

Development of Open-Source-Based Software Planetary Atmospheric Spectrum Calculator (PASCAL) Specified for Millimeter/Submillimeter Observation of Titan with ALMA

Takahiro Iino(✉)

Information Technology Center, The University of Tokyo, 2-11-16 Yayoi, Bunkyo-ku, Tokyo 113-8658, Japan
iino@nagoya-u.jp

Abstract. To illustrate planetary atmospheric environment, such as a three-dimensional structure of trace gases, combination of high precision millimeter and submillimeter-wave spectroscopic remote-sensing and an advanced information processing technology is crucial. By combining the open-source radiative transfer code Planetary Spectrum Generator (PSG) and the Python libraries SciPy and astropy, we have constructed the molecular profile derivation tool Planetary Atmospheric Spectrum CALculator (PASCAL), which focuses on analysis of data of Titan observed by the Atacama Large Millimeter/submillimeter Array (ALMA). PASCAL successfully retrieved the vertical profile of an isotopomer of a gaseous hydrogen cyanide ($H^{13}CN$) $J = 3$–2 rotational transition, where J is the rotational constant, from ALMA archival data. Further, it successfully determined that the northern high-latitude region has 27% higher column density than the south as seen in the intensity map. In addition, an installed parallelization method succeeded in decreasing the calculation time with its parallel speedup value ∼20 using 32 processors applying coordinate parallelization.

Keywords: ALMA · Planetary atmosphere · Radiative transfer · Parallel processing

1 Introduction

For the observational study of planetary atmospheres, high-dispersion spectroscopy gives us important information, such as the atmospheric structure, composition and dynamics. In particular, because the line-shape of the observed spectrum varies with respect to atmospheric pressure, complete fitting of a modeled spectrum to the observed spectrum reveals the vertical distribution of trace species. A numerical technique to derive the vertical structure of molecular abundance or temperature structure from the observed spectrum is called a retrieval method. The method consists of two different numerical techniques, a radiative

© Springer Nature Switzerland AG 2021
O. Gervasi et al. (Eds.): ICCSA 2021, LNCS 12953, pp. 234–244, 2021.
https://doi.org/10.1007/978-3-030-86976-2_16

transfer technique which produces modeled spectrum, and an optimizing method to derive the vertical physical properties convolved in the spectrum. Both techniques are very complicated because they include and consider many quantum physical processes, three-dimensional path of ray, and numerical difficulties. Each research group develops its own retrieval or at least the radiative transfer codes for their target, purpose and wavelength. For example, Non-Linear Optimal Estimator for Multivariate Spectral Analysis (NEMESIS) was developed to reduce the data observed by the *Cassini* spacecraft and is currently used mainly by NASA [6]. In addition, the Atmospheric Radiative Transfer Simulator (ARTS) is widely used in the remote sensing research [1] of both Earth and planetary atmospheres.

Recently, many astronomical observatories have made their observation data available as an archive to maximize its scientific output. To drive planetary atmospheric studies, reducing the burden of the retrieval methods is crucial for many researchers to analyze their own and archived observation data. Among such observation archive, we have focused on the archived data of the Atacama Large Millimeter-submillimeter Array (ALMA) for planetary atmospheric science. As we have reported already [4], ALMA has kept observed solar system objects as the standard absolute flux calibrator. Thus, ALMA stores a vast amount of planetary observation data. In particular, calibrated data of Titan, the largest moon of Saturn, have drawn attention. Since Titan has a thick atmosphere that contains many volatiles, understanding its atmospheric chemistry processes is one of the key science themes in our solar system planetary atmospheric science. For example, researchers have recently detected complicated molecular species, such as C_2H_5CN [3] and C_2H_3CN [8], isotopologues of nitriles (e.g. [2,5,11]), and time variation of minor species [13]. The comprehensive use of ALMA archival data is a key method for the further understanding of Titan's complicated atmospheric chemistry. For that purpose, we have developed a semi-automatic method for converting calibration data into scientific data and produced more than 2000 three-dimensional Flexible Image Transport System(FITS) datasets that store spectral data in each two-dimensional coordinate [4]. FITS is a standard data archive format widely used for astronomical data science developed in 1970s [9]. For the analysis of the data, development of a radiative transfer tool focusing on ALMA observation is needed. The recent improvement of open-source software has the potential to enhance astronomical research activity. The comprehensive use of such software reduces the burden of software development and enables us to evaluate the analyzed results using them. In this paper, we report the construction and evaluation results of our own retrieval code Planetary Atmospheric Spectrum CALculator (PASCAL) which utilizes open-source software for radiative transfer and optimization.

2 Scheme and Experiment

2.1 Overview

PASCAL mainly consists of the open-source radiative transfer code Planetary Spectrum Generator (PSG) [14], an astronomical Python code library astropy [10]

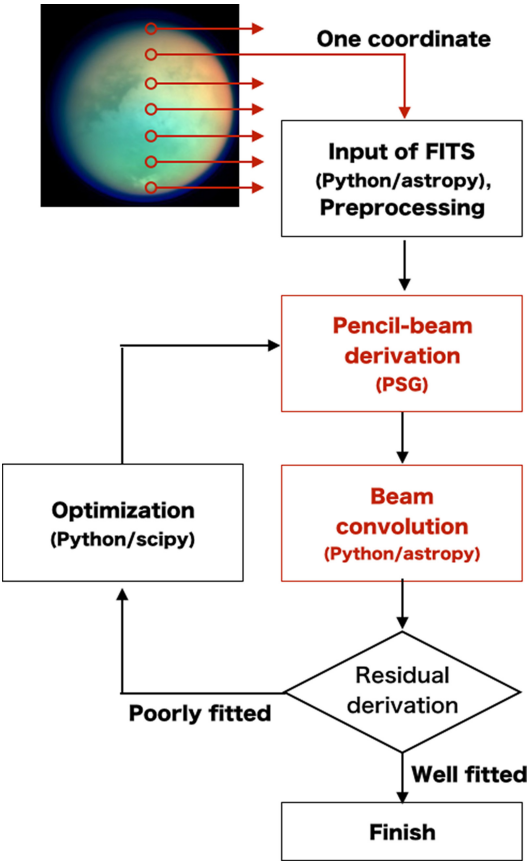

Fig. 1. A flow-chart of PASCAL to derive the vertical molecular abundance at one coordinate. Red symbols are parallelized processes (Color figure online).

and an optimizing method function included in SciPy. Figure 1 shows the flow chart of PASCAL to retrieve a molecular profile of one specific location on Titan. To reduce the calculation time, a parallelizing method is used for both radiative transfer and entire process. To obtain the spatial variation of the molecular profile, one needs to run the program to consider each coordinates.

PSG was developed by the NASA/Jet Propulsion Laboratory. PSG can consider a wide variety of planetary atmospheric observations, from in-situ satellite-based measurement to ground- and space-based observations. Frequency coverage is wide enough to consider wavelengths from UV to radio. Atmospheric structures of planets, major satellites, and minor bodies are implemented already.

PSG can be ran locally by installing it as the Docker container system. To run PSG, one has to prepare a config text file which describes the instrument setting and atmospheric structure such as pressure/altitude, temperature and composition. Besides, users can install packages developed for the specific use. In this study, a package which contains a molecular line database is installed. Once the PSG is installed, one can run the software via Application Program Interface (API) using curl command. After the calculation, PSG returns the result in the form of two column text that has channel and intensity field. An example of returned result is shown in Fig. 2. PSG can be run online by calling the address of NASA/JPL, which is easy to use but the calculation time is much longer than installed version.

For the beam convolution, the Python library, astropy, is used. Since the spatial resolution of ALMA is not infinite, observed area includes various optical paths. Thus, using PSG, PASCAL produces many spectra considering various emission angle as shown in Fig. 2, and convolves them into one spectrum by considering the shape of the beam. Because ALMA's beam typically has elliptical Gaussian shape, the contribution of each pencil beam has to be taken into account using astropy. The astropy is also used for FITS file handling.

To obtain the vertical distribution of the molecular gas, so called the retrieval analysis, we used SciPy package, a Python based package specified for the scientific research of mathematics, physics and engineering. SciPy consists of many libraries for scientific computing. Among them, we used a non-linear least-squares library for the retrieval analysis.[1] The library is a wrapper around MINPACK, a FORTRAN library for the least-squares method. Using the library, one can reproduce the vertical abundances of temperature and molecular abundance profiles by producing the molecular line spectrum and comparing it with the observed one.

The specifications of the computer used for the experiment are summarized in Table 1.

Table 1. Specifications of computer and storage used for test calculation

Item	Spec
CPU	Xeon Gold 6230 (20 cores, 40 threads, maximum frequency: 2.1 GHz) ×2
RAM	ECC Registered DDR4 2666 MHz 192 GB (16 GB × 12)
System disk	1.0 TB NVMe SSD
Data disk	214 TB RAID50 SAS
OS	Ubuntu Linux 16.04 LTS

[1] https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.leastsq.html.

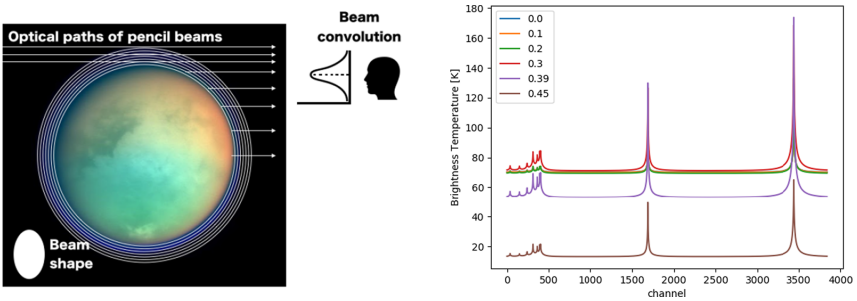

Fig. 2. Left: A model of pencil beam calculation. Pencil beams are convolved with respect to the coordinate of beam center and beam shape indicated in the bottom left. Right: Examples of pencil beam spectra calculated with respect to the distance from the disk center of Titan.

2.2 Radiative Transfer and Sample Data

PSG produces pencil beam with respect to the location on disk, frequency, temperature profile and molecular vertical abundances. A model and examples of pencil beam calculation is shown in Fig. 2. Along with the variation of optical path, the number of pencil beam produced in the code for nadir and limb viewing were 20 and 30, respectively, which correspond to ∼0.01 and ∼0.005 arcseconds spatial resolution, which is 10 times finer than the typical spatial resolution of ALMA, 0.2–0.4 arcseconds. For the test of Radiative transfer and retrieval, a strong rotational transition of the isotopologue of hydrogen cyanide ($H^{13}CN$) J = 3–2 was used. The right panel of Fig. 2 shows some examples of pencil beams for the range of distance from the apparent disk center expressed in arcseconds. The beams for 0.0 to 0.3 arcseconds are for nadir direction, 0.39 and 0.45 are for limb direction. The continuum is produced via Collision Induced Absorption processes, which has an extremely broad line-shape. Variation of the line-shape is due to the variation of optical path length with respect to the emission angles. Typically, line intensity gets increased when the path length gets longer because of increase of molecular column density.

The archived observation data of 2012.1.00453.S was used for the experiment. The observation date and an apparent diameter of Titan was July 7, 2014 and 0.78 arcseconds, respectively. An observed spectrum is shown in the left panel of Fig. 3 with a blue curve. A coordinate of the beam is the center of the disk. Lines seen in this spectrum are CH_3CN, $HC^{15}N$ and $H^{13}CN$ from left to right, respectively. The red curve over-plotted on observation is a retrieved spectrum. A red curve shown in the bottom of the panel is residual between model and observation. Vertical profile of both CH_3CN and $HC^{15}N$ were not retrieved, thus the large residual is seen under the two emission lines. A right panel of Fig. 3

is an integrated intensity map of the data shown in arbitrary unit. As clearly seen, $H^{13}CN$ emission is concentrated on the northern hemisphere due to the seasonal change. In this study, it is expected that the spatial variation of $H^{13}CN$ is obtained by the retrieval analysis. Three white cross marks are the coordinates used for the retrieval experiments described in next subsection.

The radiative transfer calculation was performed by PSG installed on a Windows 10 virtual machine set up with VMware workstation virtual machine software. The code transfers a dataset that describes atmospheric data, the molecular abundance profile, spectroscopic settings and so on to PSG. PSG returns the calculated spectrum as a standard text, which is saved to a RAM disk that has an extremely high disk I/O.

On Titan, atmospheric temperature profile is known to exhibit both spatial and time variation. Thus, the applied temperature profile is chosen from the previously obtained ones [12]. The modeled atmosphere is divided into 30 horizontal layers from the surface to the top of the atmosphere where the pressure level is 10^{-5} mbar.

2.3 Retrieval Analysis

The obtained residual between model and observation is transferred to SciPy for the retrieval. The vertical profile of the molecular abundance is controlled automatically to fit the modeled spectrum with the observed one. As an initial profile, a so-called "a priori", previously determined one is used. In this case, the profile obtained by [7] was applied. First, to coarsely determine the profile, the entire a priori profile is simply scaled to fit the spectrum. Then, the volume mixing ratio of some of the layers are modified for the fine profile determination. Values of the rest of the layers are interpolated. Fitting is performed to minimize the residual. The frequency range to measure the residual and the chi-square value is ± 100 MHz for $H^{13}CN$. Typically, the number of retrieval iterations needed to converge was up to $\sim$100 for each Titan's coordinate. A red curve in the left panel of Fig. 3 show the best-fit spectrum of disk center position. The fitting succeeded in decreasing the residual within the observed random noise level. A left panel of Fig. 4 shows a retrieved result of three coordinates. Northern hemisphere exhibits higher mixing ratio than other two coordinates other than at low altitude. The derived column densities of three coordinates converted to the main isotopologue are 4.2, 2.9 and 3.3×10^{17} molecules cm^{-2} for N, C and S coordinates, respectively.

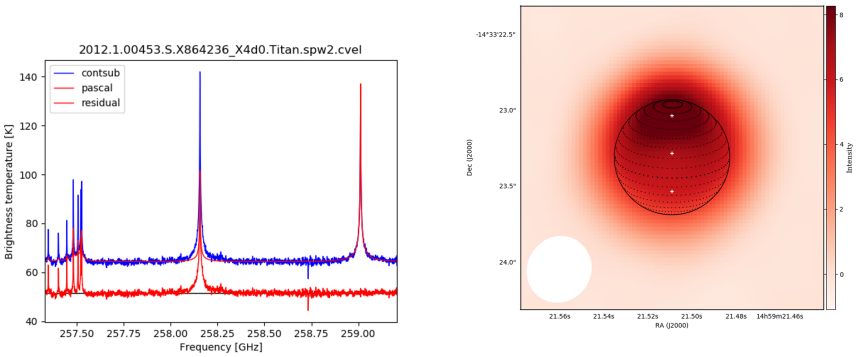

Fig. 3. Left: Observed (blue) and retrieved (red) model spectra of Titan. The residual is also plotted at the bottom of the panel. For the retrieval, the $H^{13}CN$ ($J = 3$–2) rotational transition locates at 259 GHz which is present at right side of the panel was used. Emission lines at the left edge of the panel is $J = 14$–13 transitions of CH_3CN molecule. Right: An integrated intensity emission map of $H^{13}CN$. A white ellipse at bottom left and cross markers are the synthesized beam and Titan's coordinates used for the retrieval calculation. (Color figure online)

The right panel of Fig. 4 shows the variation of chi-square values with respect to each retrieval iteration. In this case, the first 15 iterations were for coarse scaling and remaining iterations were for fine retrieval, respectively. The derived vertical profiles of three coordinates are shown in Fig. 4.

2.4 Parallelization

Since the Titan's atmospheric trace gases are known to exhibit a time variation of its vertical profile, number of ALMA data should be used for retrieval to reveal the time variation of three-dimensional distribution of each gases. Without parallelization, each iteration takes ∼30 s, thus, the total derivation time on the central meridian line with 8 points and 100 iterations is 24000 s. Thus, implementation of a parallelizing method to reduce the calculation time is crucial for its scientific application.

The entire code for one coordinate, including radiative transfer, was parallelized using the `multiprocessing` method of Python. Since the method does not use node-to-node communication, the number of used processes is limited to the maximum number of threads of the computer. In this study, the radiative transfer parallelization calculated pencil beams in parallel by running PSG simultaneously. In turn, coordinate parallelization ran the entire code in parallel to retrieve the vertical profile of different coordinates on Titan simultaneously. While the beam convolution process was parallelized also, it could not decrease

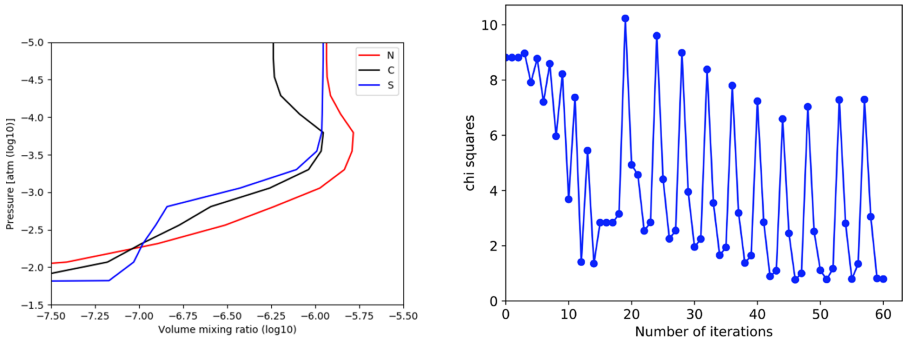

Fig. 4. Left: Retrieved vertical profile of $H^{13}CN$ at Titan's three different locations. N, C and S are for north, disk center and southern hemisphere, respectively. Right: Variation of chi-square values for the residual between modeled and observed spectra of Titan with respect to the number of iterations.

the calculation time. Effectiveness of the parallelization was estimated by calculating the parallel speedup $s(p)$, which can be obtained by the following equation:

$$s(p) = \frac{t(1)}{t(p)} \tag{1}$$

where $t(p)$ is time for one radiative transfer or vertical profile determination with p processors. Ideally, without any overhead time, the parallel speedup is equal to the number of processors. For the radiative transfer parallelization, note that the number of processors is limited to 16 because of the limitation of the virtual machine. In turn, for the coordinate parallelization, we measured the time needed for the entire process of vertical profile derivation. Figure 5 shows the parallel speedup of both parallelization tests. The tested numbers of processors were 1, 2, 4, 8 and 16 for radiative transfer and 1, 2, 4, 8, 16, 32 and 64 for coordinate parallelization. As the error bars, standard deviation of each iteration time or entire processing time is shown.

For the radiative transfer parallelization, $s(p)$ did not improve for $p > 8$. Because the number of pencil beams calculated in PASCAL is ∼30, improvement of parallel speedup for radiative transfer is still needed. Since each radiative transfer calculation does not need storage I/O and communication between processes, parallel speedup should be more efficient. In this study, we installed the PSG Docker container on the virtual machine. The indirect call of PSG may be the cause of the overhead of parallel speedup.

Coordinate parallelization exhibited a better speedup in the high-p region. Because the number of physical cores is 40, $s(p)$ does not improve above $p = 40$, although the number of threads is 80. For the coordinate parallelization, low speedup efficiency may be caused by the Hyper Threading function, or the storage I/O to read a large FITS file that has 1.5 GB in size. From a scientific point of view, calculations for the derivation of the latitudinal variation of a

molecular profile should be performed only on central meridian line. Thus, the number of coordinates on Titan is no more than ∼10, which is equivalent to 0.08 arcseconds spatial step, considering that the finest spatial resolution of ALMA at Band 6 or 7 as 0.1 arcseconds. Thus, the combination of both parallelization methods is useful considering the scientific purpose and benefit of parallelization.

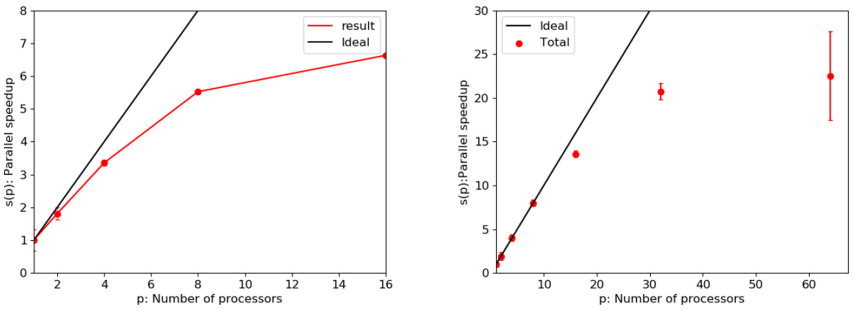

Fig. 5. Parallel speedup measured for the radiative transfer and vertical profile measured at multiple coordinates. Coordinate parallelization shows a higher speedup than the radiative transfer.

3 Summary and Future Prospects

In this study, we successfully retrieved the vertical structure of the isotopologue of hydrogen cyanide at different coordinates of Titan. The obtained result is consistent with the distribution of the integrated intensity, and it demonstrated that the PASCAL code is useful for the automated derivation of three-dimensional molecular abundances on Titan.

To use the results of PASCAL, solution of the following problems is needed:

- Application and testing for other molecules are insufficient. For example, for the HC_3N molecule, SciPy could not converge after many iterations to decrease the residual within the standard deviation of the spectrum.
- The error bar produced by SciPy retrieval is much smaller than that of retrieval results. For example, the error for $H^{13}CN$ derivation was as low as several percent for the entire altitude. However, previous submillimeter observations of HCN reported an error of ∼50% at high altitude [13]. Evaluation and new implement of the errors are needed.

Acknowledgements. TI is grateful to Mitaro Namiki of Tokyo University of Agriculture and Technology for his advices on the entire study. Also, TI is also grateful to Hideo Sagawa for discussions about the improvement of retrieval program and usage of ALMA data. Authors are grateful to Geronimo Villanueva, a developer of PSG, for his kind support for the usage of PSG. This work was financially supported by a Telecommunications Advancement Foundation, JSPS Kakenhi grants (17K14420, 19K14782) and an

Astrobiology Center Program of National Institutes of Natural Sciences (NINS) grants. This paper makes use of the following ALMA data: ADS/JAO.ALMA#2012.1.00453.S. ALMA is a partnership of ESO (representing its member states), NSF (USA) and NINS (Japan), together with NRC (Canada), MOST and ASIAA (Taiwan), and KASI (Republic of Korea), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, AUI/NRAO and NAOJ.

References

1. Buehler, S.A., Mendrok, J., Eriksson, P., Perrin, A., Larsson, R., Lemke, O.: ARTS, the atmospheric radiative transfer simulator - version 2.2, the planetary toolbox edition. Geosci. Model Dev. **11**(4), 1537–1556 (2018). https://doi.org/10.5194/gmd-11-1537-2018
2. Cordiner, M.A., et al.: Interferometric imaging of Titan's HC_3N, $H^{13}CCCN$, and $HCCC^{15}N$. Astrophys. J. **859**(1), L15 (2018). https://doi.org/10.3847/2041-8213/aac38d, http://stacks.iop.org/2041-8205/859/i=1/a=L15?key=crossref.3db3b0a1e328504fc7cc09382363bbb3
3. Cordiner, M.A., et al.: Ethyl Cyanide on Titan: spectroscopic detection and mapping using ALMA. Astrophys. J. **800**(1), L14 (2015). https://doi.org/10.1088/2041-8205/800/1/L14, http://stacks.iop.org/2041-8205/800/i=1/a=L14?key=crossref.d304cda239aa0a7fa1c22fbf25847529
4. Iino, T., Namiki, M., Yamada, T.: A feasibility study of exhaustive analysis of ALMA calibration data for the creation of big-data driven solar system astronomy. JAXA Res. Dev. Rep. J. Space Sci. Inform. Jpn. **7**(1), 19–32 (2018). https://doi.org/10.20637/JAXA-RR-17-009/0003
5. Iino, T., Sagawa, H., Tsukagoshi, T.: $^{14}N/^{15}N$ isotopic ratio in CH_3CN of Titan's atmosphere measured with ALMA. Astrophys. J. **890**(2), 95 (2020). https://doi.org/10.3847/1538-4357/ab66b0, https://iopscience.iop.org/article/10.3847/1538-4357/ab66b0
6. Irwin, P.G.J., et al.: The NEMESIS planetary atmosphere radiative transfer and retrieval tool. J. Quant. Spectrosc. Radiat. Transf. **109**(6), 1136–1150 (2008). https://doi.org/10.1016/j.jqsrt.2007.11.006
7. Marten, A., Hidayat, T., Biraud, Y., Moreno, R.: New millimeter heterodyne observations of Titan: vertical distributions of nitriles HCN, HC_3N, CH_3CN, and the isotopic ratio $^{15}N/^{14}N$ in its atmosphere. Icarus **158**(2), 532–544 (2002). https://doi.org/10.1006/icar.2002.6897, http://linkinghub.elsevier.com/retrieve/pii/S0019103502968971
8. Palmer, M.Y., et al.: ALMA detection and astrobiological potential of vinyl cyanide on Titan. Sci. Adv. **3**(7), 1–7 (2017). https://doi.org/10.1126/sciadv.1700022
9. Pence, W.D., Chiappetti, L., Page, C.G., Shaw, R.A., Stobie, E.: Definition of the Flexible Image Transport System (FITS), version 3.0. Astron. Astrophys. **524**(10), A42 (2010). https://doi.org/10.1051/0004-6361/201015362
10. Robitaille, T.P., et al.: Astropy: a community Python package for astronomy. Astron. Astrophys. **558**, A33 (2013). https://doi.org/10.1051/0004-6361/201322068
11. Serigano, J., et al.: Isotopic ratios of carbon and oxygen in Titan's co using ALMA. Astrophys. J. **821**(1), L8 (2016). https://doi.org/10.3847/2041-8205/821/1/L8, http://arxiv.org/abs/1602.07707, http://stacks.iop.org/2041-8205/821/i=1/a=L8?key=crossref.65db31dc4452e5223c4a0886fb6e5050

12. Thelen, A.E., et al.: Spatial variations in Titan's atmospheric temperature: ALMA and Cassini comparisons from 2012 to 2015. Icarus **307**, 380–390 (2018). https://doi.org/10.1016/j.icarus.2017.10.042
13. Thelen, A.E., et al.: Abundance measurements of Titan's stratospheric HCN, HC_3N, C_3H_4, and CH_3CN from ALMA observations. Icarus **319**(June 2018), 417–432 (2019). https://doi.org/10.1016/j.icarus.2018.09.023, https://linkinghub.elsevier.com/retrieve/pii/S0019103518304184
14. Villanueva, G.L., Smith, M.D., Protopapa, S., Faggi, S., Mandell, A.M.: Planetary spectrum generator: an accurate online radiative transfer suite for atmospheres, comets, small bodies and exoplanets. J. Quant. Spectrosc. Radiat. Transf. **217**, 86–104 (2018). https://doi.org/10.1016/j.jqsrt.2018.05.023