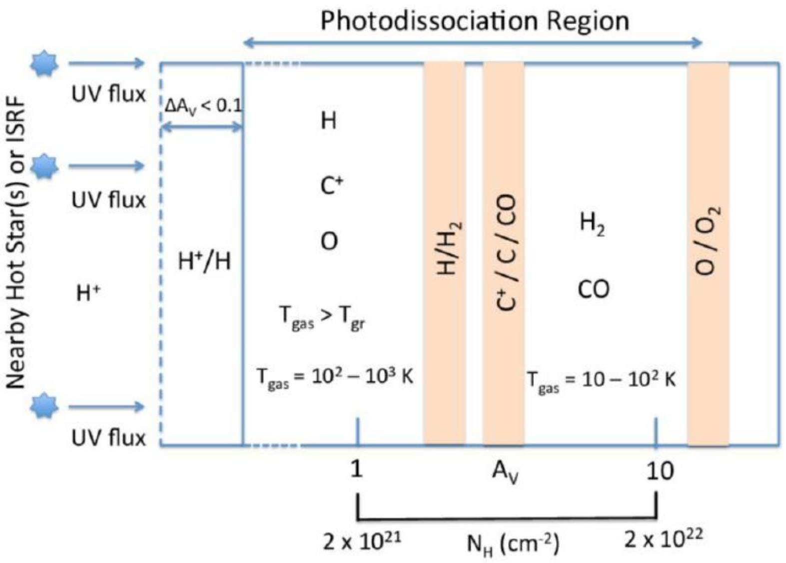
Machine learning techniques to turn rotational lines into powerful diagnostics of the physical conditions inside a Giant Molecular Cloud

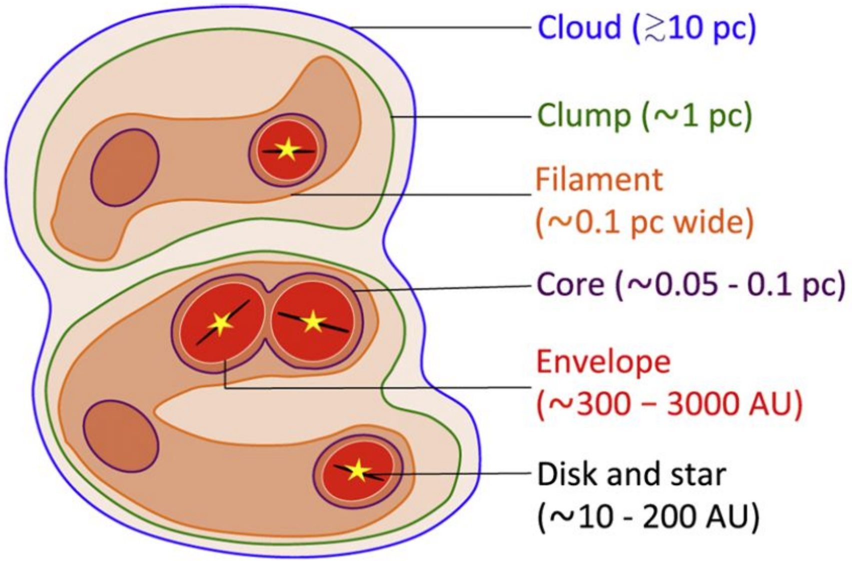
The Orion B Case

Léo BELLOIR, Eleonore DANN, Lilly KORMANN, Taavishi JINDEL, Ziming PENG, Saurav SEN
Supervisors: Laure BOUSCASSE & Jérôme PETY

Star-Forming Regions

Molecular clouds fragment into regions of higher densities (filaments & cores) traced by dust and different molecular lines:

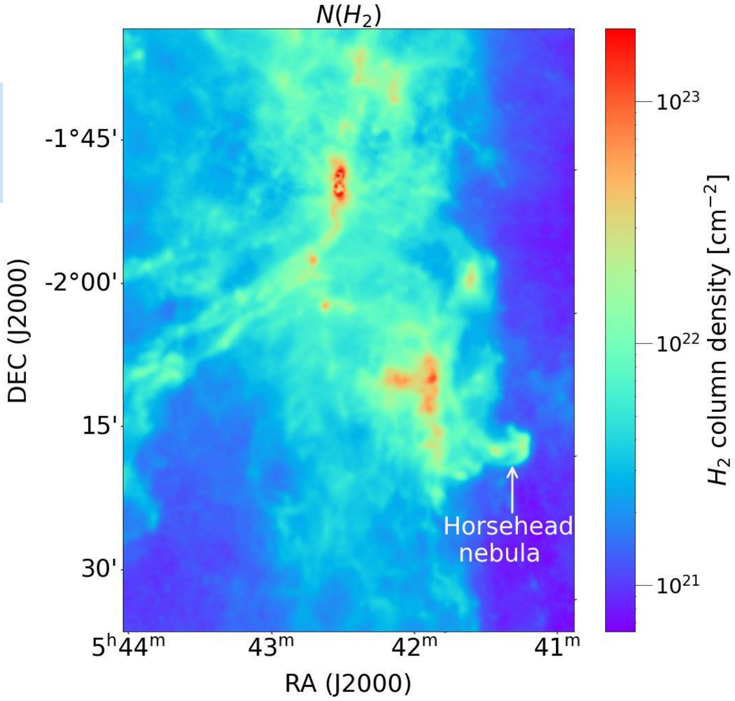




Credit: Hollenbach & Tielens (1999) adapted by S. Topal (left), Pokhrel et al. 2018 (right)

Project Goal

Estimate gas column density $N(H_2)$ using machine learning techniques and multi-line data.



Existing Methods:

CO Based

- CO(1-0) integrated line intensity:

$$N(H_2) = X_{CO(1-0)} I_{CO(1-0)}$$

- CO-to-H_2 conversion factor:

$$X_{CO(1-0)} = 2 \times 10^{20}\ \mathrm{cm}^{-2}\ (\mathrm{K\ km\ s}^{-1})^{-1}$$

Dust Based

- SED fitting ⇒ Dust column density
- Estimate $N(H_2)$ using the A_v-to-H_2 conversion: $N(H_2) = \frac{1.8}{2} \times 10^{21}\ \mathrm{cm}^{-2}\ \mathrm{mag}^{-1}\ A_v$

Data and Molecular Lines

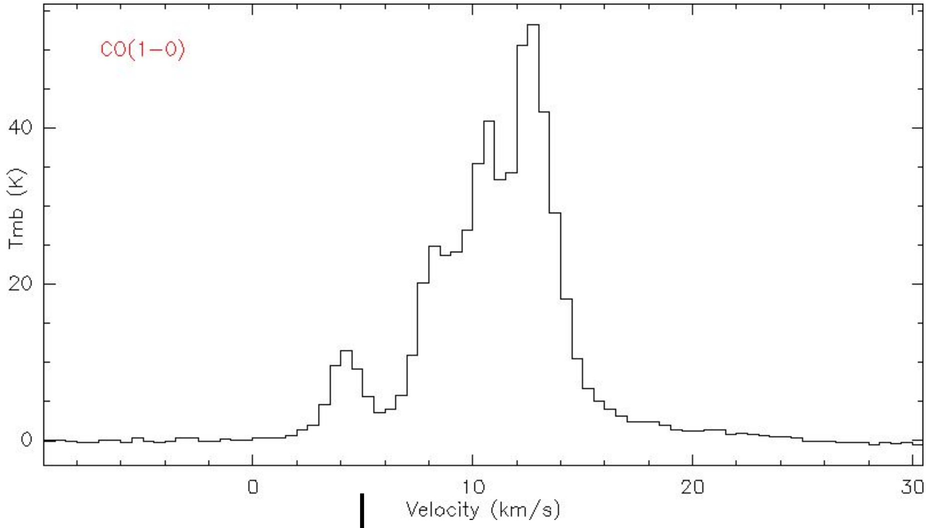
- data from the **O**utstanding **R**adio-**I**maging of **O**rio**N**-**B** program (co-PIs: J. Pety & M. Gerin)
- radio program covering the Orion B molecular cloud (18×13 pc)
- spatial resolution of 27” (< 0.1 pc) and spectral resolution of 0.5 km/s for the full 3 mm atmospheric band

Available Lines for this Project

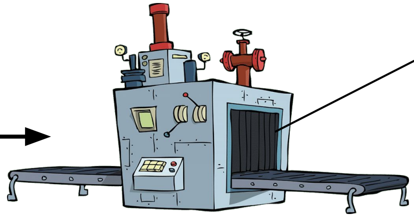
$^{12}CO(1\text{–}0)$, $^{13}CO(1\text{–}0)$, $C^{18}O(1\text{–}0)$, $C^{17}O(1\text{–}0)$, $^{12}CS(2\text{–}1)$, $CH_3OH(2\text{–}1)$, $^{32}SO(2\text{–}1)$, $HCN(1\text{–}0)$, $C_3H_2(2\text{–}1)$, $^{12}CN(1\text{–}0)$, $HCO^+(1\text{–}0)$, $HNC(1\text{–}0)$


30 m IRAM radio telescope

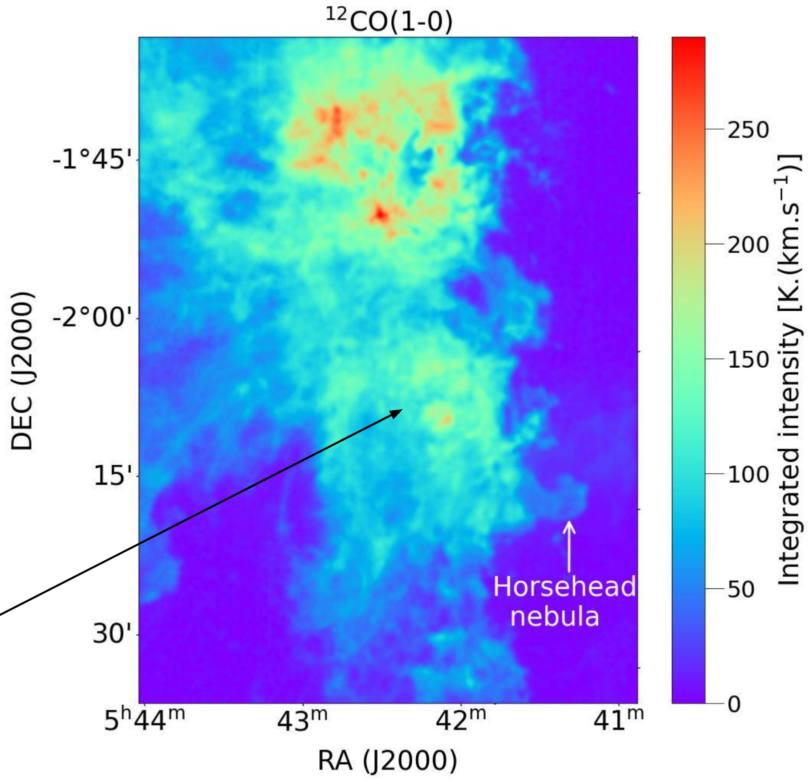
Spectra Integration



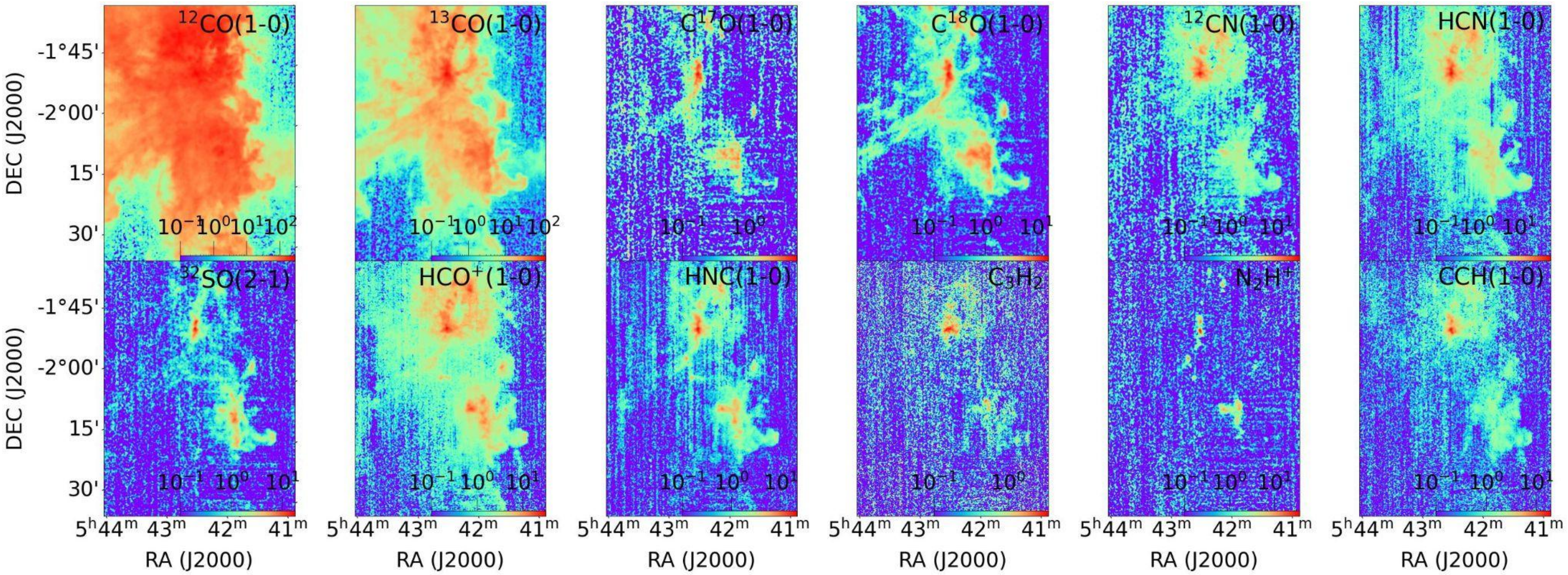
Integration for each pixel



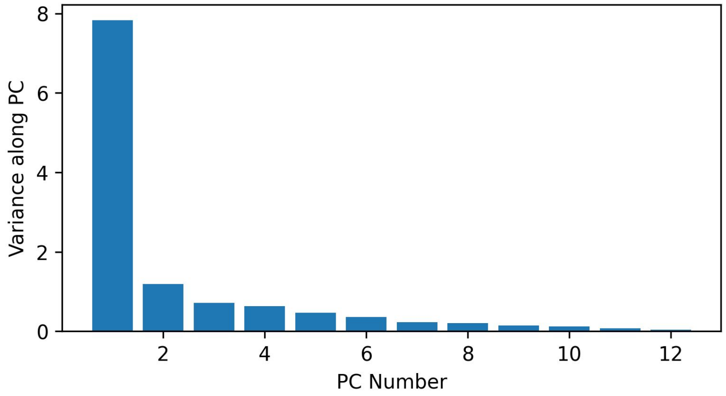
GILDAS Software

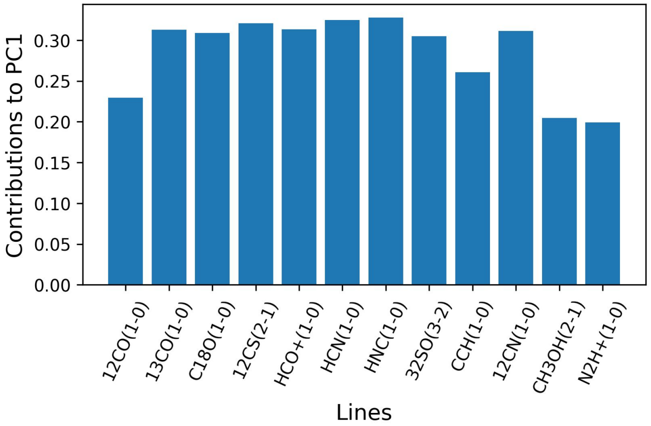


All Integrated Intensity



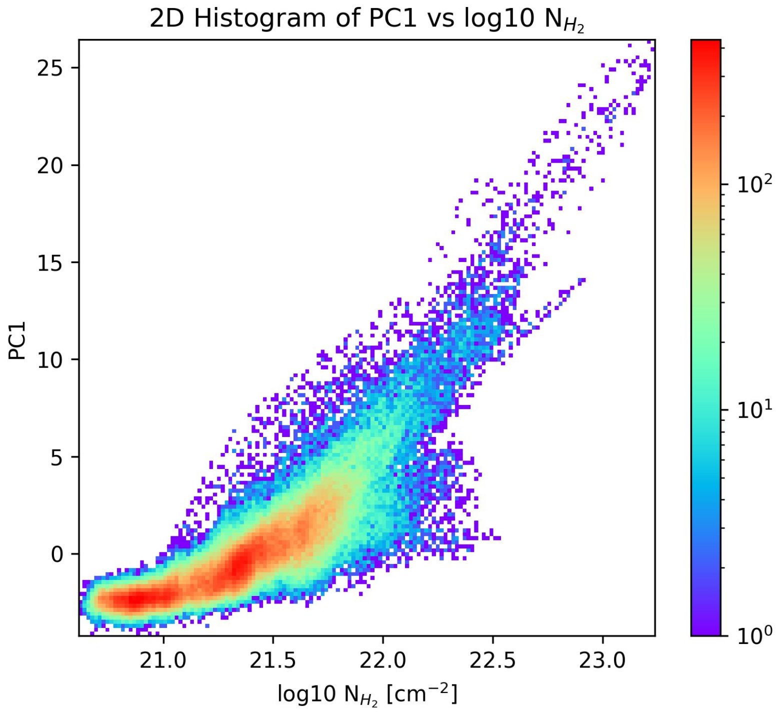
Principal Component Analysis





- First step to understand multidimensional line data correlation
- Correlation implies similar underlying physical processes
- First principal component gives direction of largest variance
- Positive correlations of lines with PC1 may be due to the line intensities positively correlating with N_{H2}

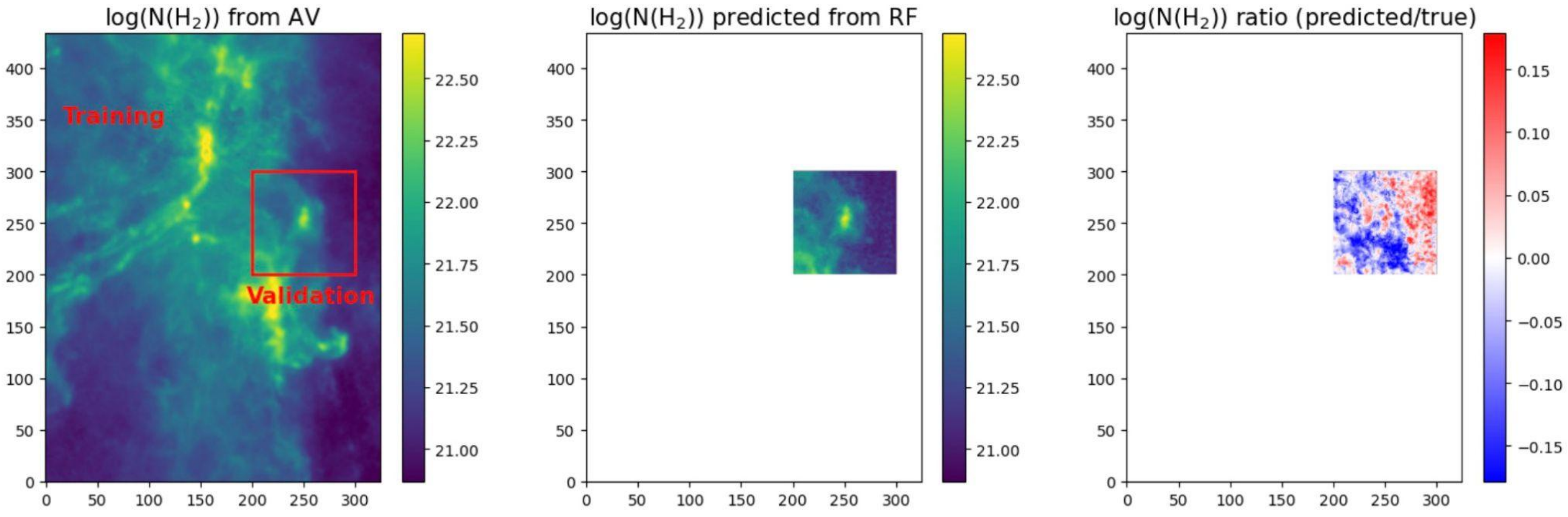
Principal Component Analysis



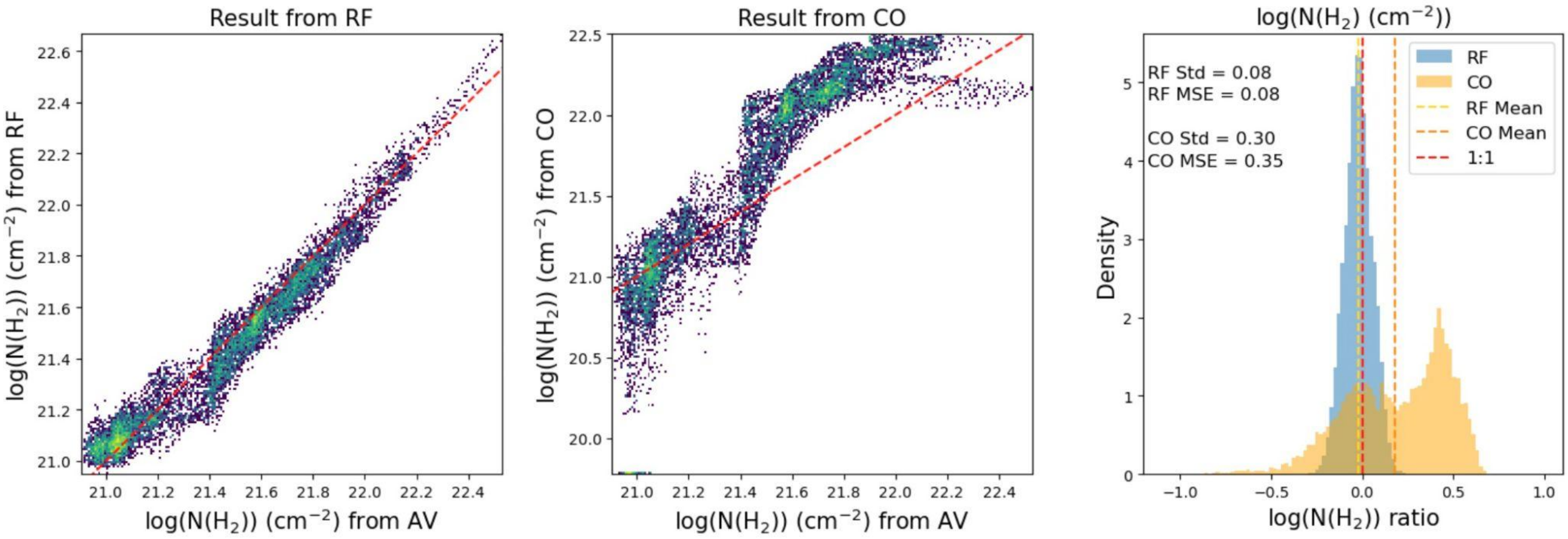
- First step to understand multidimensional line data correlation
- Correlation implies similar underlying physical processes
- First principal component gives direction of largest variance
- Positive correlations of lines with PC1 may be due to the line intensities positively correlating with N_{H2}

⇒ Not a linear correlation at lower column densities! Multiple power laws possible

Random Forest Method



Random Forest vs X_{CO} Method



Conclusion and Outlook

Project Result

- Using RF techniques, the $N(H_2)$ estimated within a factor of 2 compared to a factor of 10 with the X_{CO} method in Orion B.

Future Work

- Is the learnt relation generic? Needs this method to be applied on other nearby molecular clouds like Orion A, Taurus, Perseus etc.
  - It works → gain confidence that the learnt relations has some genericity
  - It doesn't work → need to train on more clouds to increase the training dataset
- What about other galaxies? Relation is probably metallicity dependent. This will mean we need more datasets from other galaxies.