

品質評価結果の使い方

- その後のデータ利用の方針を定める材料とする
 - ① 捨てる（そのデータは使わない）
 - ② 補正をかけて使う（補正がうまく効くと確証が取れればの話）
 - ③ （科学的）目的に対して十分な精度であるならばそのまま使う
- 品質情報を開示（図示などが一般的）、利用者の判断を促す
 - ✓ データ検索の段階で開示情報をも利用したスクリーニング

例えば、製造業の製品品質調査と基本的考え方は同じ

- ✓ 不良品廃棄
- ✓ 訳あり商品を安価で販売
- ✓ 手を入れて作り直す 等

現在までの成果（探索的データ分析）

• 変光天体

1. 変光天体の高速かつ高精度検出に向けた取り組み

- 阪大チームとの共同作業（isolation-forest等の異常検知の手法を用いる）
 - 既存手法との処理速度や検出精度の比較を検討・試験中
 - 目視などの手法による正答データの確立と機械学習への発展
 - 変動天体に対する様々な時系列解析（単OU（DRW）、無限OU、CARMA、周期解析等） -> 分類

1 視野で約1000個ほどの変動天体候補を同定

2. 変光天体検出のより大規模なデータへの適用

- 1桁大きなデータセットに開発中の手法を適用（3.5億レコード -> 39億レコード）（1桁増）
 - データ品質の確認（開発手法の適用性の確認に必須）

手法の工夫でより高精度・高感度を実現へ

3. 銀河進化プロジェクトとの連携（e.g. 銀河とBHの共進化の研究）

- 変動情報を用いた独立な記述パラメータの導入など

4. 移動天体プロジェクトとの連携

- 変動天体検出処理の中で見つかった移動天体の調査・研究

変光天体と移動天体の取り組みで用いる
天体リストの種類が異なる

• 移動天体

1. 既知移動天体の同定

- 既知でNASAのデータベースに登録されている天体の画像（データベース）上での検出
 - 軌道精度の改良
 - 変化の確認（形・色・運動など） => 太陽系外縁部に居る周期型彗星の同定に成功！

2. 新規移動天体の検出

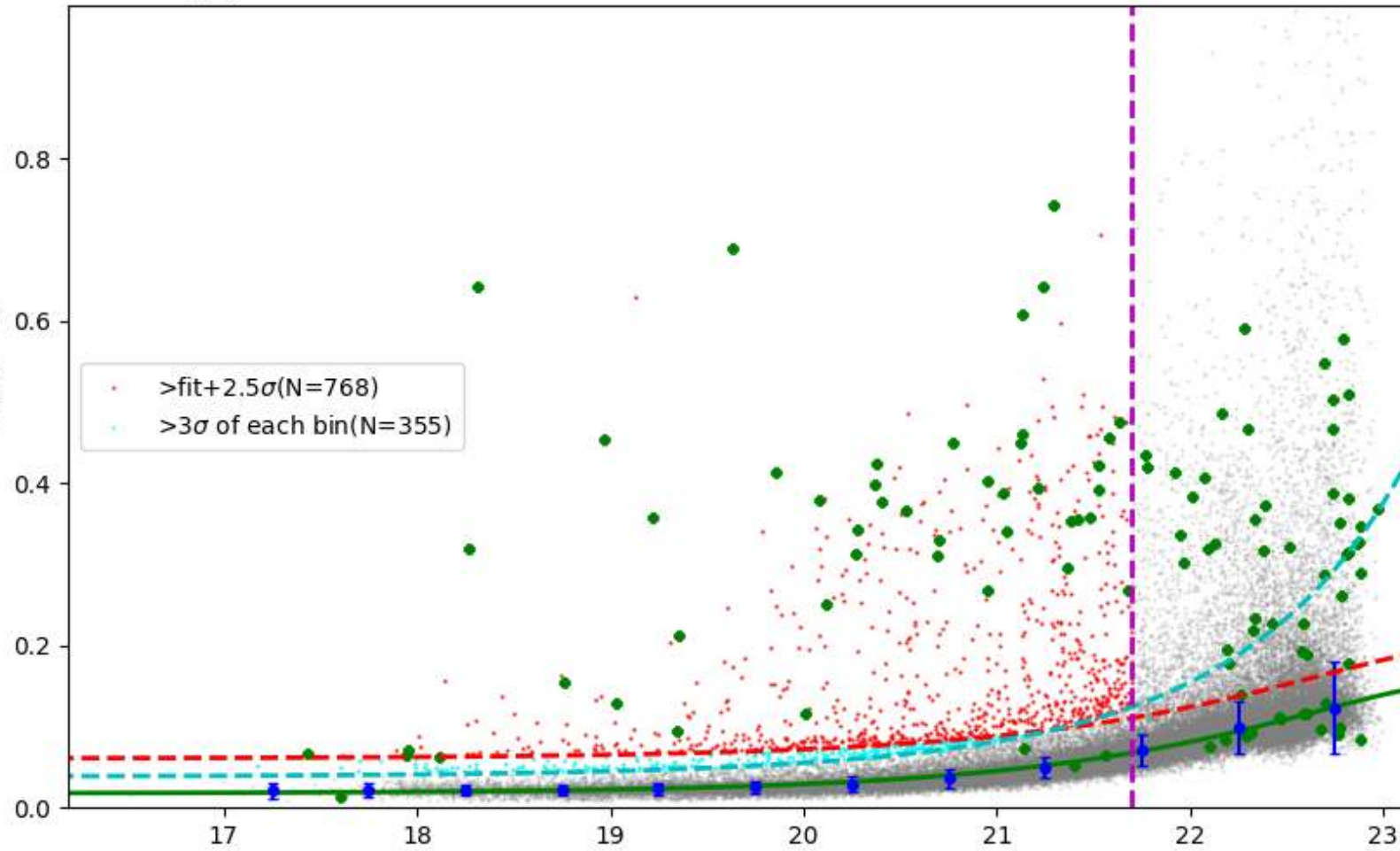
- より暗い移動天体の検出 -> 太陽系の起源に迫る研究

異常検知手法を用いた変動天体検出

Isolation-forestを用いた異常検知手法による変動天体検出の試験的実装 by 阪大

多数回測定における各天体の明るさの標準偏差値

$$\xi(m) = 1.72083942e-02 + 1.16396967e-10 \cdot 10^{0.4m} + -2.49651343e-20 \cdot 10^{0.8m}$$



多数回測定における各天体の平均的明るさ (中央値)

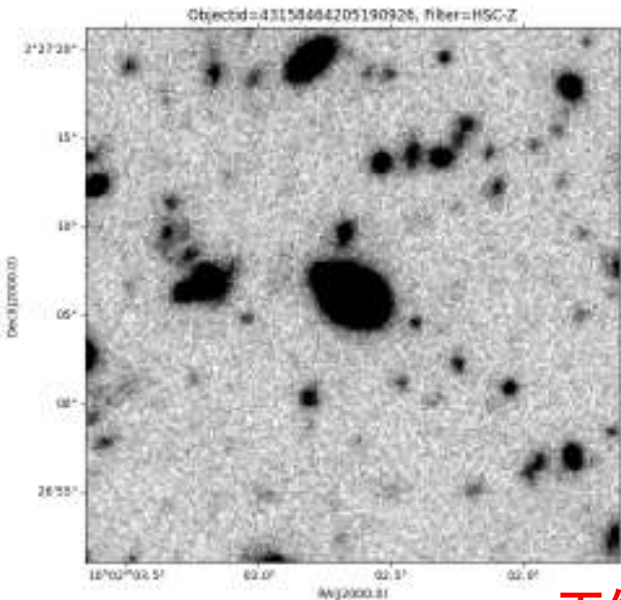
赤点：変動天体候補
緑点：サンプルデータについて異常検知した天体

現状2.4億個のデータについて40分弱で分析を終了。

より高速化を目指している。

突発天体の検出

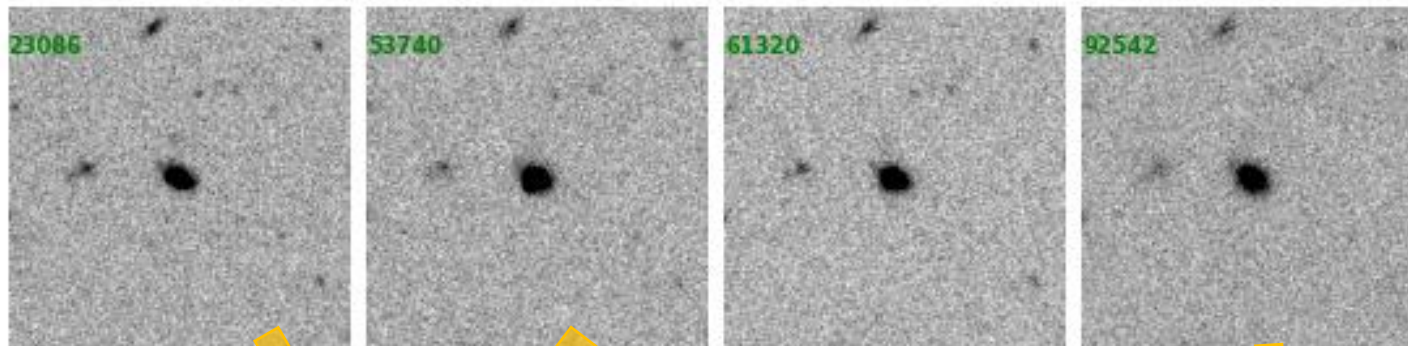
Coadd



赤方偏移 0.2118の銀河 (25.8億年前)

objectid=43158464205190926 Filter=HSC-Z

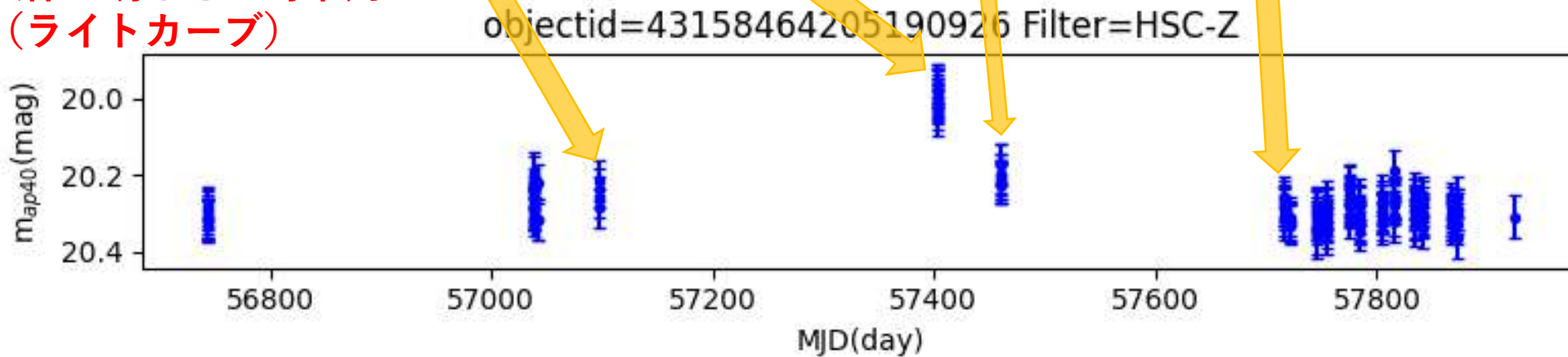
Exposures 今回の解析では9個の突発天体を同定



天体の明るさの時系列
(ライトカーブ)



系外銀河で発生した超新星
(extragalactic supernova)



2021/10/12

NEDOプロジェクトユーザー会

15

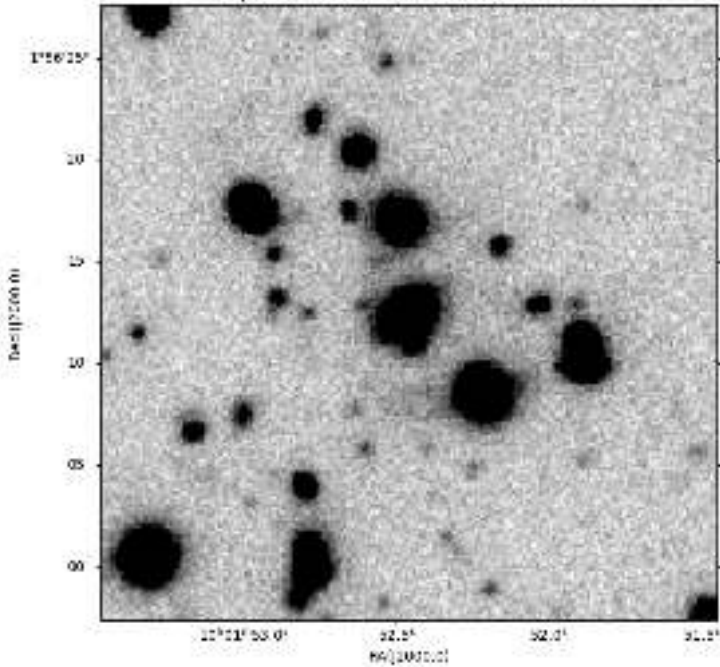
移動天体も副産物として見つかる

objectid=43158451320295165 Filter=HSC-Z

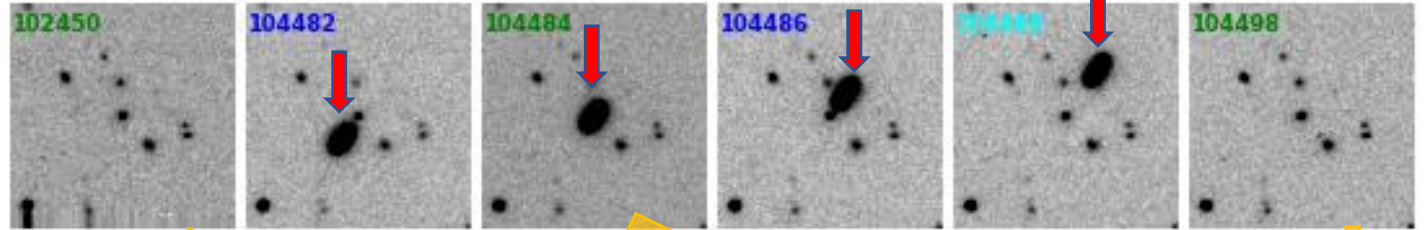
今回の解析で79個を同定

Coadd

Objectid=43158451320295165, Filter=HSC-Z

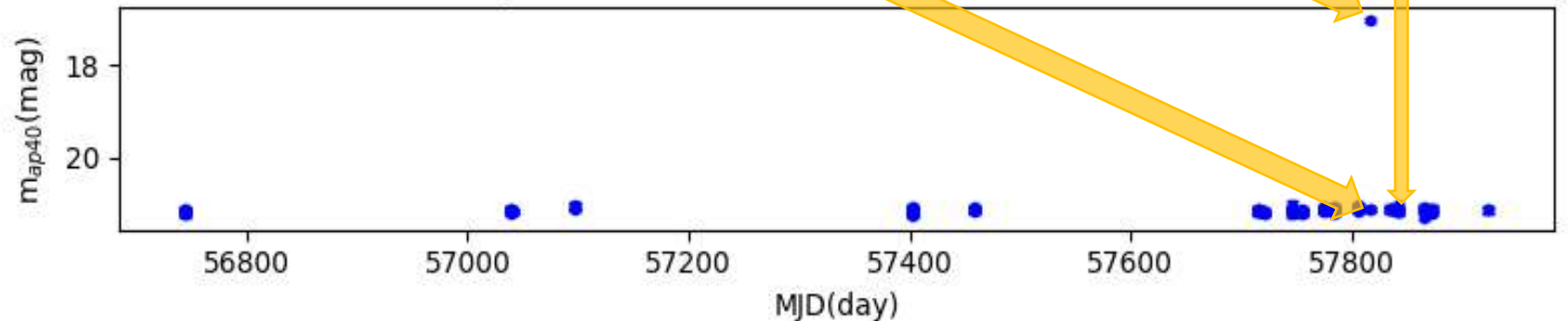


Exposures



天体の明るさの時系列
(ライトカーブ)

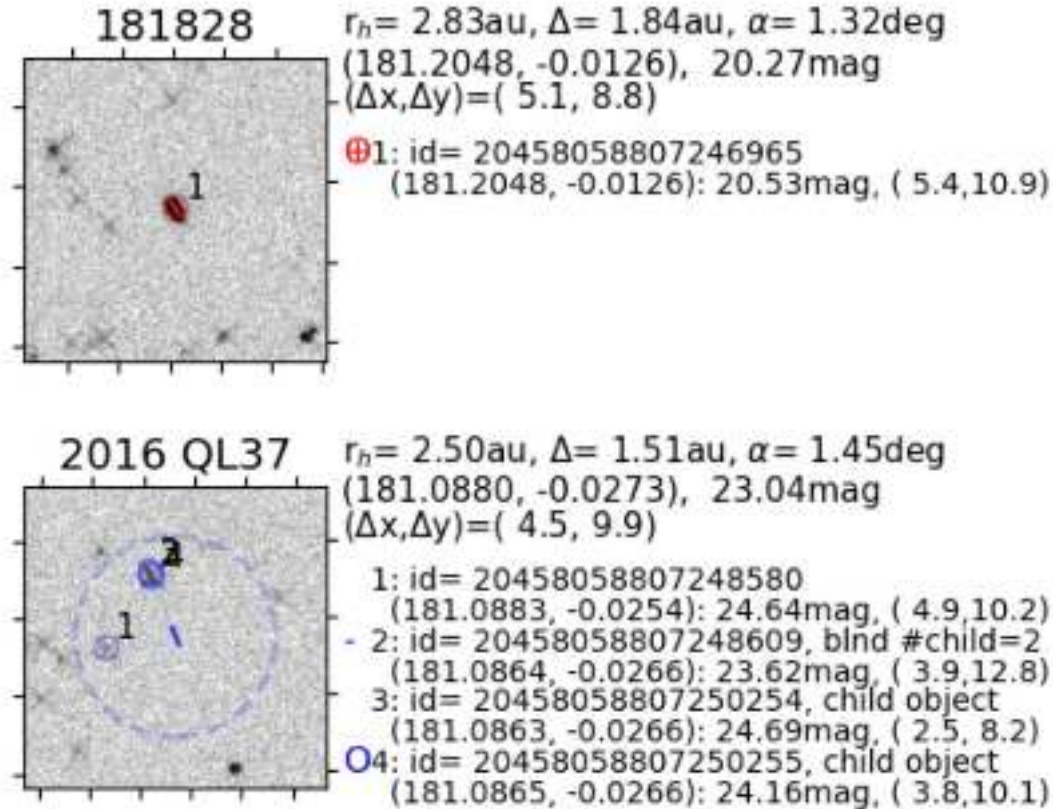
objectid=43158451320295165 Filter=HSC-Z



小惑星などの移動天体

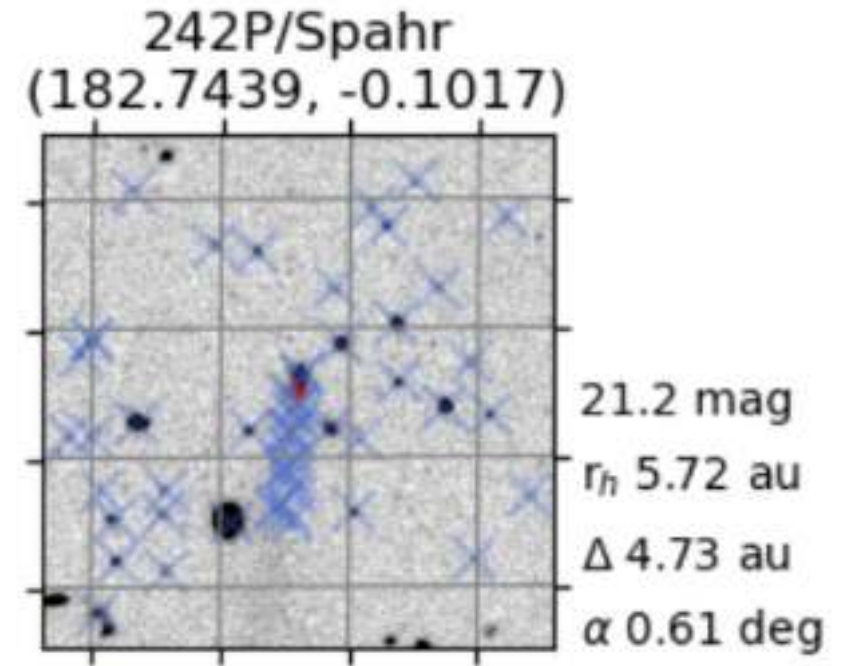
既知移動天体の同定

asteroids (小惑星)



点源天体である小惑星は移動による像の伸びや明るさも条件に入れることで自動判定できそうである

comets (彗星)



彗星はコマ・尾が複数の source として検出されている。まずは目視確認が必要

今後の展開 (1) --継続・拡張--

• クエリーの高速化

- ✓ 阪大との協力による高速化の実装->Cardinality Estimation等 (現存サービスの性能・機能向上)
- ✓ (より) 多くのパラメータを用いたデータ分析
 - 例えば、天体までの距離推定の高精度化や高速化に関する研究->銀河進化の研究につなげる

• 変光天体探査

- ✓ 20視野以上についての分析 (延べ~40億天体 (5つのフィルター (波長) の情報) について) -> BH形成進化の研究、(特殊な) 変光星の検出と詳細研究 など
- ✓ 1万個に及ぶ変動天体候補の同定が可能に
- ✓ 今後数年で1000視野について同様の解析を行う事が目標 (延べ約400億)

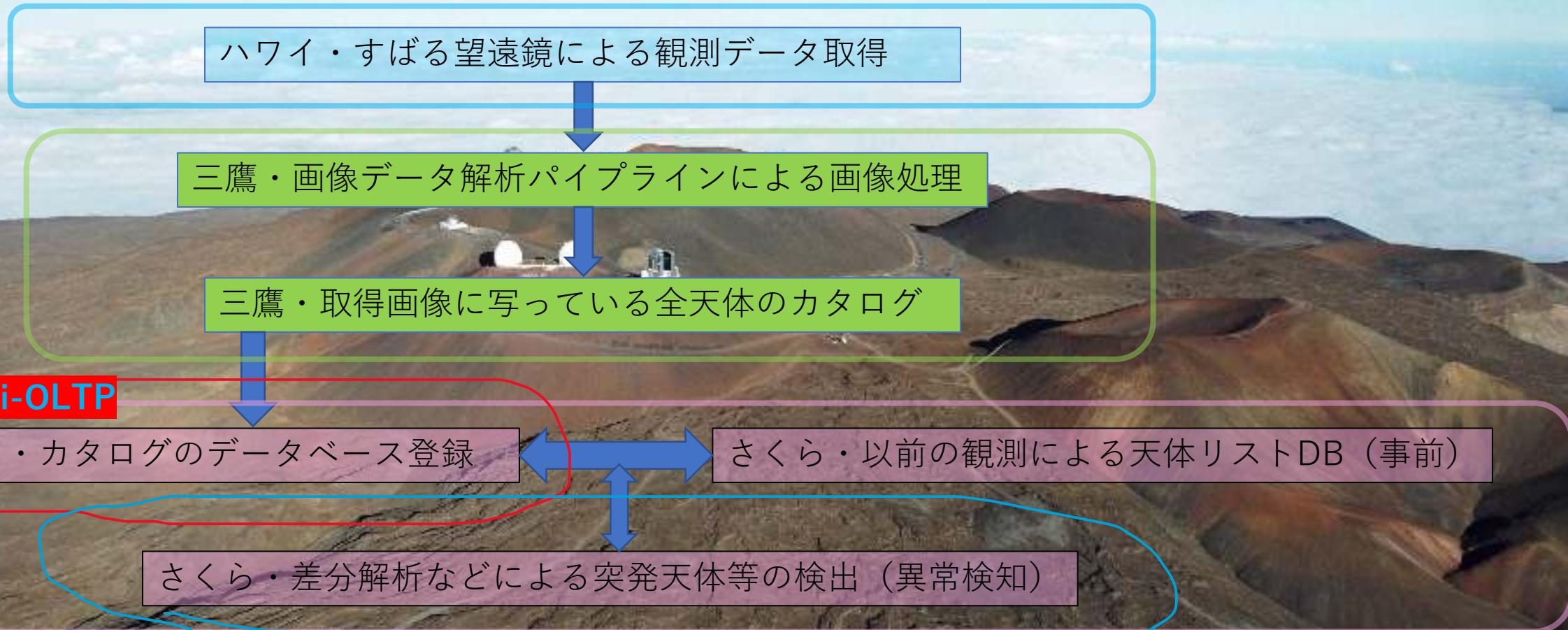
• 移動天体探査

- ✓ 既存天体 (小惑星・彗星) のデータベース上での存在確認->太陽系の成り立ちを探る
発見前の遠方にあった (12au) 時を検出、土星軌道より外側時の核の検出等、興味深い発見がなされている。(2021年秋季天文学会で発表)
- ✓ 新移動天体の発見 → 地球 (我々) への影響 → 世界中の天文台による追観測

高速に検索するだけでなく、分析も行うプラットフォームとして利用

今後の展開 (2) --発展--

- (準) リアルタイム分析による突発・変光・移動天体検出
➤ 今までのSpark-OLAPでの成果をTsurugi-OLTPとつないでより高速に

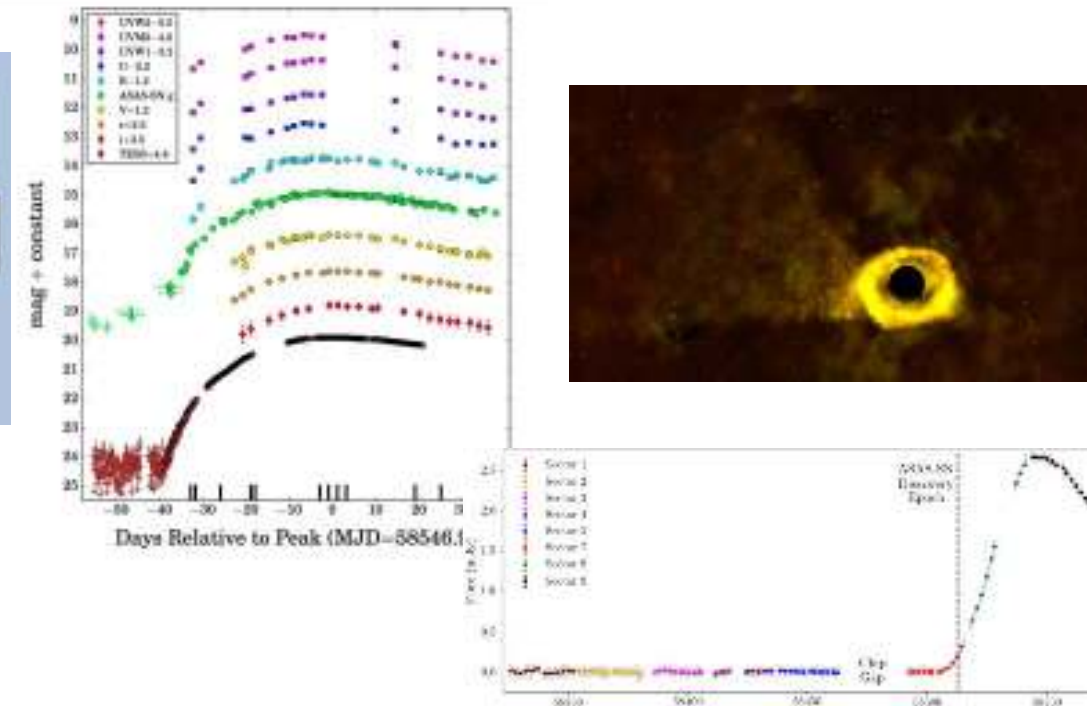
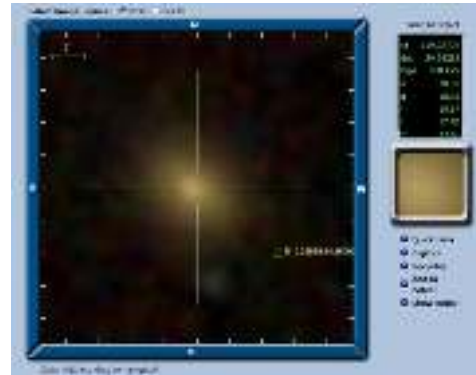
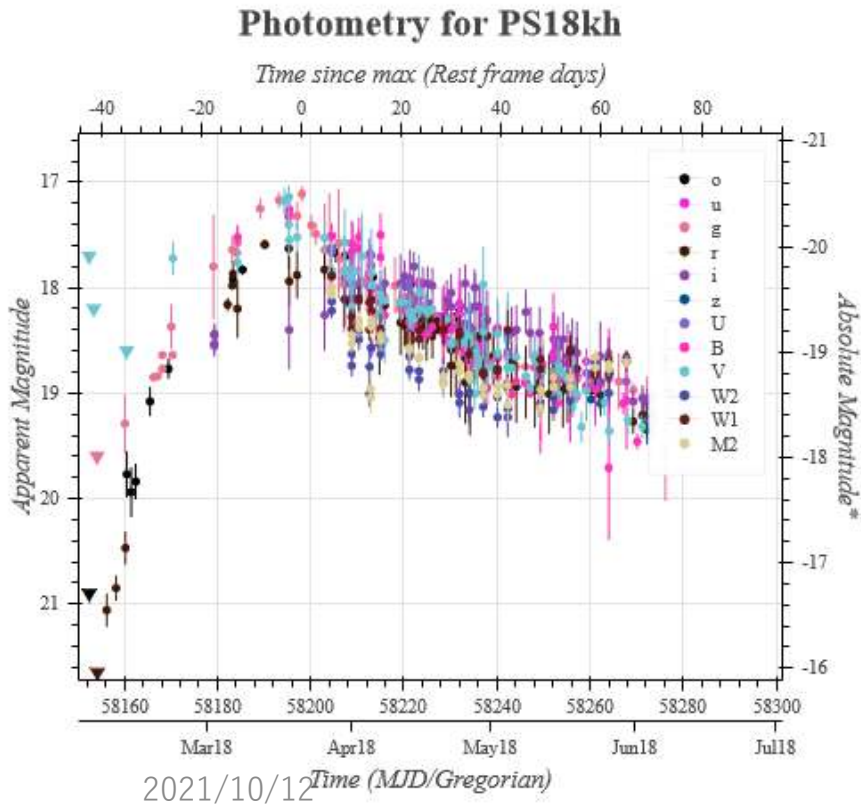


一例：Tidal Disruption（潮汐破壊）現象の検出

- （巨大）BHの重力場によって、近くを通った恒星が引きちぎられて飲み込まれ、破壊された星の質量の一部がブラックホール周囲の降着円盤に取り込まれた結果、円盤内の物質がブラックホールによって消費され、一時的に電磁放射のフレア（爆発的に明るくなる）現象が発生する。

[PS18kh](#)

ASASSN-19bt



本アプリ・システムに関する今後・情報（予定）

- すばる望遠鏡HSCの戦略枠観測のうち、特にデータ量が多い各露出に関する天体カタログのデータベースについては、本システムの構築後に公開サーバー化することを予定している。（但し、運用資金の調達に関しては未定。現在模索中です。乞うご支援。）
- 同じようなデータ戦略スキーム（大量のデータ及びそれから派生するデータや情報を用いたデータ品質評価、制御、前処理なども含めてデータ分析をする）に関する実装や性能の詳細については、別途、エンジニアリング論文またはホワイトペーパー等の形で公開する予定。

Sparkを用いたクエリ高速化についてはSPIE（The International Society for Optical Engineering）のカンファレンス“Astronomical Telescopes + Instrumentation 2020” (<https://spie.org/conferences-and-exhibitions/astronomical-telescopes-and-instrumentation?SSO=1>) において**J. Furusawa et al.(2020)**論文として発表（ご興味のある方はご参照下さい。）