

基于时序数据库的 任意时标光变曲线生成

国家天文台

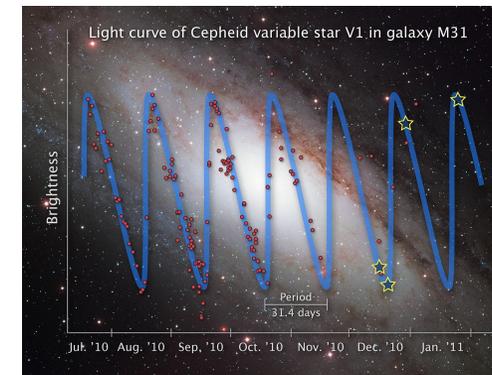
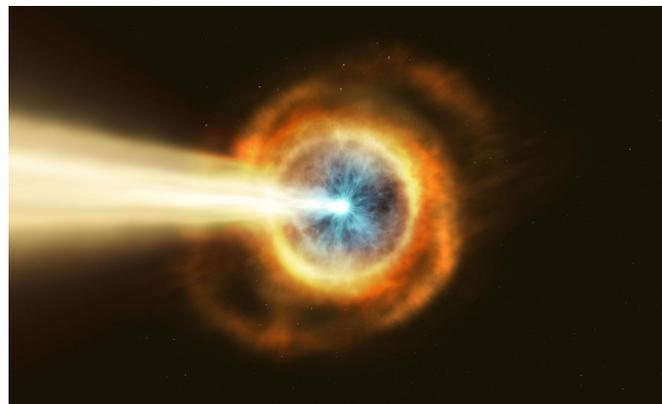
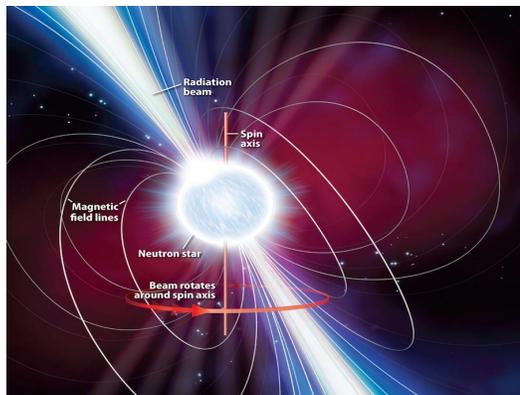
国家天文科学数据中心

张震

目录

- 需求和背景
- 传统方案存在的问题
- 基于时序数据模型的解决方案
- 优势
- 应用场景和未来工作

科学引领：宇宙中的变源跨越数个量级的时标



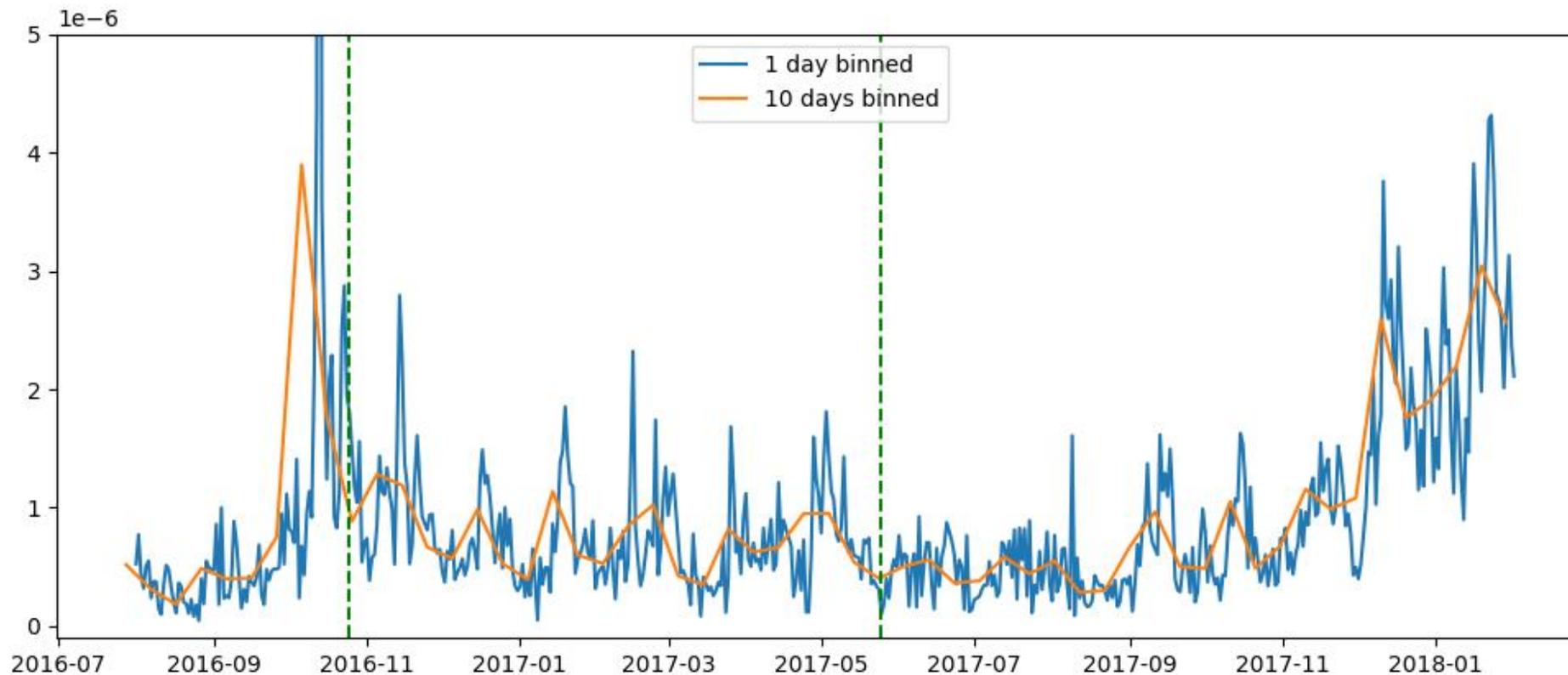
毫秒脉冲星 恒星耀发 伽马暴 双星 系外行星 造父变星

毫秒

年

不同的研究对象，需要选取不同的timebin

耀变体 PKS 2247-131

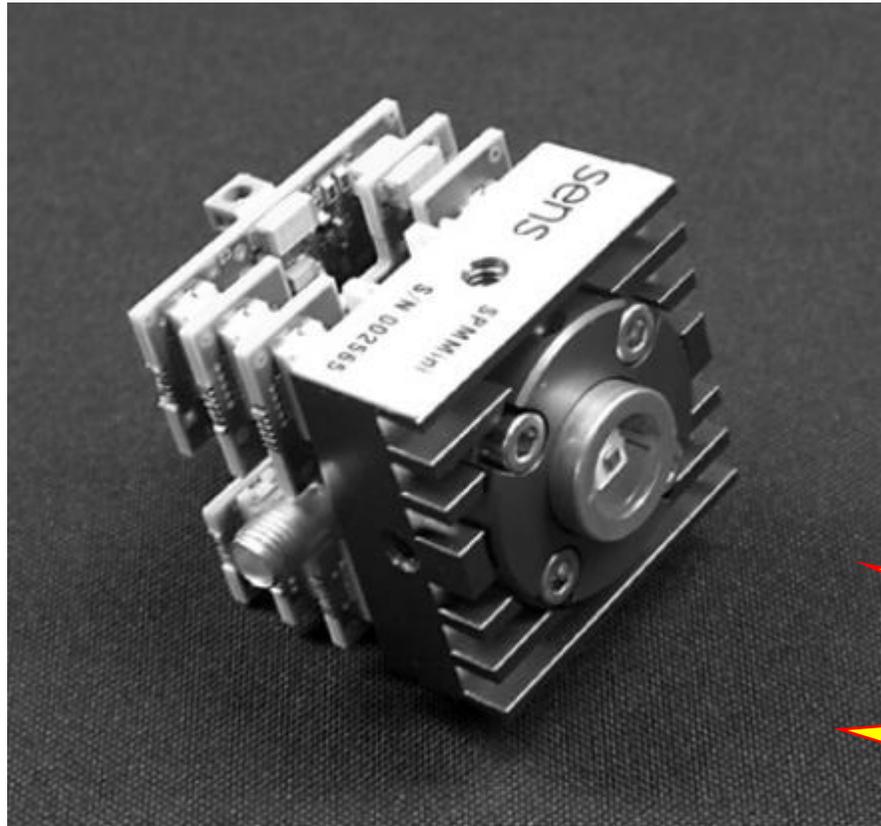


timebin < 5day : Test Statistic Value<9

timebin > 20day: 不能找到34.5天的周期

——Zhou, Jianeng 等. 2018. 《A 34.5 day quasi-periodic oscillation in γ -ray emission from the blazar PKS 2247-131》

技术推动： 望远镜更高的时间分辨率



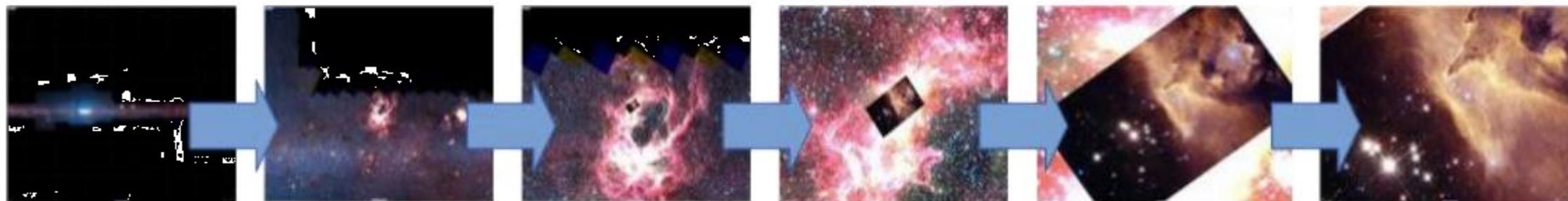
High Speed Photometry
时间分辨率提高到了

ms*、*ns级别

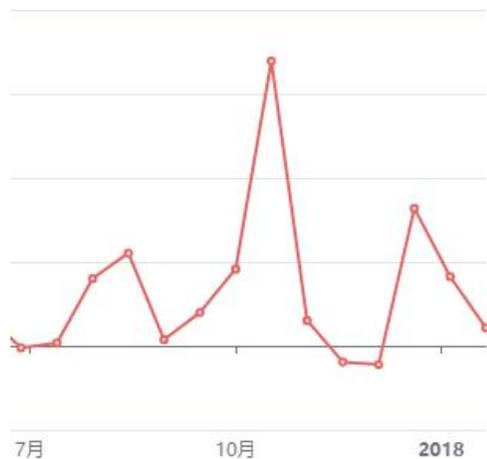
Data Volume

硅光电倍增器

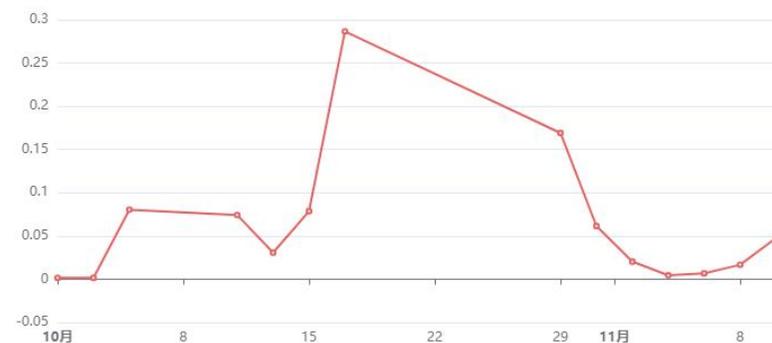
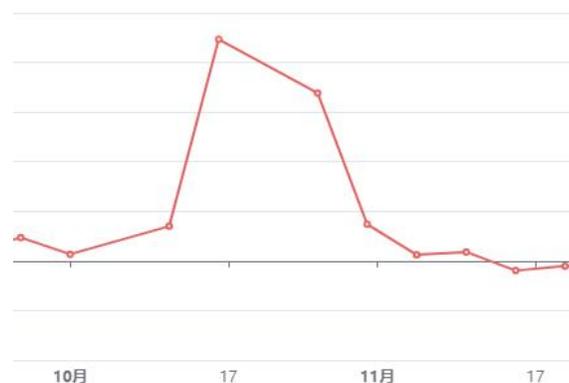
他山之石：天文数据分层可视化的HiPS技术



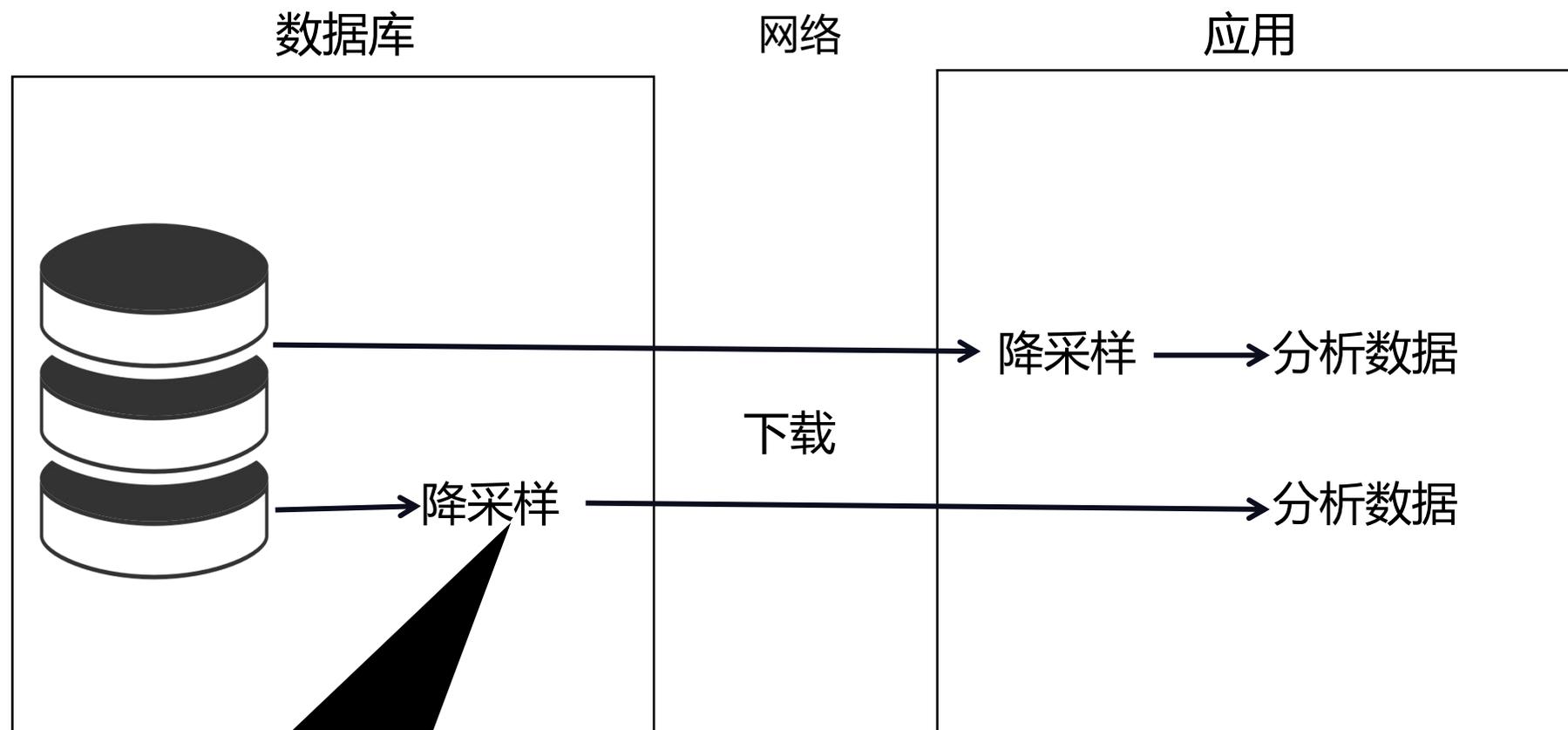
低分辨率全景



高分辨率局部



方法变革：先聚合，再下载



需求：在数据库完成降采样

方法变革：关系数据模型难以满足需求

Srcid	Ra	Dec	Time(mjd)	Flux / 1e-8
1	1.581	20.203	55055.232025	0.044991	
1	1.581	20.203	55055.359097	0.028534	
1	1.581	20.203	55055.422627	0.021105	
1	1.581	20.203	55055.486157	0.013430	

关系数据模型将一行数据存储在—起

列数据模型将—列数据存储在—起

针对时序数据的特点，工业界对包含时间范围的存储和查询进行了针对性优化，提出**时序数据模型**，在监控等领域得到广泛应用

曲折前行：传统的建表方案不再合适

Time	Src_id	Instru_id	Energy Band	Flux	Flux_err
55055.232025	1	1	2-20 keV	0.0355	0.0086
55055.232025	1	2	4-10 keV	0.0355	0.0086
55055.232025	2	1	2-20 keV	0.0385	0.0282

Same Timestamps

曲折前行：基于时序数据模型的方案

STable

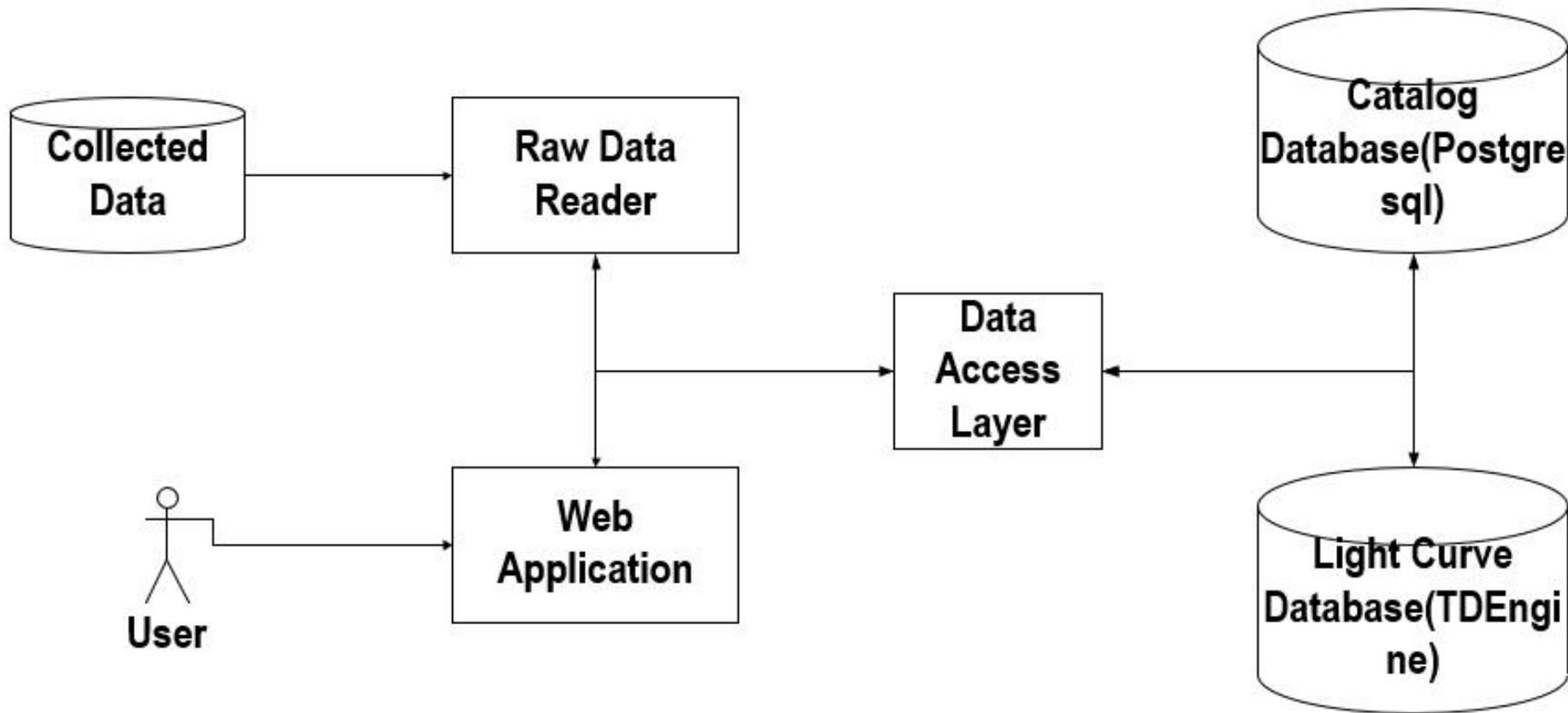
Timestamp	Fields			Columns		
start_time	src_id	instru_id	energy band	exp_time	flux	flux_err

Table

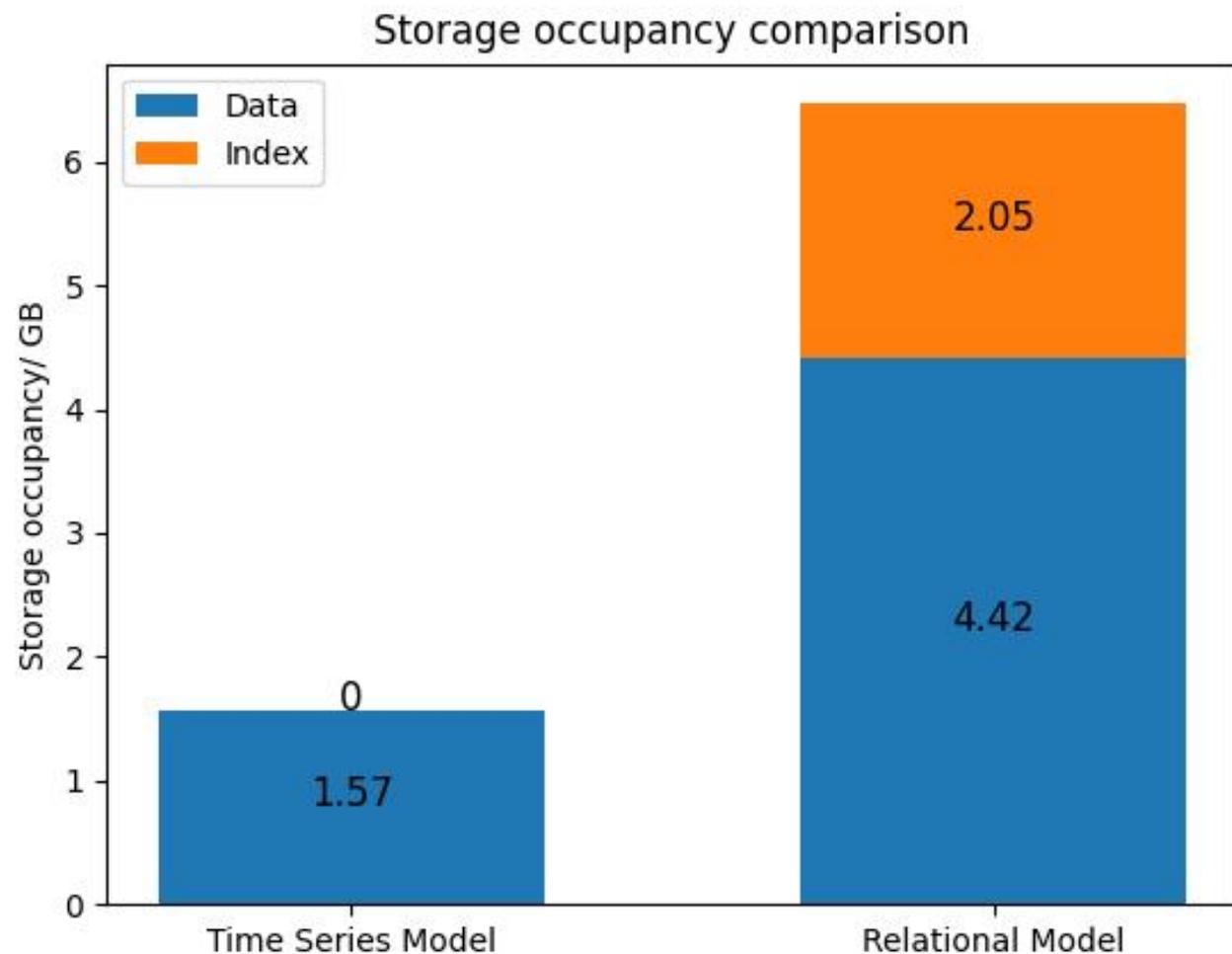
src1instru1			
src_id=1 instru_id=1, energy_band=0.5-2keV			
start_time	exp_time	flux	flux_err
...

src1instru2			
src_id=1, instru_id=2, energy_band=2-4keV			
start_time	exp_time	flux	flux_err
...

应用层设计



高效压缩

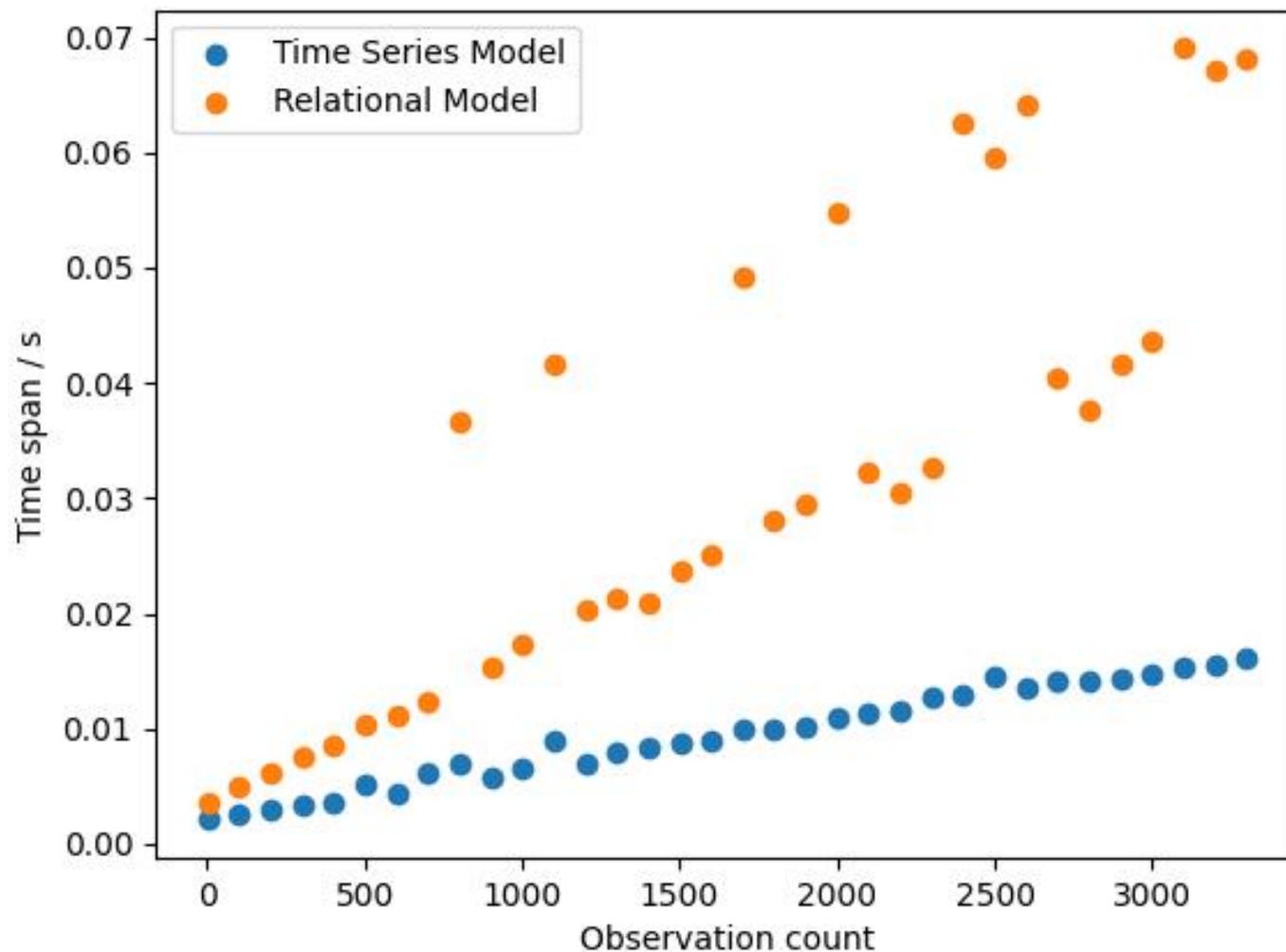


以MAXI 数据为例,

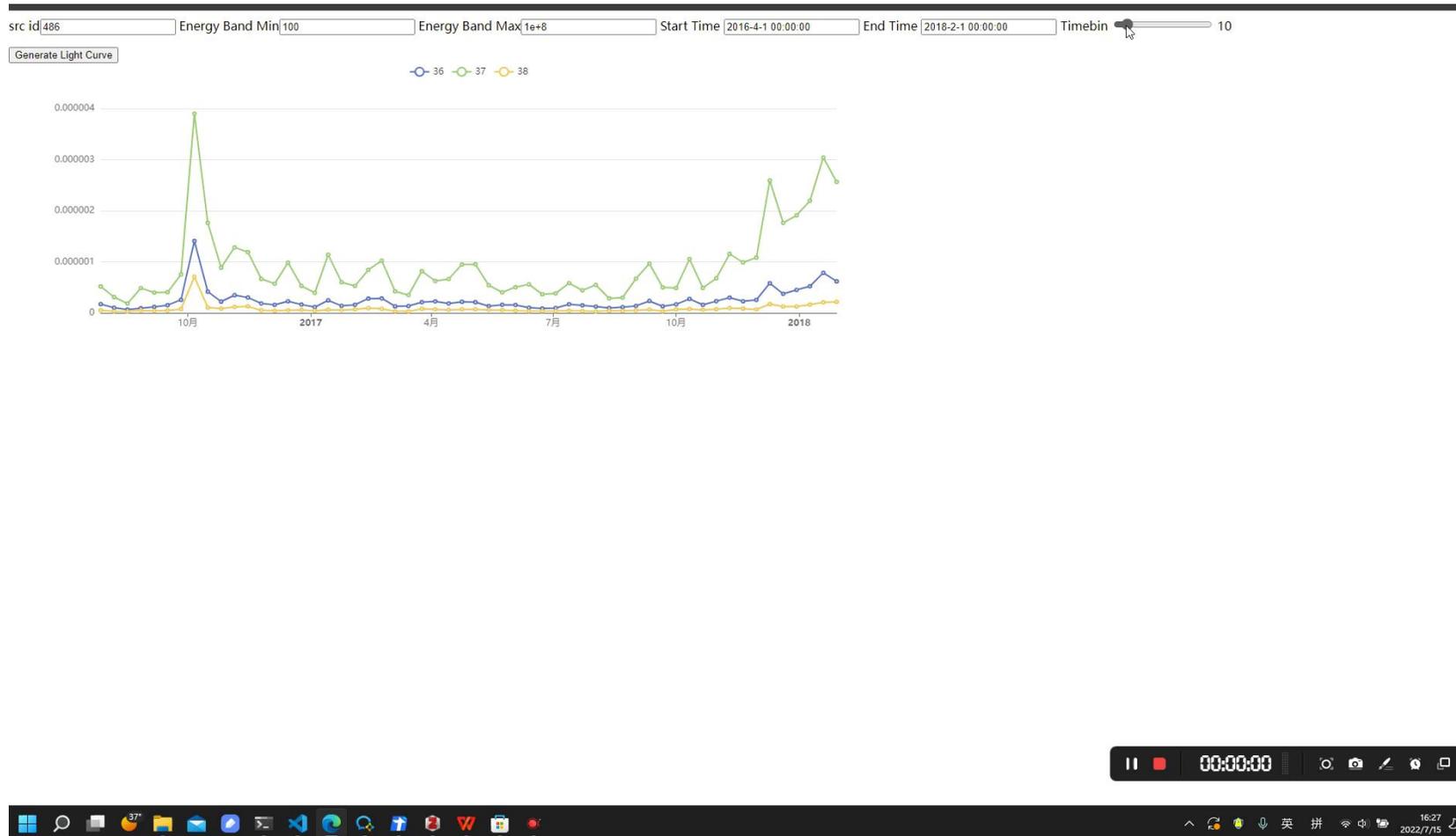
存储空间仅占用原方案**25%**

高效读取

查询速度提高**2倍**



便捷使用



未来工作

- 更友好的界面
- 更多的数据源：Gaia DR3测光数据
- 考虑应用场景：例如基于时序数据模型的变源证认
-

谢谢大家

敬请批评指正！