

HII-CHI-mistry: A collection of python scripts for the analysis of emission lines in ionized gaseous nebulae

Enrique Pérez-Montero

Instituto de Astrofísica de Andalucía - CSIC. Apdo. 3004, 15080, Granada, Spain

1. General description

1.1. *What is HII-CHI-MISTRY and where can it be found?*

Ionized gaseous nebulae are ubiquitous objects in the Universe that provide information about the galaxies where they are located by means of their bright emission lines. Among the different properties that we can derive from them are the relative chemical abundances of the observed ions, the excitation of the gas or the effective hardening of the ionizing source. HII-CHI-MISTRY (hereinafter, HCM) is a collection of python scripts that help to analyze the observational information from several bright emission lines observed in the ultraviolet, optical, or infrared ranges of the spectrum, comparing them with the predictions from large grids of photoionization models and providing estimates with their errors for some derived properties, adapting the accuracy of the solutions to the used input lines and their uncertainties.

HCM for the calculation of chemical abundances presents three main advantages in relation to other model-based solutions, including: 1) the results are totally consistent with the so-called direct method, based on the previous determination of the electron temperature. This is true even when the flux of an auroral line, such as $[[\text{O III}]] \lambda 4363 \text{ \AA}$ is not used (see Pérez-Montero, 2014). 2) HCM provides consistent solutions independently of the set of input emission lines and their errors. This is especially useful when we want to compare results for different sets of objects observed with different spectral ranges or redshift. 3) As HCM provides independent solutions for N/O (in the optical and the infrared) or C/O (in the ultraviolet), it allows the use of N or C for the derivation of O/H without any previous assumption about the relation between O/H and N/O or C/O.

In this tutorial I explain the features of the different versions of the code and its instructions of use, along with its limitations and future expected improvements. There are four versions of HCM depending on the spectral range or its use for the determination of chemical abundances or the calculation of the hardening of the ionizing source. These are

- HII-CHI-MISTRY. The original HCM package described in Pérez-Montero, 2014 for the analysis of gaseous nebulae ionized by massive young star clusters and in Pérez-Montero et al., 2019a for Narrow Line Regions (NLR) ionized by active galactic nuclei (AGN), used to derive the total oxygen abundance ($12 + \log(\text{O}/\text{H})$, hereafter O/H), the nitrogen-to-oxygen chemical abundance ratio ($\log(\text{N}/\text{O})$, hereafter N/O), and the ionization parameter ($\log U$) using optical emission lines from $[[\text{O II}]] \lambda 3727 \text{ \AA}$ up to $[[\text{S II}]] \lambda 6717 + 6731 \text{ \AA}$.
- HII-CHI-MISTRY-UV. Described in Pérez-Montero and Amorín, 2017. It calculates O/H, the carbon-to-oxygen chemical abundance ratio ($\log(\text{C}/\text{O})$, hereafter C/O), and $\log U$, using ultraviolet emission lines from $\text{Ly}\alpha \lambda 1216 \text{ \AA}$ up to $\text{C III} \lambda 1909 \text{ \AA}$.
- HII-CHI-MISTRY-IR. Described in Fernández-Ontiveros et al (*submitted*). It calcu-

lates O/H, N/O, and $\log U$ using mid-infrared emission lines from Br α λ 4.05 μm up to [[N II]] λ 122 μm .

- HII-CHI-MISTRY-TEFF. It calculates the equivalent effective temperature (T_*) and $\log U$ using optical emission lines and O/H, as described in Pérez-Montero et al., 2019b. It can be used to derive the fraction of escaping photons for objects with He II λ 4686 Å as described in Pérez-Montero et al., 2020.

All versions of HCM have been written in `python` and should work for both versions 2 and 3. All can be downloaded from the webpage of HCM at <http://www.iaa.es/epm/HII-CHI-mistry.html>. In this document I describe features for the public versions of HCM 5.0 and equivalent, which require the `python` library `numpy`. Each package contains the corresponding script (named with `.py` extension), a `txt` file with instructions, an input text file example, and different libraries containing the information from the models, all compressed in a `tgz` file. These can be uncompressed using from a terminal prompt the command

```
> tar xvfz HCM_v5.0.tar.gz
```

1.2. How does the code work?

All versions of HCM performs a bayesian-like approach to derive the chemical abundances, ionization parameter or T_* . In brief, for a given property X , the final result is

$$X_f = \frac{\sum_i X_i / \chi_i}{\sum_i 1 / \chi_i} \quad (1.1)$$

where X_f is the result, X_i are the input values in each one of the models of the grid, and χ are the weights assigned to each one of the models, calculated as the quadratic difference between the observed and the predicted values for some specific emission-line ratios

$$\chi_i^2 = \sum_j \frac{(O_j - T_{ji})^2}{O_j} \quad (1.2)$$

being O_j and T_{ji} the observed and model-based values, respectively, for the considered emission-line dependent ratios. These are described in each one of the papers of the different versions of HCM as a function of the input emission lines.

The errors assigned to each result is calculated as the quadratic sum between the dispersion of all results obtained following a Monte-Carlo iteration through the nominal values perturbed with the nominal input observational errors, and the mean of all intrinsic uncertainties assigned to the bayesian process, calculated as:

$$(\Delta X)^2 = \frac{\sum_i (X_f - X_i)^2 / \chi_i}{\sum_i 1 / \chi_i} \quad (1.3)$$

For those versions of HCM aimed at the calculation of chemical abundances, the first iteration is used to provide an estimation for the abundance of a secondary ion (N for HCM and HCM-IR, or C for HCM-UV), relative to oxygen, as these are based on emission-line ratios very low sensitive to $\log U$. This also has the advantage that the grid can be constrained in a second iteration, once N/O or C/O are fixed, to calculate both O/O

and $\log U$ using N or C lines without any a-priori assumption for the relation between O and the respective secondary element.

In the case of HCM-TEFF this previous iteration is not performed, but the grid of models is interpolated to fix the value of O/H to the input value in order to minimize the dependence of both T_{*} and U on metallicity.

2. HCM in the optical

2.1. Running the program

To run the HCM code in version 5.0, we must typed in the terminal prompt:

```
> python HCM_v5.0.py
```

A text describing the code will appear on the screen and it will ask for the input file containing the observational information. Alternatively, this name can be typed when the code is invoked, along with the desired number of iterations for the Monte-Carlo simulation:

```
> python HCM_v5.0.py input.txt 100
```

When the number of iterations is not specified, this is put by default to 25.

2.2. The input file

The input file must be an `ascii` file, with as many rows as the number of objects or pointing for which we want to perform the calculation. Each introduced column represents the identification for each one and the reddening corrected relative to $H\beta$ fluxes with their corresponding errors. In previous versions of HCM, all columns should be introduced, but from version 5.0 a first row with the labels of the introduced columns is just required and the order is not essential. The columns assigned to the observational relative errors are neither mandatory. The labels for the emission lines in this version are:

- ID. To identify each row with a name..
- OII_3727 and eOII_3727, for $[[\text{O II}]] \lambda 3727 \text{ \AA}$ and its error. This is assumed to be the addition of $\lambda 3726 \text{ \AA}$ and $\lambda 3729 \text{ \AA}$ in resolved spectra.
- NeIII_3868 and eNeIII_3868, for $[[\text{Ne III}]] \lambda 3868 \text{ \AA}$ and its error.
- OIII_4363 and eOIII_4363, for $[[\text{O III}]] \lambda 4363 \text{ \AA}$ and its error.
- OIII_4959 and eOIII_4959, for $[\text{oeiii}] \lambda 4959 \text{ \AA}$ and its error.
- OIII_5007 and eOIII_5007, for $[[\text{O III}]] \lambda 5007 \text{ \AA}$ and its error. When only of the two stron $[[\text{O III}]]$ lines is given, the code assumes the theoretical ratio between them into the account.
- NII_6584 and eNII_6584, for $[[\text{N II}]] \lambda 6584 \text{ \AA}$ and its error.
- SII_6716 and eNII_6716, for $[[\text{S II}]] \lambda 6716 \text{ \AA}$ and its error.
- SII_6731 and eNII_6731, for $[[\text{S II}]] \lambda 6731 \text{ \AA}$ and its error.
- Alternatively, it can be given the addition of these two last with SII.6725 and SII.6725.

2.3. Selecting the grid of models

If the input file is correct, the code will ask for the grid of models to perform the calculation:

- (1) POPSTAR with Chabrier IMF, age = 1 Myr with binaries
- (2) AGN, double component, $a(\text{OX}) = -0.8$, $a(\text{UV}) = -1.0$
- (3) AGN, double component, $a(\text{OX}) = -1.2$, $a(\text{UV}) = -1.0$

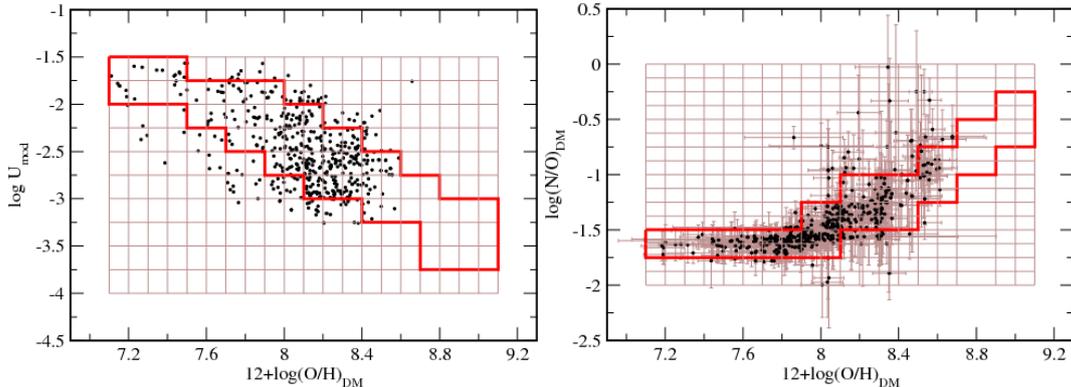


FIGURE 1. At left, relation between total oxygen abundance and ionization parameter for the sample studied in Pérez-Montero, 2014. The solid red line encompasses the most probable combination of parameters occupied by the objects. At right, empirical relation between $12+\log(\text{O}/\text{H})_{\text{DM}}$ and $\log(\text{N}/\text{O})_{\text{DM}}$ with the region occupied by the grid of models in case of no available observational information to constrain N/O.

Choose SED of the models:

All grids have been calculated using the code CLOUDY v.17 (Ferland et al., 2017) assuming a central ionizing source and a plane-parallel geometry. The grid 1 is described in Pérez-Montero, 2014 and is calculated using POPSTAR (Mollá et al., 2009) cluster model atmospheres with instantaneous burst at an age of 1Myr, assuming a Chabrier, 2003 initial mass function (IMF) and a constant electron density of 100 cm^{-3} . Grids 2 and 3 can be used to derive chemical abundances in the NLR of AGN and are described in Pérez-Montero et al., 2019a. These assume a doubled peak power law spectral energy distribution with a parameter $\alpha_{UV} = -1.0$ and $\alpha_{ox} = -0.8$ (grid 2) and -1.2 (grid 3). The assumed electron density is 500 cm^{-3} .

All grids cover an input O/H value in the range $[6.9, 9.1]$ in bins of 0.1 dex and N/O in the range $[-2.0, 0.0]$ in bins of 0.125 dex. Regarding $\log U$ is covered in the range $[-4.0, -1.5]$ for those grids for star-forming regions (i.e. 1 and 2) and in the range $[-2.5, -0.5]$ for the grids for AGN (i.e. 2 and 3). In both cases with a resolution of 0.25 dex.

Once the grid of models has been selected, the code will ask for the use of interpolation between the models of the grid:

Choose models [0] No interpolated [1] Interpolated:

In the interpolated mode, the code makes a linear interpolation of all variables multiplying the resolution by a factor 10. This preserves the results from clustering around certain knots of the grid. On the contrary, this mode slows the calculation time.

2.4. Results

The code will show on the screen the results for the calculations for each one of the rows, showing as well the ratio of completeness of the task. At the end, it will create another `ascii` file with all the results named in the same way as the input file, but adding `_hcm-output.dat`. The first column lists the identification of each row. If this has not been specified in the input file, a number is assigned. The 12 next columns of this file contain the emission-line fluxes used as input, with their corresponding errors. Notice that all lines are in this file, even if they have not been used in the input file.

The next column is an index that indicates if the whole grid of models has been used or, in contrast, a constrained grid has been used instead. This depends mainly in each row on the introduced emission-lines, as a limited observational set implies additional

assumptions to calculate ionic abundances. An index 1 indicates that the whole grid of models has been used. This is done when an estimation of the electron temperature, similarly as in the direct method (see for instance Pérez-Montero, 2017) can be made because both auroral and nebular $[[\text{O III}]]$ emission lines are present. If $[[\text{O III}]] \lambda 4363 \text{ \AA}$ is not given, what it is a common situation in faint or metal-rich nebulae (Pérez-Montero and Díaz, 2005), additional assumptions should be made. Nevertheless, in the case of AGN, not such a relation has been observed, so this relation is not assumed even when the auroral $[[\text{O III}]]$ is not observed, although the range for $\log U$ is restricted in all cases for values > -2.5 , owing to the degeneracy of the O2O3 emission-line ratio as discussed in Pérez-Montero et al., 2019a.

The index 2 then represents the case when electron temperature cannot be estimated and the code assumes an empirical law between O/H and $\log U$ (see Figure 1). This relation assumes that metal-poor objects present in average higher excitation and, on the contrary, metal-rich objects have lower excitation. This is the same assumption behind the use of many high- to low-excitation emission line flux ratios (e.g. O3O2, O3N2) to derive chemical abundances. Finally, index 3 denotes the use of a grid constrained assuming an empirical relation between O/H and N/O as shown in Figure 1. This is necessary in the case that N/O cannot be calculated independently in a first iteration using emission-line ratios such as N2O2 or N2S2. This assumption is behind all strong-line calibrations based on $[[\text{N II}]]$ and implies a constant N/O value for low metallicities due to a mostly primary production of N, and an increasing N/O with O/H when a certain production of secondary N is assumed. This assumption can lead to non-negligible deviations from the real O/H if N/O does not lie in the expected regime, as discussed in Pérez-Montero and Contini, 2009. The final six columns of the output file give the results for O/H, N/O, and $\log U$ with their corresponding errors. If no solution is found, O/H and $\log U$ are denoted by 0, and N/O is denoted by -10

The results and consistency for the abundances derived from HCM both for star-forming objects whose abundances were derived from the direct method (Pérez-Montero, 2014) and for NLR in AGN as derived using detailed photoionization models (Pérez-Montero et al., 2019a) are well discussed. Nevertheless, in Table 1 I provide a list of the mean offsets and the standard deviation of the residuals of the resulting O/H, N/O, and $\log U$ as compared with the input values from the models as a function of the used emission lines. These values illustrate how the code can recover the abundances from the values using only the emission-lines as input, but cannot be taken as true uncertainties.

3. HCM in the ultraviolet

The HII-CHI-MISTRY-UV code (hereafter HCM-UV) is similar to the version in the optical but using a different set of emission lines in the ultraviolet regime and estimating in a first iteration C/O instead N/O. It also admits two emission lines in the optical in order to provide an estimate of the electron temperature, although this implies a larger inaccuracy due to reddening uncertainties. It is described in Pérez-Montero and Amorín, 2017.

3.1. Running the program and preparing the Input file

Similarly to the version in the optical, it can be executed from the terminal prompt with python and allows either just calling the script or to precise the name of the input file and the number of iterations for the Monte Carlo simulations:

```
> python HCM-UV_v4.0.py input.txt 100
```

TABLE 1. Mean offsets and standard deviation of the residuals of the resulting properties derived by HCM when the model from POPSTAR emission lines are used as input as a function of the used emission-line ratios. $[[\text{O III}]]_a$ stands for the $[[\text{O III}]]$ auroral line at λ 4363 Å and $[[\text{O III}]]_n$ for the nebular lines at λ 4959,5007 Å.

Used lines	Grid	O/H		N/O		log U	
		Mean Δ	σ res.	Mean Δ	σ res.	Mean Δ	σ res.
All lines	1	+0.03	0.09	-0.07	0.08	+0.06	0.15
$[[\text{O III}]]_a, [[\text{O III}]]_n, [[\text{N II}]], [[\text{S II}]]$	1	+0.02	0.09	-0.01	0.07	+0.06	0.15
$[[\text{O II}]], [[\text{O III}]]_n, [[\text{N II}]], [[\text{S II}]]$	2	-0.04	0.08	+0.00	0.03	+0.05	0.10
$[[\text{O III}]]_n, [[\text{N II}]], [[\text{S II}]]$	2	-0.03	0.07	+0.00	0.04	+0.05	0.10
$[[\text{O II}]], [[\text{O III}]]_n, [[\text{N II}]]$	2	-0.04	0.24	+0.00	0.23	+0.03	0.09
$[[\text{N II}]], [[\text{S II}]]$	2	+0.00	0.19	+0.00	0.04	+0.02	0.20
$[[\text{O III}]]_n, [[\text{N II}]]$	3	-0.04	0.08	–	–	+0.07	0.13
$[[\text{N II}]]$	3	+0.01	0.15	–	–	+0.03	0.21
$[[\text{O II}]], [[\text{O III}]]_n$	3	+0.00	0.01	–	–	+0.00	0.02
$[[\text{O II}]], [[\text{Ne III}]]$	3	-0.01	0.02	–	–	-0.01	0.03

The input file is an `ascii` file whose first row must the labels of the used columns, including:

- ID. Identification name for each row.
- `Lya_1216` and `eLya_1216`, for $\text{Ly}\alpha$ λ 1216 Å and its error.
- `CIV_1549` and `eCIV_1549` for $\text{C IV}]$ λ 1549 Å and its error.
- `HeII_1640` and `eHeII_1640` for $\text{He II}\lambda$ 1640 Å and its error.
- `OIII_1665` and `eOIII_1665` for $[\text{O III}]$ λ 1665 Å and its error. This includes all lines of the $[\text{O III}]$ multiplet lines between 1660 Å and 1666 Å.
- `CIII_1909` and `eCIII_1909` for $\text{C III}]$ 1909 Å and its error. Again, this includes all the lines of the $\text{C III}]$ multiplet.
- `Hb_4861` and `eHb_4861` for $\text{H}\beta$ λ 4861 Å **AND** its error.
- `OIII_5007` and `eOIII_5007` for $[[\text{O III}]]$ λ 5007 Å and its error.

The following rows must be the names and extinction corrected fluxes in arbitrary units. A value zero can be used for missing values and not all columns are mandatory to get an estimation for O/H, C/O or U .

If the input file is correct the code will ask for the use of interpolation. At the moment, only models for POPSTAR are available and models for AGN will included and described in forthcoming papers.

3.2. Output file and analysis of the results

Once the grid of models is selected, the code will show on the screen the results for each one of the input rows, along with the ratio of completeness. At the end, it will create and output file in `ascii` format with the results. The first column is the corresponding ID and the next 14 will be the input emission lines along with their errors. Notice that all lines are included even if they have not been added to the input file. The next column is an index that informs about if the complete grid has been used (index 1), only when both $[[\text{O III}]]$ λ 1665 and 5007 Å are given) to provide an estimate of the electron temperature. Index 2 denotes a grid with an implicit relation between O/H and log U as shown in Figure 1. Finally, index 3 denotes a grid constrained assuming a relation between C/O and O/H, similar to that shown in Figure 1 for N/O and O/O, considering a fixed C/N to the solar ratio. This is used when C/O cannot be calculated by means of the emission

TABLE 2. Mean offsets and standard deviation of the residuals of the resulting properties derived by HCM-UV when the model from POPSTAR emission lines are used as input as a function of the used emission-line ratios.

Used lines	Grid	O/H		C/O		log U	
		Mean Δ	σ res.	Mean Δ	σ res.	Mean Δ	σ res.
All lines	1	+0.02	0.28	-0.05	0.08	+0.04	0.25
Ly α , C IV],He II,[O III]],C III]	2	+0.02	0.27	-0.06	0.19	+0.05	0.10
C IV],He II,[O III]],C III]	2	+0.10	0.30	-0.14	0.13	-0.11	0.20
Ly α ,C IV],He II,C III]	3	+0.00	0.02	–	–	+0.00	0.01
C IV],He II,C III]	3	-0.01	0.03	–	–	+0.00	0.02

line ratio C3O3, depending on C III] 1909 Å and [O III] λ 1665 Å but an estimate for both O/H and log U can be given using C lines. The six final columns give the results for O/H, C/O, and log U with their corresponding errors. As in the case for the optical, if no solution can be found for both O/H and log U these are denoted as 0 in the file, and -10 for N/O

In Table 2 I list the mean offsets and the standard deviation of the residuals when we use as input for the code the same predictions from the models of the grid, as a function of the different combination of emission lines that lead to a solution.

4. HCM in the infrared

The program HII-CHI-MISTRY-IR (hereafter HCM-IR) calculates total oxygen abundance, nitrogen-to-oxygen abundance ratio, and the ionization parameter from a set of observed emission-lines in the mid infrared. It is described in Fernández-Ontiveros et al (*submitted*).

4.1. Running the program and preparing the input file

Similarly to the other versions of the program, HCM-IR is executed from a terminal prompt using python, and the input text file can be invoked in the same sentence, along with the number of iterations for the Monte Carlo simulations:

```
> python HCM-IR_v2.0.py input.txt 50
```

The input file must be typed in `ascii` format. The first row specifies the input emission lines, with the following accepted labels:

- ID. For the identification name for each row.
- HI_4m and eHI_4m for Pa α at λ 4.07 μm and its error.
- HI_7m and eHI_7m for Br α at λ 7.46 μm and its error.
- SIV_10m and eSIV_10m for [[S IV]] at λ at 10.5 μm and its error
- HI_12m and eHI_12m for Hu α at λ 12.46 μm and its error.
- NeII_12m and eNeII_12m for [[Ne II]] λ 12.8 μm and its error.
- NeIII_15m and eNeIII_15m for [[Ne III]] λ 15.5 μm and its error.
- SIII_18m and eSIII_18m for [[S III]] λ 18.8 μm and its error.
- SIII_33m and eSIII_33m for [[S III]] λ at 33.7 μm and its error.
- OIII_52m and eOIII_52m for [[O III]] λ 52 μm and its error.
- NII_57m and eNII_57m for [[N II]] λ 57 μm and its error
- OIII_88m and eOIII_88m for [[O III]] λ 88 μm and its error

TABLE 3. Mean offsets and standard deviation of the residuals of the resulting properties derived by HCM-IR when the models from POPSTAR emission lines are used as input as a function of the used emission-line ratios.

Used lines	Grid	O/H		N/O		log U	
		Mean Δ	σ res.	Mean Δ	σ res.	Mean Δ	σ res.
All lines	2	+0.04	0.09	-0.00	0.01	+0.02	0.07
[[Ne II]], [[Ne III]], [[S II]], [[S IV]], [[O III]], [[N II]], [[N III]]	2	+0.03	0.13	+0.00	0.01	+0.02	0.07
[[O III]], [[N II]], [[N III]]	2	-0.01	0.18	+0.00	0.01	+0.02	0.15
<i>HI</i> , [[S III]], [[S IV]], [[Ne II]], [[Ne III]]	3	+0.04	0.04	–	–	-0.01	0.01
[[S III]], [[S IV]], [[Ne II]], [[Ne III]]	3	+0.04	0.04	–	–	+0.00	0.01
[[S III]], [[S IV]]	3	+0.02	0.02	–	–	+0.00	0.01
[[Ne II]], [[Ne III]]	3	+0.03	0.09	–	–	+0.00	0.06
[[N II]], [[N III]]	3	+0.00	0.11	–	–	+0.00	0.06

- NII_122m and eNII_122m for [[N II]] λ 122 μ m and its error.

Then the program will ask for the chosen grid of models. For version 2.0, only models from POPSTAR are available, but future versions will also incorporate options for AGN input SEDS. The will also ask for the possibility of model interpolation to increase by a factor 10 the resolution of the grid, but this can slow the time of calculation.

4.2. Output file and analysis of the results

As in previous versions of HCM, once the grid of models has been selected, the code will show in the screen results for O/H, N/O, and log U along with the ratio of completeness of the task. It will also create another `ascii` file whose first column is the identification of each row and the next 28 are the possible emission lines with their errors. Again, all lines are listed, even if they have not been included in the input file. The next column indicates if the complete grid is used (index 1). However, as no auroral lines are available in this spectral range, this is not used in any case. The index 2 corresponds to the grid constrained following an empirical relation between O/H and log U , as shown in Figure 1. Finally, index 3 is used when no previous estimation of N/O, using the N3O3 parameter can be done, and an empirical relation between O/H and N/O is assumed as shown in right panel of Figure 1.

In Table 3 I list mean offsets and standard deviations of the residuals for the resulting O/H, N/O, and log U derived from the code from different sets of input emission lines, as compared with the values used as input in each grid of the model.

5. HCM for the calculation of T_*

The HII-CHI-MISTRY-TEFF)hereafter HCM-TEFF is different to the previously described versions of HCM, as its aim is not the derivation of chemical abundances, but the calculation of the equivalent effective temperature of the ionizing source or the fraction of escaping photons. This code makes use of the so-called softness parameter (Vilchez and Pagel, 1988, Pérez-Montero and Vilchez, 2009) and it is well described in Pérez-Montero et al., 2019b and Pérez-Montero et al., 2020.

5.1. Running the program and preparing the input file

Running the program is equivalent to the previous versions of HCM, through the terminal prompt:

```
> HCM-Teff_v5.0.py input.txt 50
```

The input file must be written in `ascii` format. The first row should correspond to the labels of the used emission lines. The code admits the following lines:

- ID. identification name for each row.
- `12logOH` and `e12logOH` for the total oxygen abundance $12+\log(\text{O}/\text{H})$ and its error. Maybe this can be calculated using HCM in a previous iteration.
- `OII_3727` and `eOII_3727` for $[[\text{O II}]] \lambda 3727 \text{ \AA}$ and its error. As in the case of HCM this represents the addition of the two lines of the doublet of $[[\text{O II}]]$ if there is good spectral resolution.
- `OIII_4959` and `eOIII_4959` for $[[\text{O III}]] \lambda 4959 \text{ \AA}$ and its error.
- `OIII_5007` and `eOIII_5007` for $[[\text{O III}]] \lambda 5007 \text{ \AA}$ and its error. As in the case of HCM if only one of the two lines of the $[[\text{O III}]]$ doublet 4959,5007 is introduced, the code assumes its addition taking into account its theoretical relation.
- `HeI_4471` and `eHeI_4471` for $\text{He I } \lambda 4471 \text{ \AA}$ and its error.
- `HeI_5876` and `eHeI_5876` for $\text{He I } \lambda 5876 \text{ \AA}$ and its error.
- `HeII_4686` and `eHeII_4686` for $\text{He II } \lambda 4686 \text{ \AA}$ and its error.
- `SII_6716` and `eSII_6716` for $[[\text{S II}]] \lambda 6716 \text{ \AA}$ and its error.
- `SII_6731` and `eSII_6731` for $[[[\text{S II}]] \lambda 6731 \text{ \AA}$ and its error. It is also possible to use the addition of the two $[[\text{S II}]]$ lines using `SII_6725` and `SII_6725`.
- `SIII_9069` and `eSIII_9069` for $[[\text{S III}]] \lambda 9069 \text{ \AA}$ and its error.
- `SIII_9532` and `eSIII_9532` for $[[[\text{S III}]] \lambda 9532 \text{ \AA}$ and its error. As in the case of $[[\text{O III}]]$, if one of the two $[[\text{S III}]]$ nebular lines is not introduced, the code assumes its addition taking the theoretical expected ratio.

The following rows of the table correspond to the different values for which the code should derive T_* and $\log U$. The lines must be reddening corrected, but it is not necessary that they are relative to $H\beta$, as only ratios of lines of different ionization are used. Values `z` must be used when no value is found in a line.

5.2. The grids of models

If the input file is correct, the code will ask for the grid of models and the type of calculation we want to perform:

- (1) Effective temperature, WM-Basic stellar atmospheres (30–60 kK)
- (2) Effective temperature, Black body (30–90 kK)
- (3) Photon escape fraction, BPASS cluster atmospheres, age = 4 Myr, $M_{\text{up}} = 300$, $x = 1.35$, w/binaries

Choose parameter and models:

The option 1 calculate T_* and $\log U$ using WM-BASICsingle star atmospheres from Pauldrach et al., 2001 from 30 to 60 kK, while option 2 extends the range up to 90 kK using black-body spectral energy distributions. On the other hand option 3 gives an estimation of the absorption factor (f_{abs}) that is the ratio of ionizing hydrogen photons that do not escape from the nebula, along with $\log U$. This is calculated using cluster model atmospheres from BPASS v.2.1 from Eldridge et al., 2017 assuming an instantaneous burst at 4 Myr, with binaries and an IMF with slope $X = -1.35$ and an upper mass limit of $300 M_{\odot}$.

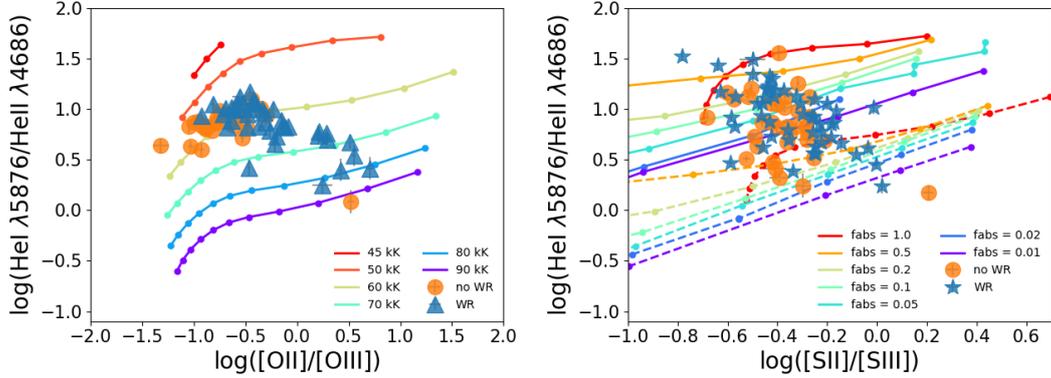


FIGURE 2. Examples of two softness diagrams for the sample of He II-emitters presented in Pérez-Montero et al., 2020. At left $[[\text{O II}]/[\text{O III}]]$ vs $\text{He I}/\text{He II}/$ with models at $12+\log(\text{O}/\text{H})$ and different values for T_* and, at right, using $[[\text{S II}]/[\text{S III}]]$ and BPASS models assuming different values for f_{abs} . In all sequences values in the lower left part correspond to higher values for $\log U$.

In Figure 2 it can be seen two examples of the behaviour of the grids of models in one of the softness diagrams as shown in Pérez-Montero et al., 2020 and how the grids of models cover the space of the emission line ratios.

It is known that the geometry assumed in the models can affect its position on the softness diagrams, so the code will also ask about this:

- (1) Plane-parallel geometry
- (2) Spherical geometry

Choose geometry of the models:

And finally it will ask about the possibility of using interpolation in the final chosen grid:

Choose models [0] No interpolated [1] Interpolated:

5.3. Output file and analysis of the results

As in other versions of HCM, the code will show on the screen the results for each row of the input file along with the ratio of completeness of the task. At the end it will create an output file called as the name of the input file adding `_hcm-teff--ouput.dat`. This file contains all the information for the assumed grids and the results for each one of the rows. The first column corresponds to the identification of each row and the next 18 correspond to the input fluxes of the emission lines. The next two correspond to the assumed O/H, with 0 if no value was introduced. Finally, the four last columns correspond to the calculated T_* (or f_{abs} if the grid 3 was chosen) and $\log U$ with their errors. A value 0 is assigned if no solution is found or the introduced lines are not enough.

In Table 4 I list the mean offsets and the standard deviation of the residuals for the obtained final results as a function of the input emission lines when compared with the input information from the code.

TABLE 4. Mean offsets and standard deviation of the residuals of the resulting properties derived by HCM-TEFF when the model emission lines are used as input as a function of the used emission-line ratios.

Used line ratios	T_* (kK) [†]		$\log U^a$		$\log f_{abs}^\ddagger$		$\log U^b$	
	Mean Δ	σ res.	Mean Δ	σ res.	Mean Δ	σ res.	Mean Δ	σ res.
[[O II]]/[[O III]], [[S II]]/[[S III]], He I/He II	+0.8	2.1	+0.01	0.10	+0.11	0.34	-0.10	0.53
[[O II]]/[[O III]], He I/He II	+0.8	2.1	+0.01	0.10	-0.12	0.27	+0.14	0.47
[[S II]]/[[S III]], He I/He II	+0.6	2.0	+0.03	0.13	-0.19	0.35	+0.10	0.62
[[S II]]/[[O III]], He I/He II	-0.3	5.1	+0.04	0.51	-0.07	0.27	-0.10	0.59
He I/He II	-0.5	6.8	+0.04	0.72	-0.12	0.35	-0.25	0.63

Acknowledgements

All versions and articles of the project HII-CHI-MISTRY have been possible to the financial support from the different editions of the Coordinated project of the Spanish Plan Nacional de astronomía y Astrofísica Estallidos de Formación Estelar en Galaxias: Estallidos 5, Estallidos 6, and Estallidos 7.

REFERENCES

- Chabrier, G. (2003). The Galactic Disk Mass Function: Reconciliation of the Hubble Space Telescope and Nearby Determinations. *ApJ*, 586:L133–L136.
- Eldridge, J. J., Stanway, E. R., Xiao, L., McClelland, L. A. S., Taylor, G., Ng, M., Greis, S. M. L., and Bray, J. C. (2017). Binary Population and Spectral Synthesis Version 2.1: Construction, Observational Verification, and New Results. *PASA*, 34:e058.
- Ferland, G. J., Chatzikos, M., Guzmán, F., Lykins, M. L., van Hoof, P. A. M., Williams, R. J. R., Abel, N. P., Badnell, N. R., Keenan, F. P., Porter, R. L., and Stancil, P. C. (2017). The 2017 Release Cloudy. *RMAA*, 53:385–438.
- Mollá, M., García-Vargas, M. L., and Bressan, A. (2009). PopStar I: evolutionary synthesis model description. *MNRAS*, 398:451–470.
- Pauldrach, A. W. A., Hoffmann, T. L., and Lennon, M. (2001). Radiation-driven winds of hot luminous stars. XIII. A description of NLTE line blocking and blanketing towards realistic models for expanding atmospheres. *A&A*, 375:161–195.
- Pérez-Montero, E. (2014). Deriving model-based T_e -consistent chemical abundances in ionized gaseous nebulae. *MNRAS*, 441(3):2663–2675.
- Pérez-Montero, E. (2017). Ionized Gaseous Nebulae Abundance Determination from the Direct Method. *PASP*, 129(974):043001.
- Pérez-Montero, E. and Amorín, R. (2017). Using photo-ionisation models to derive carbon and oxygen gas-phase abundances in the rest UV. *MNRAS*, 467(2):1287–1293.
- Pérez-Montero, E. and Contini, T. (2009). The impact of the nitrogen-to-oxygen ratio on ionized nebula diagnostics based on [NII] emission lines. *MNRAS*, 398:949–960.
- Pérez-Montero, E. and Díaz, A. I. (2005). A comparative analysis of empirical calibrators for nebular metallicity. *MNRAS*, 361(3):1063–1076.
- Pérez-Montero, E., Dors, O. L., Vílchez, J. M., García-Benito, R., Cardaci, M. V., and Hägele, G. F. (2019a). A bayesian-like approach to derive chemical abundances in type-2 active galactic nuclei based on photoionization models. *MNRAS*, 489(2):2652–2668.
- Pérez-Montero, E., García-Benito, R., and Vílchez, J. M. (2019b). Revisiting the hardening of the stellar ionizing radiation in galaxy discs. *MNRAS*, 483:3322–3335.
- Pérez-Montero, E., Kehrig, C., Vílchez, J. M., García-Benito, R., Duarte Puertas, S., and Iglesias-Páramo, J. (2020). Photon leaking or very hard ionizing radiation? Unveiling the nature of He II-emitters using the softness diagram. *A&A*, 643:A80.

- Pérez-Montero, E. and Vílchez, J. M. (2009). On the hardening of the ionizing radiation in HII regions across galactic discs through softness parameters. *MNRAS*, 400:1721–1725.
- Vílchez, J. M. and Pagel, B. E. J. (1988). On the determination of temperatures of ionizing stars in H II regions. *MNRAS*, 231:257–267.