend-then-search” paradigm, including smoothing and fitting methods such as running medians and splines, signal decomposition methods based on wavelets or cosine filtering, and probabilistic or time-series models such as Gaussian processes and autoregressive models. The second category is joint modeling, in which noise and transit signals are fitted simultaneously to reduce signal distortion introduced by conventional stepwise procedures. The third category is data-driven “direct transit detection,” where deep-learning models are used to identify transit events directly from noisy light curves. As a related case study, we also describe ProDeNet, a model based on the DeepLabv3+ architecture. [Although its current test set is small, preliminary results suggest that segmentation-based direct detection has potential for identifying low-S/N transit events.](#)

Comparisons among existing methods indicate that there is still no universally optimal detrending algorithm. The most suitable approach depends on the scientific objective, stellar type, and data quality. This review summarizes the applicability and future prospects of different methods, and provides a reference for the analysis of high-precision photometric data and the extraction of transit signals in future studies.

Key words: exoplanet; transit method; detrending; light curve; joint modeling; deep learning