

ALMA 較正用天体の自動解析ツールの開発とその科学的応用

成田佳奈香¹・西田明正²

〈^{1,2} 東京大学大学院理学系研究科 〒113-0033 東京都文京区本郷 7-3-1〉

e-mail: ¹naritakn@gcc.u-tokyo.ac.jp, ²nishida-akimasa990@g.ecc.u-tokyo.ac.jp

アタカマ大型ミリ波サブミリ波干渉計（ALMA）は運用開始から10年が経過し、観測データが日々蓄積されている。新たな観測提案に加え、膨大なアーカイブデータを有効活用することが科学的成果を最大限に引き出すうえで重要である。本記事前半では、ALMAで較正用天体として観測されているクエーサーのデータダウンロードから解析までの自動化を紹介する。後半では、較正用天体を用いたサイエンスの一例として、クエーサーの連続光を背景光源とした、銀河系内の分子雲の吸収線解析を紹介し、分子雲の化学的多様性や構造について得られた知見を紹介する。

1. はじめに

観測開始から10年を迎えたALMAのアーカイブデータサイトでは、PIデータ占有期間を経たサイエンスターゲット天体のデータのみならず、較正用天体の観測データも公開されている。較正用天体とは、大気の影響による電波信号の振幅や位相の乱れを補正するために観測される天体で、サイエンスターゲット天体の近くに位置する別の天体である。較正用天体として多く選ばれているのが、クエーサーと呼ばれる天体である。

このクエーサーの連続光を背景に、手前の分子ガスが影絵（吸収線）として図1のように観測される。分子吸収線系として観測されるような希薄な分子ガスでは、水素分子との衝突励起が十分に

吸収線系においては、現在の宇宙背景放射（CMB）温度を仮定することで、複数の回転遷移を観測していなくとも分子の柱密度を決定することができる。宇宙論的遠方天体の場合は、逆に複数の吸収線から励起温度を求めることで、その赤方偏移のCMB温度を測定できる。分子吸収線系は物理量を高い精度で測定できる特長があり、分子雲の形成における初期条件を制限したり、輝線では捉えにくい宇宙論的距離に存在する希薄な分子ガスの物理的・化学的進化を探ったりできるなどの大きな可能性を秘めている。しかしながら、既知の吸収線系は限られており、較正用天体データも用いた大規模な探査が求められている。

2. 自動解析ツールの開発

2.1 概要と開発の経緯—なぜ自動化／パッケージ化が必要か？

本来サイエンスターゲットの較正に用いられる較正用天体に吸収線があると、サイエンスターゲットの較正精度を劣化させてしまう。そのため、観測所から配布されるスクリプトを用いたキャリブレーションでは、較正用天体の吸収線は

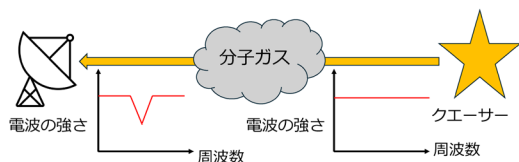


図1 吸収線観測の模式図

マスクされる。しかし今回紹介する解析では、まさしくこの吸収線をターゲットとするため、配布スクリプトは使用できず、マニュアルでスクリプトを作成する必要がある。さらに、較正用天体データはPIのサイエンスに合わせた様々な観測モードや観測時期で取得されている。単に吸収線をマスクしないだけでなく、それらに対応する解析コードを作成する必要がある。これらの要件を満たして大規模な探索を行うには、データのダウンロードから適切なキャリブレーションまでを自動で行うツールの開発が必要である。

このような背景のもと、吉村勇紀氏によって自動解析コードの作成が行われた。それを用いて2019年から2020年にかけて、当時のアーカイブデータにある702個の較正用のクエーサーの2000個を超えるデータに対して大規模な探索が行われた[1]。

図2は、ALMAで観測されたデータに対する一般的な処理と、吉村氏により開発されたツールによる自動化の概要図である。観測データを解析するためには、先んじて生データに対して較正を行う必要がある。通常であれば、配布される較正用スクリプトを実行するだけでよい。しかし、今回は吸収線をマスクしない較正用スクリプトが必要であるため、これを作成する必要がある。吉村氏のツールでは、生データのダウンロードから、較正

用スクリプトの作成とその実行までを自動化した。

吉村氏による開発から年数が経過し、本ツールは解析ソフトウェアの新しいバージョンに対応しなくなってしまった。また、ソースコードを読み解く中で、機能を改善・追加できる箇所や不要な箇所も見つかった。さらに、インストールの方法が煩雑という問題点も浮かび上がった。

そこで今回の開発では、

- ・現在の解析ソフトウェアへの対応
- ・機能の整理・追加
- ・ツールのインストールの簡易化

を目標に、コード全体の再設計から行った。

2.2 ツールの開発

吉村氏が作成した吸収線の解析のスクリプトの素案[2]を元にバグとりを行い、誰もが解析環境を手に入れられるようにPythonのパッケージ化を行なった。以下では、自動化ツールの開発（プログラミング）の様子を述べる。開発は図3のようなサイクルを繰り返すことで行われた。

Step 1：目標の明確化

Step 1として、ツールが行うべき操作を明確化した。本ツールの場合、最終目標は「較正を自動で行う」ことである。その実現のため、最初に「データをインポートする」必要があり、その次に「較正用のスクリプトを生成する」という操作が必要である。こういったことを明確にすること

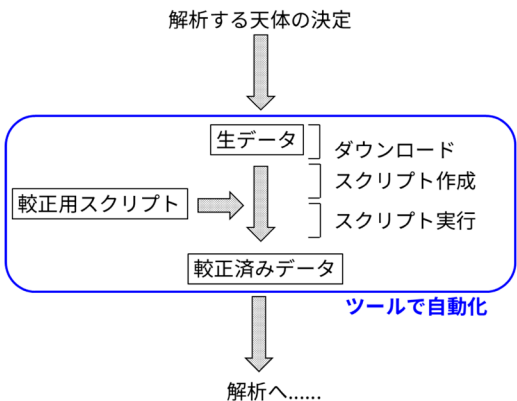


図2 ALMAデータの解析を始めるまで

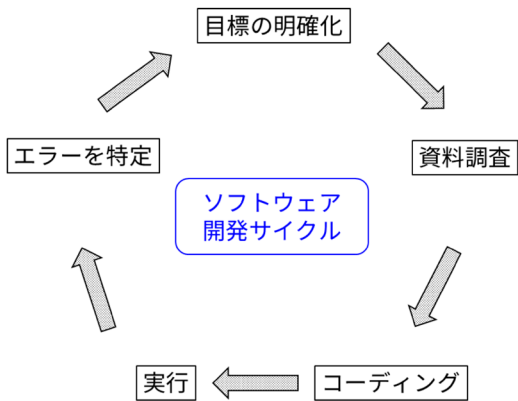


図3 ツール開発のサイクル

で、自動化のための道標が得られる。研究目標を定めることに近いかもしれない。

Step 2：資料調査

Step 2では、Step 1で設定した目標を達成するために必要なコマンドや関数の特定を行った。必要なコマンドや関数とその使い方を理解するために、解析ソフトのマニュアルやライブラリのドキュメントなどを調査した。研究では論文調査にあたるだろう。今回の開発での「データをインポートする」操作を例にとると、解析ソフト「CASA」にALMAのrawデータをインポートするコマンドは「importasdm」である。

Step 3：コーディング

Step 3で実際にコードを書いた。ゼロからのコーディングは大変な作業であるから、生成AIを使うのは1つの選択肢だと考える。目標の明確化や資料調査を行っているならば、生成AIに「このような機能を実装したい。このコマンド/関数を使って、コードを書いて」と具体的に指示を出すことができる。具体的な指示に対する生成AIの回答精度は比較的高い。

Step 4：実行 & Step 5: エラーを特定

Step 4でコードを実行し、意図通りの動作をするか確認する。1回で正しく実装できればよいが、悲しいことにほとんどの場合でそうはならない。ここまで慎重に各Stepを踏んだとしても、必ずどこかに誤りが潜んでいる。そこでStep 5として、このエラーの原因を特定することになる。ログや出力結果を確認したり文献に再度あたったりする必要がある。生成AIとのやり取りも行いつつ、原因を切り分けていく。これが作業全体で最も時間のかかる部分である。研究での試行錯誤がまさにこのStepに対応しているだろう。辛い作業ではあるが、エラーが解消したときの爽快感は格別である。

ここまでで、Step 1で定めた目標を実現する

コードが完成する。その後は再度Step 1に戻り、次の機能の追加を行う。このサイクルを繰り返して、本ツールは完成した。

開発にあたって取り組んだ課題

最後に、開発にあたって取り組んだ課題についても述べる。

CASAはPythonの上に構築されたソフトウェアであり、スクリプトもPythonで記述することができる。これ自体は非常に便利なのだが、CASAが独自のPython環境を持つがゆえに、本ツールに必要な追加のライブラリを自由にインストールすることができない状況だった^{*1}。CASAと本ツールの環境をそれぞれ保ったまま使い分けるとというのが本開発の課題の一つであった。

詳細はテクニカルな話題となるので立ち入らないが、外部からのCASAの呼び出し方法を調査・試行錯誤することで、この課題を解決することができた。

2.3 自動化の成果

本ツール、すなわちALMA観測の未校正データ(ASDM)から校正済みのmeasurement setを作成するツールalmaqsoはZenodo [3]よりインストール可能である。

また、今後の開発として、校正のみならず画像作成機能の追加も予定している。ご興味のある方はそちらもお待ちいただくと幸いである。

3. 科学的应用

ALMA校正用天体のデータを用いた科学的な応用例を紹介する。

J1851+0035方向の銀河系内吸収線探索

ここでは、J1851+0035と呼ばれるクエーサーに着目した研究を紹介する。本天体はALMAで校正用天体として頻繁に観測されている。本研究では、J1851+0035の全受信バンドにおける大量の校正用観測データをダウンロードし、銀河系内

^{*1} 正確には「技術的には可能ではあるが望ましくない」

吸収線を探査・解析した [5].

観測結果

ダウンロードしたアーカイブデータを解析することにより, J1851+0035 方向の視線上で, 14 種 (CO , ^{13}CO , HCO^+ , H^{13}CO^+ , HCO , C_2H , $c\text{-C}_3\text{H}$, $c\text{-C}_3\text{H}_2$, CN , HCN , HNC , CS , SO , SiO) の分子が検出され, そのうちの 5 種 (HCO , C_2H , $c\text{-C}_3\text{H}$, CN , HCN) には顕著な超微細構造が見られた. ALMA で得られた代表的な吸収線プロファイル

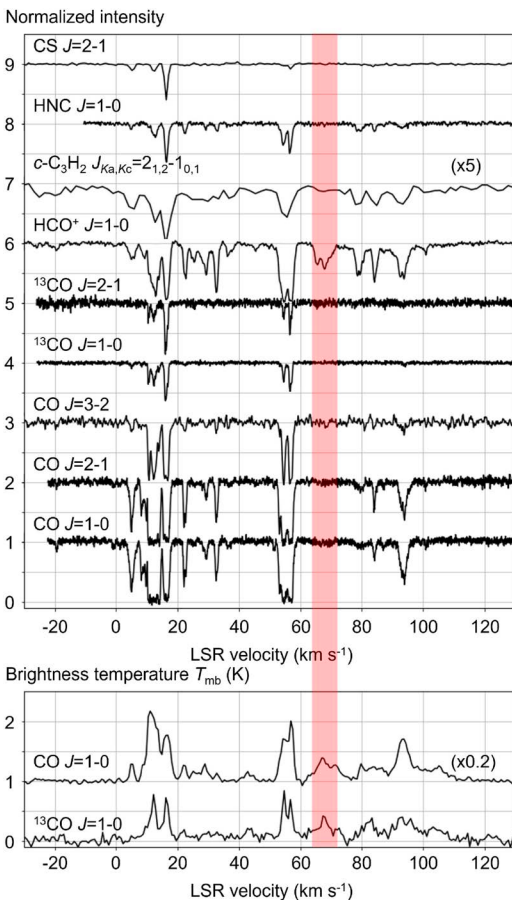


図4 ALMA で検出された吸収線のプロファイル (上パネル) と NRO 45 m での CO 輝線のプロファイル (下パネル). 薄く塗り潰された領域は, ALMA では CO 吸収線が未検出だが, NRO 45 m では CO 輝線が観測されている. [5] より転載.

と, 同じ領域を国立天文台 野辺山 45 m 電波望遠鏡 (NRO 45 m) で観測した輝線と吸収線プロファイルでは明らかな違いが見られる. 例えば, 図4の薄く塗りつぶした領域 (視線速度 $\sim 70 \text{ km/s}$) のように, ALMA では CO の吸収がないにもかかわらず, NRO 45 m による輝線データでは CO の強い輝線が確認できる.

吸収線・輝線解析で見てきた分子雲の構造

吸収線と輝線プロファイルの違いは, 雲が内部構造を持っていると考えることで説明できる. 図5は吸収線と輝線の観測を比較したものである. NRO 45 m のビームサイズは 15 秒角であるのに対し, クエーサーの大きさはミリ秒角オーダーであり, 秒角以下の分解能を実現する ALMA の極めて細いペンシルビームによってクエーサーの大きさに肉薄することが可能である. 分子雲の内部に単一鏡では分解できないコンパクトな構造があり, またそれらがクエーサーの視線方向になく ALMA のペンシルビームから外れてしまう場合, 単一鏡の広いビームでは輝線として検出される一方, ALMA では吸収線が見えないという結果が説明できる. そこで, HCO^+ と CO の吸収線, CO の輝線のプロファイルを比較することにより, 拡散した分子ガスの中に分子ガス塊が浮いているような雲の構造 (図6) を提唱した.

Extream CO-poor ガスの発見

吸収線のプロファイルより, 速度成分から個々の雲を同定した. また, 分子が CMB 温度で励起

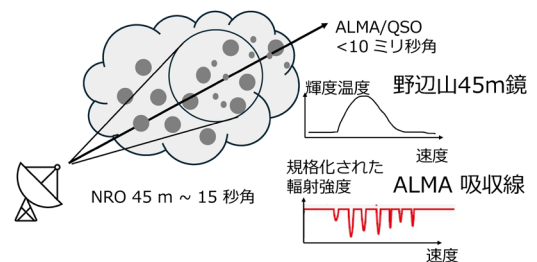


図5 ALMA での吸収線観測と NRO 45 m での輝線観測の比較図.

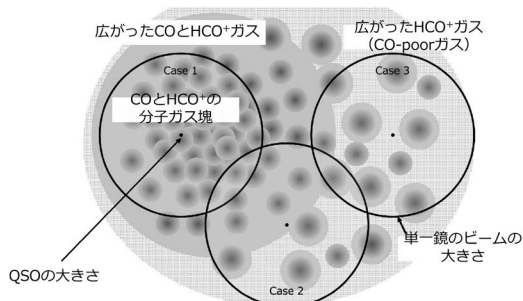


図6 我々が提唱した雲のモデル。[5]より転載。

されていると仮定すれば、複数の回転遷移で得られた吸収線や他分子の同一速度コンポーネント内で分子の光学的厚みを積分することで、柱密度を導出することが可能である。検出された分子吸収線で同定したコンポーネント毎に柱密度を導出した。その中でも HCO^+ と CO の結果について紹介する。

先行研究[6]より、diffuse cloudのような希薄な分子雲環境では、 $N(\text{HCO}^+)/N(\text{H}_2) = 3 \times 10^{-9}$ 程度の柱密度比が一般的に見出されている。この値は、星形成雲の外縁部などの低密度領域 ($< 10^3 \text{ cm}^{-3}$) における観測結果[7]とも一致する。これは、希薄な環境において HCO^+ が H_2 の良好なトレーサーであることを示している。

では、同じく水素分子ガスのトレーサーとされてきた CO との比をとってみるとどうであろうか？ 結果は図7と図8の通りである。 CO/HCO^+ 存在比は ~ 3 桁の大きなばらつきがあり、 $\sim 2 \times 10^4$ から ~ 30 の範囲であった。図8の左パネルから分かるように $N(\text{HCO}^+)$ の低い領域は CO/HCO^+ 存在比が低いという傾向がみられた。

HCO^+ は CO を含む他の主要な分子種と正の相関を示していた。さらに図7と図8からわかるように、 CO は3桁もの柱密度の分散をもっていた。このことから、 HCO^+ や他の分子種の増加というよりも、むしろ CO の存在量が減少していると考えられる。

CO/HCO^+ 存在比を HCO^+ の柱密度と銀河中心距離 (R_G) の関数としてプロットしたのが図8であ

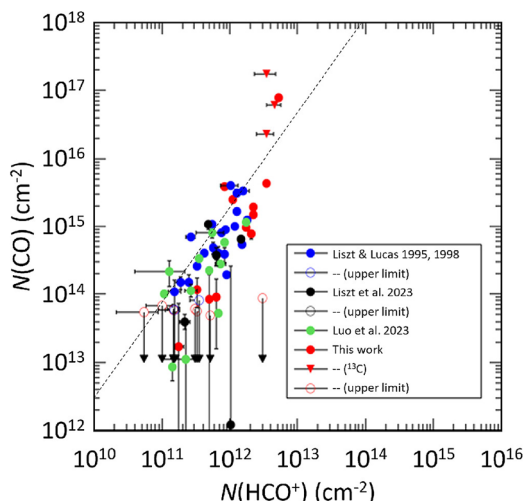


図7 $\text{H}^{13}\text{CO}^+/\text{HCO}^+$ の存在量比が50であると仮定して、 H^{13}CO^+ の柱密度から求めた HCO^+ データ点を三角マークで表す。銀河系内の吸収線の先行研究[6, 8, 9]も合わせてプロットした。[8]でフィットにより得られたべき乗則の傾きは1.4であった。[5]より転載。

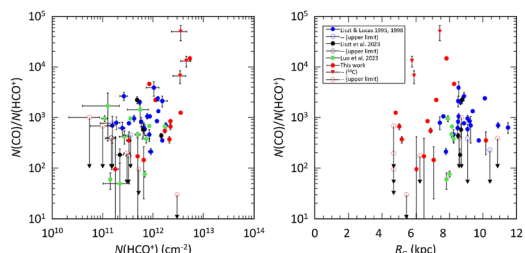


図8 HCO^+ 柱密度と銀河中心距離 (R_G) を横軸に CO/HCO^+ 比をプロットした。[5]より転載。

る。大きな分散を伴っているが、 HCO^+ の柱密度、すなわち A_V (減光量) への依存性は見えている一方で、銀河中心距離への依存性は明らかではない。“CO-dark” (ここでは“CO-poor”と呼ぶ) 分子ガスの存在は、紫外線に照射されている環境 (e.g., [10, 11]) や希薄な環境 (e.g., [12–14]) で指摘されている。さらにこの HCO^+ に対する CO の存在量の相対的な減少は、先行研究[9, 15–17]でも指摘されているが、これらの結果では最大10分の1程度とその減少幅はあまり大きくはなかった。今回の我々の結果では、暗黒雲に対する測定 (図7

で $N(\text{CO}) \gtrsim 10^{16} \text{ cm}^{-2}$ のデータ点に対応) を加えることにより, 3桁にも及ぶCOの存在量の減少が見出された. さらに, CO/HCO^+ 存在比が HCO^+ 柱密度に依存する可能性が示された. これらの結果は, COの欠乏した希薄な分子ガス(CO-poorガス)の中に微小なCO分子塊が浮いているという我々が提唱した雲の構造モデル(図6)とも整合的である. 特に, このモデルはCO/ HCO^+ 存在比の大きな分散を自然に説明し, ALMAのペンシルビームと単一鏡の広いビームによる観測結果の違いも統一的に理解する手がかりを与える.

紙面の制約上, 論文内の全ての成果を紹介することはできなかったが, この記事を読んで興味を持たれた方は是非論文[5]を手にとってみていただきたい.

4. おわりに

この記事では, ALMAの観測データの自動較正ツールの開発について説明した. ツール開発において最大の障壁となるのはエラーの特定およびその修正である. 先にも述べた通り, 本プロジェクトにおいてもこのステップに最も時間が割かれている. しかし, そのような開発コストを鑑みても, 自動化のメリットは非常に大きい. 時間と労力を大きく取られる作業がある場合は, 自動化を検討することは有力である.

一見すると, ソフトウェア開発は天文学の研究からは離れた存在のようにも感じられる. しかし, 実際には似た思考方法や試行錯誤を経て行われる. 本投稿を通じてより多くの方に自動化の利点を感じていただけたならば幸いである.

また, ここで紹介した開発は, プロフェッショナルに取り組まれている方にとっては日曜大工のようなものであろう. お気づきになった点をご教示いただけるとありがたく存ずる(西田).

ALMAの較正用天体のデータを用いた研究について紹介をさせていただいた. 普段はサイエン

スとして解析されずに見過ごされているデータであるが, このような中にもお宝となるデータが沢山眠っていることを知っていただけたら幸いである. 今回紹介したのは手動解析で得られた結果であるが, 自動化ツールにより他の天体に対して高速で適用できると考えている. 今後の進展が楽しみである. また, 今回の自動化プロジェクトは学生のみで行われた. 一緒に取り組んでくれた西田氏には大変感謝している(成田).

謝 辞

まずは本記事の校閲をされ有益な提案をしてくださった天文月報編集委の宮本祐介さんと小山翔子さんに感謝いたします. 較正用天体を用いた吸収線の自動解析のスキプトの素案を共有していただいた吉村勇紀氏には感謝いたします. 較正用天体のデータを含め, ALMAの豊富なアーカイブデータでサイエンスを行うことができるのは, アルマ地域センター(ARC)をはじめとしてALMAに関わる全ての方々の多大なるご尽力のおかげであり, この場を借りて御礼申し上げます.

参考文献

- [1] Yoshimura, Y., 2020, Masters Thesis, The University of Tokyo
- [2] <https://github.com/astroysmr/almaqso> (2025.7.30)
- [3] <https://doi.org/10.5281/zenodo.14752251> (2025.7.30)
- [4] Ando, R., et al., 2016, PASJ, 68, 6
- [5] Narita, K., et al., 2024, ApJ, 969, 102
- [6] Liszt, H., & Gerin, M., 2023, ApJ, 943, 172
- [7] Rodríguez-Baras, M., et al., 2021, A&A, 648, A120
- [8] Liszt, H. S., & Lucas, R., 1998, A&A, 339, 561
- [9] Liszt, H., et al., 2019, A&A, 627, A95
- [10] van Dishoeck, E. F., & Black, J. H., 1988, ApJ, 334, 771
- [11] Wolfire, M. G., et al., 2010, ApJ, 716, 1191
- [12] Sonnentrucker, P., et al., 2007, ApJS, 168, 58
- [13] Sheffer, Y., et al., 2008, ApJ, 687, 1075
- [14] Liszt, H., et al., 2018, A&A, 617, A54
- [15] Lucas, R., & Liszt, H. S., 1995, eds. Winnewisser, G., & Pelz, G. C. in The Physics and Chemistry of Interstellar Molecular Clouds, 459, eds.(Springer), 120
- [16] Luo, G., et al., 2020, ApJ, 889, L4
- [17] Luo, G., et al., 2023, ApJ, 946, 91



Development of an Automatic Analysis Tool for ALMA Calibration Sources and Its Scientific Applications

Kanako NARITA and Akimasa NISHIDA

The University of Tokyo, Graduate School of Science, Department of Astronomy, 7-3-1 Hongo, Bunkyo, Tokyo 113-0033, Japan

Abstract: The Atacama Large Millimeter/submillimeter Array (ALMA) has been in operation for over a decade, accumulating vast observational data. In addition to proposing new observations, effectively utilizing the extensive archive data is crucial for maximizing scientific outcomes. This study presents an absorption line analysis using quasars, which are observed as calibration sources with ALMA, as background sources. In the first part, we report on developing the Python package `almaqso`, which facilitates the downloading and calibration of archive data. The second part introduces scientific research results derived from calibrated data.