

### Research Article

Gwenaël Van Looveren\*, Olga Kochina, and Dmitry Wiebe

# PRESTALINE: a package for simulation and analysis of molecular spectra of star forming regions

https://doi.org/10.1515/astro-2021-0012 Received Oct 14, 2021; accepted Nov 16, 2021

**Abstract:** We present the PRESTALINE package, a novel tool to simulate and analyse spectra of star forming objects. PRESTALINE allows for a direct comparison of theoretical models with observations and simplifies an analysis of observed spectra. This allows researchers to estimate physical conditions in a studied object and to interpret its chemical composition in a quantitative way. The goal of the project is to set up a consistent framework which would bridge a gap between theoretical studies and observations of star forming regions. In this paper, we present the results of applying PRESTALINE to the test object DR21(OH) and discuss the project general possibilities.

Keywords: astrochemistry, star formation, interstellar medium

# 1 Introduction

Star-forming regions are tricky objects to study. Molecular hydrogen, the main component of the regions, does not emit in conditions of cold molecular clouds, which warrants a usage of other tracers to get the necessary information, with admixture molecules being the most prominent candidates.

Molecules in that regard have many positive features: they emit multiple lines and provide information on kinematics and the temperature of the gas. Spectral surveys allow us to conduct studies of large numbers of lines of different molecules. Unfortunately, they are not uniformly mixed with the gas so additional efforts are required to make them a reliable diagnostic tool. That is where theoretical astrochemistry, building chemical evolution models, which connect chemical composition and characteristics of the source, comes into play.

In order to apply theoretical models to observations quantities for comparison are needed, identical for both theoretical studies and observations. Most often column density is used, as it is the direct output of the theoretical model and can be computed for a large number of molecules. But

deriving column densities from observations requires certain assumptions that may seriously compromise accuracy in some cases. The better approach is to directly compare line intensities of the real spectrum with the synthetic ones. Line intensities can be obtained from observations directly and are derived from chemical models by radiative transfer calculations. The most significant shortcoming of that approach is that it is limited by available molecular data (term energies, statistical weights, etc.).

The package PRESTALINE presented here is able to provide the comparison in terms of both intensities and column densities. The ability of PRESTALINE to create synthetic spectra allows for a more direct comparison of theoretical models with observations. This way, the simplifications needed to calculate the column densities from observations are avoided. The goal of this project is to set up a consistent framework which would allow evaluating different models of star forming regions. Yet another possible application of this tool would be to create an atlas of spectra for a grid of objects with given parameters, to facilitate estimating specifics of studied objects. PRESTALINE does not only model synthetic spectra, but allows analysing the observed ones. As a separate branch of tools, PRESTALINE is able to estimate the mean temperature of the object out of spectral lines and to use this information to determine approximate distributions of the species in the object.

Corresponding Author: Gwenaël Van Looveren: University of Vienna; Email: gwenael.van.looveren@univie.ac.at

Second Author's name: Olga Kochina Institute of Astronomy, Russian Academy of Sciences

Third Author's name: Dmitry Wiebe Lebedev Physical Institute, Samara 443011, Russia; Institute of Astronomy, Russian Academy of Sciences

# 2 Package description

The basic PRESTALINE package consists out of 3 main components. The first part is PRESTA, a chemical evolution code. The second component uses RADEX, a radiative transfer code, to simulate the flux of the modelled cloud. In the visualisation part, the code creates a synthetic spectrum, which is confronted with the observed spectrum. Additional tools devoted to estimations of physical and chemical parameters of an obtained spectrum are discussed in section 4.

### 2.1 PRESTA

Calculations of the chemical evolution are performed with PRESTA, a multi-point model of the chemical evolution of prestellar and protostellar objects, which is being developed at the Institute of Astronomy of the Russian Academy of Sciences (Kochina et al. 2013). Its current version enables modeling of the collapse of a spherically symmetric core, irradiated by both diffuse external radiation and a central source (protostar) with specified parameters. The studied region is divided into concentric cells, and in each cell the dust temperature, gas temperature, density, and extinction for external and internal radiation are specified. The distributions of the temperature, density, and radiation field in a spherically symmetric approximation, are calculated using the radiative transfer model or set by hand. The physical parameters are taken from literature for the studied region or specified from some other considerations. The possibility to take into account the warming up phase and evolution of the density structure are present in the model as well (Kochina & Wiebe 2015).

The model takes into account cosmic-ray reactions, two-body reactions, photoreactions (including cosmic-ray induced photoreactions), dissociative recombination reactions with electrons and charged and neutral dust grains and reactions with the dust particles and surface chemistry. Chemical equations are integrated in each cell up to a specified age. PRESTA works with different initial elemental chemical composition sets for both low and high metallicities (Kochina & Wiebe 2017). The rates of gas-phase reaction can be calculated using either the UMIST database or the chemical network from the work by Albertsson et al. (2014) to include deuterated species.

The model may contain several dust populations. This is implemented technically by creating analogs of the surface components for each dust population. Surface reactions on grains, accretion of gas-phase components onto

grains, photodesorption, thermal desorption, and cosmicray induced desorption were all taken into account. The rates of surface reactions, as well as the accretion and desorption rates, were calculated separately for each dust population at each point of the cloud, taking into account the grain temperatures and mass fractions. The temperatures of the dust populations are computed using a radiativetransfer model taking into account stochastic heating of grains with various sizes and chemical compositions by the external and internal radiation fields (Pavlyuchenkov et al. 2014).

In order to understand the chemical pathways and determine the main factors influencing the chemical evolution, an analysing tool for chemical rates is used (Kochina & Wiebe 2014). The analyser uses the network of reactions and current abundances of species to determine the most productive and destructive reactions for selected species for any required moment in time and distance from the centre of the cloud. This way it provides sufficient information for the understanding of the main factors influencing the chemical evolution of the object. On the output, PRESTA provides information on both relative abundances and column densities of every species for any moment.

As a part of the PRESTALINE package, PRESTA creates files containing all the information needed by RADEX for further calculations. These include kinetic temperature (the average for the region), density of collision partners (currently only H<sub>2</sub> is considered), molecular column density of the considered species. Being a one-dimensional model and taking into account density and temperature distribution, it is more accurate to divide the object into three parts with different average physical parameters (density and temperature) and provide three different input files for RADEX to be calculated separately. These are the inner region (hot and dense), the medium region (much cooler and dense) and the outer region (cold and less dense). Information on averaged temperatures, densities of the collision partner and column densities of the species of interest for the regions are provided as separate RADEX input files for the same output moment in PRESTA. As PRESTA does not distinguishes ortho- and para- modifications of species, the RADEX input of the column densities is manually divided and considered to be 75% for para- and 25% for ortho-.

### 2.2 RADEX

The column densities provided by PRESTA are passed on to the radiative transfer code RADEX, developed by van der Tak et al. (2007). Level populations are calculated through statistical equilibrium, which holds up in the low density conditions of molecular clouds. Statistical equilibrium, however, relies on the knowledge of collisional data. Currently PRESTALINE uses 50 species, including some isotopologues, from the LAMDA database (van der Tak *et al.* 2020). The selection of collisional data can strongly influence the final result of the simulation. Section 3 will look at impact of including hyperfine structure in the collisional data. For each of the species the intensities and optical depths of their transitions are calculated. The temperature of background radiation is taken to be the cosmic microwave background radiation, and the line width is assumed to be 0.1 km/s. The results are passed to the third part of the code.

### 2.3 Visualisation

The visualisation module of PRESTALINE combines all radiative transfer data created by RADEX into a single synthetic spectrum. As mentioned before, PRESTA produces a three layered model of the cloud. This accounts for variations of physical parameters and the fact that an outer layer can obscure a deeper layer. These layers are combined by comparing each transition to the optical depth of more outward layers. If in a specific wavelength range the optical depth is too large, the lower layers do not contribute to the emergent spectrum.

The selected line is then modelled by a Gaussian, as Doppler broadening is assumed to be the primary form of line broadening. The integrated intensity of this line corresponds to the flux calculated by RADEX. To ensure that each line is rendered properly, PRESTALINE models a section around the central frequency based on the line width. These sections are then combined to form a full spectrum. At frequencies where lines overlap, their flux is combined to handle any line blending.

The combined spectrum is saved as a text file to easily create and manipulate graphical representations. Additionally PRESTALINE keeps track of each line and the corresponding levels, which allows for easy identification of the different spectrum sections.

In order to visually assess the quality of the modelled spectrum, PRESTALINE calculates an upper and lower boundary. The upper boundary is represented by multiplying the column densities by 5 and the lower boundary is obtained by dividing the column densities by 5.

The quality of a model compared to an observation, needs to be quantified in order to select the best model. To this end, PRESTALINE calculates the mean absolute error (MAE) between the model and the observation. The modelled spectrum is made in such a way that a resolution is

high around lines, but there are no points in the rest of the spectrum. This allows the error to focus only on the modelled lines and does not add any penalty for areas for which no molecular data files are available.

# 3 DR21(OH): modelling the synthetic spectrum

As a test object to which PRESTALINE is applied DR21(OH) was chosen. DR21(OH) is the region of massive star formation located in a giant molecular cloud in the Cygnus-X complex at a distance of  $\sim$ 2 kpc from the Sun. Evidence of active star formation in DR21(OH)M is provided by the existence of powerful OH, H<sub>2</sub>O, and CH<sub>3</sub>OH masers (Norris *et al.* 1982; Genzel & Downes 1977; Batrla & Menten 1988, see). The absence of compact HII regions indicates that the stages of evolution of the protostars in the region are very early. PRESTA has already been used to model the object successfully in Kochina *et al.* (2013), which determined the choice. The spectrum (Kalenskii & Johansson 2010) of DR21(OH) was kindly provided by Sergei Kalenskii. Same physical structure as in Kochina *et al.* (2013), but an upgraded chemical network, were applied for modelling.

The synthetic spectrum covers a frequency range from 83 GHz to 116 GHz. Due to this large range, a single figure would be illegible. A first section of interest to focus on is line blending. Figure 1 shows how overlapping lines are combined into a smooth spectrum.

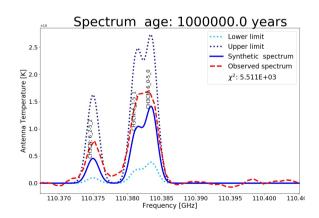


Figure 1. Example of multiple lines blending together.

A particular case of line blending is represented by transitions with hyperfine splitting. This effect can be taken into account if the corresponding collisional data are available. An example of such a line is the  $HC_3N$  transition at 109.17

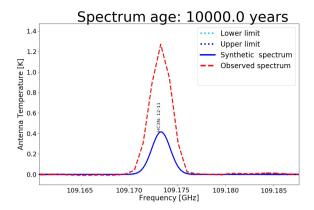


Figure 2. Example of a transition without hyperfine structure. Note: at this scale the upper and lower limit coincide with the synthetic spectrum.

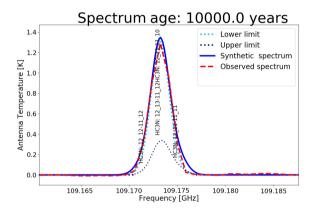


Figure 3. Example of a transition with hyperfine structure. Note: due to the large column density of the upper limit, the line becomes optically thick, blocking its own signal.

GHz. Figure 2 shows this line when the hyperfine structure is not present. Modelling the same transition taking into account the hyperfine structure, results in Figure 3. Even though the individual transitions are not resolved, it is clear that the blended line fits the observed spectrum much better.

# 4 Parameter estimator

Star forming regions come in a wide variety of evolutional stages, each with their own set of physical parameters. For this reason, in addition to creating synthetic spectra, PRE-STALINE includes tools to estimate parameters from an observed spectrum. These tools can be used on their own to analyse observations or in combination with the package described in section 2. In the latter case, estimating the parameters of a real object can reduce the number of possible input parameters, when trying to recreate the spectrum.

# 4.1 Temperature

The first parameter is the temperature of the molecular cloud. The rotation temperature can be derived using a rotation diagram, which is described in detail in Kalenskii & Johansson (2010). This method makes use of optically thin lines, to relate the intensity of a transition to its energy through following relation (equation 3 from Kalenskii & Johansson (2010)):

$$\log\left(\frac{N_u}{g_u}\right) = \log\left(\frac{3kW}{8\pi^3\nu_0 S\mu^2 g_I g_k}\right)$$

$$= \log\left(\frac{N}{Q_{\text{rot}}}\right) - \log\left(e\left(\frac{E_u}{kT_{\text{rot}}}\right)\right).$$
 (2)

$$= \log \left( \frac{N}{Q_{\text{rot}}} \right) - \log \left( e \left( \frac{E_u}{kT_{\text{rot}}} \right) \right).$$
 (2)

In this equation, the terms related to the transition intensity are represented by the right side of equation 1. Here k is the Boltzmann constant,  $v_0$  the central frequency, S the line strength,  $\mu$  the permanent dipole moment of the molecule and  $g_I$  and  $g_k$  the spin and K degeneracy, respectively. The term W is the integrated intensity of the line, which can be derived from observations. Equation 2 is a linear function where the intercept is defined by the ratio of the column density N, and the rotational partition function  $Q_{\text{rot}}$ . The upper energy level of each transition is given by  $E_u$ . Once this energy is divided by the Boltzmann constant, we can determine the slope described by these points, which is related to the inverse of the temperature.

In order to use this method, a molecule needs several spectral lines close to each other. Due to the assumptions of the equations, these lines also need to be optically thin and have a high enough excitation temperature. For this purpose, PRESTALINE uses CH<sub>3</sub>CCH and CH<sub>3</sub>CN. Section 5 takes a closer look at the application of the method to an observation.

# 4.2 Column density

The second parameter we can estimate is the column density. This process shares many similarities with the second and third subsections described in section 2. Transitions are calculated for a list of included species. For this, RADEX will need a temperature, a column density and the density of the collision partner, here taken to be H<sub>2</sub>. These parameters can vary strongly between the core regions of the molecular cloud compared to the outer regions. Just as the model created by PRESTA, the column density estimator also uses a 3 layer model. The inner region uses the upper limit of the temperature range estimated at the previous step (see previous subsection), a middle region uses the average temperature and finally an outer layer is modeled with the lower limit of the temperature estimate.

To find the best set of column densities, we perform a random search over the full range of column densities which RADEX can handle ( $10^5 - 10^{25} \ cm^{-2}$ ). A random search is used rather than a gird search due to the high number of combinations a 3 dimensional parameters space brings along. Bergstra & Bengio (2012) showed that random search preforms at least as good as a grid search, whilst increasing efficiency by not giving each dimension the same importance. For each combination the Mean Absolute Error (MAE) is calculated. We opt for the MAE rather than the Mean Squared Error (MSE), in order to give both weak and strong lines similar weights as the MSE is much more sensitive to outliers. The smallest error is considered the best fit for this molecule. This processes is repeated for each species in the spectrum.

# 5 DR21(OH): parameters estimated

The first step is to determine the temperature for the observed spectrum. For this we used the rotation diagrams of  $CH_3CCH$  and  $CH_3CN$ . Figure 4 shows the diagram for  $CH_3CCH$ , from which a temperature of  $34 \pm 12K$  is derived. The rotation diagram of  $CH_3CN$  is shown in Figure 5. The lines of this molecule result in a temperature of  $51 \pm 9K$ . These values coincide with those derived in Kalenskii & Johansson (2010).

With a temperature set, the next step is to determine the column densities of the various species. Table 1 summarises the column densities for the species found in the spectrum. To compare the obtained values with the ones determined by the observers themselves, the column densities estimated via LTE approach in Kalenskii & Johansson (2010) are also presented. In this work, these values have been rounded to an order of magnitude to present conservative results during the testing phase of the tool. The planned addition of a more advanced optimisation algorithm will allow for more detailed results. However, current results are already promising as seen in Figures 6 and 7.

The use of a 3 layered model of the cloud, results in significantly better fits to observations. Figure 6 shows how a 1 layer model is fitted to an observed spectrum with a column density of  $10^{15}~\rm cm^{-2}$ . Figure 7 shows the same lines, but fitted using a 3 layered model, each with a column density of  $10^{14}~\rm cm^{-2}$ . It is immediately clear that the latter is a much closer fit. The difference lies especially in the ratio between the lines. Where a single layer is limited to form all lines from a single temperature and column density, the 3 layered model has greater freedom to vary the intensity of lines individually. This effect is illustrated in Figure 8.

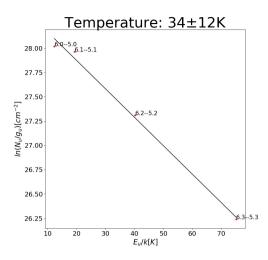


Figure 4. The rotation diagram of CH<sub>3</sub>CCH. Each point represents an energy transition. The derived temperature is 34 ± 12K.

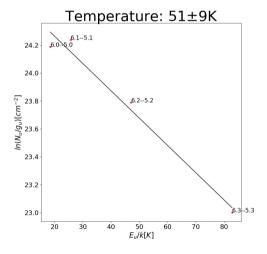


Figure 5. The rotation diagram of  $CH_3CN$ . Each point represents an energy transition. The derived temperature is  $51 \pm 9K$ .

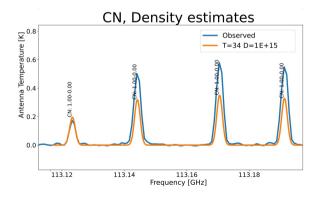


Figure 6. Fitting of CN lines with a single layered model.

Here we can see how the lowest temperature puts greater emphasis on the left most transition, whereas the highest temperature results in a stronger second transition.

An additional note must be made about CO, which is determined to have column densities of  $10^{17}$  cm<sup>-2</sup>,  $10^{16}$  $cm^{-2}$  and  $10^{20}$   $cm^{-2}$ . The modelled line fits the observation very poorly, as can be seen in Figure 9. The strong intensity and width of this transition are caused by outflows (Zapata et al. 2011). This type effect falls outside of the scope of PRESTALINE, so in order to determine the column density of CO properly, a different approach is needed. As no other

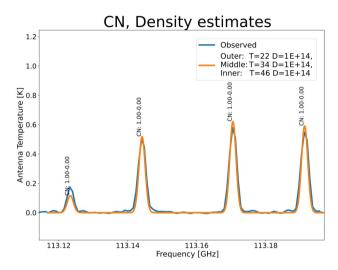


Figure 7. Fitting of CN lines with a 3 layered model.

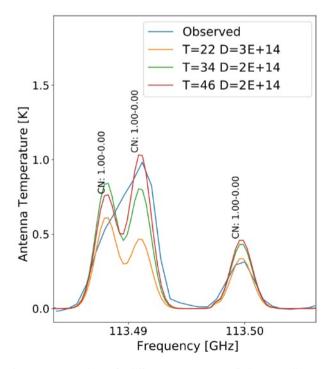


Figure 8. Comparison of 3 different temperatures fitting to CN lines.

transitions of CO fall within the observed spectrum, the best alternative would be to look at isotopologues.

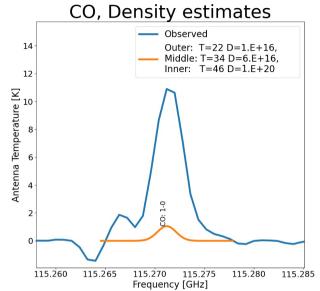


Figure 9. Fitting a modelled line to the CO transition at 115.27 GHz.

Table 1. Overview of the estimated column densities for some species in the outer, middle and inner layer at 22K, 34K and 46K respectively. All column densities are in cm<sup>-2</sup>

element	outer	middle	inner	KJ
A-CH <sub>3</sub> OH	3E+13	4E+11	7E+14	8E+13
$C_2H$	7E+14	7E+14	1E+22	3E+15
CF+	5E+05	1E+05	7E+12	9E+12
СН	3E+17	1E+05	2E+06	N/D
CH <sub>3</sub> CN	3E+13	5E+05	4E+07	N/D
CN	3E+13	2E+12	7E+14	5E+14
CO	1E+16	6E+16	1E+20	3E+16*
CS	1E+14	7E+14	1E+22	4E+14
E-CH <sub>3</sub> OH	2E+21	7E+14	1E+14	N/D
HCN	1E+14	1E+14	3E+17	2E+14
HCO+	1E+14	3E+13	1E+24	1E+14
HCS+	7E+12	2E+12	4E+09	1E+14
HNC	1E+14	1E+14	1E+18	2E+14
HNCO	3E+13	5E+05	5E+05	3E+14
$N_2H^+$	7E+14	3E+13	3E+19	9E+13
ocs	3E+13	2E+12	3E+13	2E+14
SO	1E+14	1E+14	7E+14	7E+13*
$SO_2$	2E+12	3E+13	9E+08	N/D
$SiC_2$	1E+05	5E+05	4E+11	N/D
SiO	4E+11	2E+12	1E+16	2E+13

KJ - column densities determined in Kalenskii & Johansson (2010)

N/D — no data in Kalenskii & Johansson (2010)

<sup>\* -</sup> determined for <sup>18</sup>O isotopologues.

# 6 Conclusions

We present a new tool to simulate and analyse molecular spectra of star forming objects. The package PRESTALINE, based on PRESTA astrochemical model and RADEX radiative transfer code, is a multipurpose instrument that is able to solve a wide range of scientific tasks. In this paper, it has been applied to a well-known object DR21(OH). It is shown that properties of the object derived with PRESTALINE are consistent with previous studies. The package PRESTALINE may also become an indispensable tool for connecting observational results to laboratory studies.

Acknowledgment: The studies at Lebedev Physical Institute were supported by the Ministry of Science and Higher Education of the Russian Federation by the grant No. 075-15-2021-597.

We thank Sergei Kalenskii for valuable support and for the spectral data provided.

**Author contributions:** All authors have accepted responsibility for the entire content of this manuscript and approved its submission.

**Conflict of interest:** The authors state no conflict of interest.

# References

Albertsson T, Semenov D, Henning T. 2014. Chemo-dynamical deuterium fractionation in the early solar nebula: The origin of water on earth and in asteroids and comets. ApJ. 784(1):39.

- Batrla W, Menten KM. 1988. Detection of a strong new maser line of methanol toward DR21(OH). ApJ. 329:L117
- Bergstra J, Bengio Y. 2012. Random search for hyper-parameter optimization. J Mach Learn Res. 13(10):281-305.
- Genzel R, Downes D. 1977. H<sub>2</sub>O in the Galaxy: sites of newly formed OB stars. A&AS. 30:145.
- Kalenskii SV, Johansson LEB, 2010, Spectral scan of the star-forming region DR21(OH). Observations and LTE analysis. Astron Rep. 54(4):295-316.
- Kochina OV, Wiebe DS, Kalenskii SV, Vasyunin AI. 2013. Modeling of the formation of complex molecules in protostellar objects. Astron Rep. 57(11):818-832.
- Kochina OV, Wiebe DS. 2014. Different-sized dust grains and the chemical evolution of protostellar objects. Astron Rep. 58(4):228-239.
- Kochina OV, Wiebe DS. 2015. The possibility of diagnosing evolutionary stages of a protostellar object using observations of complex molecules. Astron Rep. 59(8):762-771.
- Kochina OV, Wiebe DS. 2017. Diagnostics for the elemental composition of protostellar objects. Astron Rep. 61(2):103-114.
- Norris RP, Booth RS, Diamond PJ, Porter ND. 1982. Interferometric observations of four protostellar OH maser sources. Mon Not R Astron Soc. 201(2):191-207.
- Pavlyuchenkov YN, Wiebe DS, Akimkin VV, Khramtsova MS, Henning T. 2012. Stochastic grain heating and mid-infrared emission in protostellar cores. Mon Not R Astron Soc. 421(3):2430-2441.
- van der Tak FFS, Black JH, Schöier FL, Jansen DJ, van Dishoeck EF. 2007.A computer program for fast non-LTE analysis of interstellar line spectra: With diagnostic plots to interpret observed line intensity ratios. Astron Astrophys. 468(2):627-
- van der Tak F. Lique F. Faure A. Black I. van Dishoeck E. 2020. The Leiden Atomic and Molecular Database (LAMDA): Current status, recent updates, and future plans. Atoms. 8(2):15.
- Zapata LA, Loinard L, Su Y-N, Rodríguez LF, Menten KM, Patel N, Galván-Madrid R. 2012. Millimeter multiplicity in DR21(OH): Outflows, molecular cores, and envelopes. ApJ. 744(2):86.