University of Alabama in Huntsville

LOUIS

Summer Community of Scholars (RCEU and HCR) Project Proposals

Faculty Scholarship

1-1-2023

Mining the Data Cubes in Astronomy

Ming Sun
University of Alabama in Huntsville

Nathaniel Pichette
University of Alabama in Huntsville

Follow this and additional works at: https://louis.uah.edu/rceu-proposals

Recommended Citation

Sun, Ming and Pichette, Nathaniel, "Mining the Data Cubes in Astronomy" (2023). *Summer Community of Scholars (RCEU and HCR) Project Proposals*. 424.
https://louis.uah.edu/rceu-proposals/424

This Proposal is brought to you for free and open access by the Faculty Scholarship at LOUIS. It has been accepted for inclusion in Summer Community of Scholars (RCEU and HCR) Project Proposals by an authorized administrator of LOUIS.

Mining the Data Cubes in Astronomy

Current GPA: [redacted]

Data cubes are an important aspect of modern Astronomy research. They are the main format of the final data to be used for analysis from some cutting-edge instruments on major telescopes such as the Atacama Large Millimeter/Submillimeter Array (ALMA), the James Webb Space Telescope (JWST) and the Very Large Telescope (VLT) (e.g., Multi Unit Spectroscopic Explorer - MUSE). Data cubes are especially useful for modern research because they have the spectra for each pixel in a telescope image. This is what makes the data a cube, as the x-axis and y-axis are the standard position axes with the z-axis being wavelength, frequency, or any other unit to express the spectra of the light received. This allows for very in depth analysis of the celestial targets these telescopes observe (e.g., galaxies, nebulae). Such analysis includes comparing the amount of light received by the telescope at various wavelengths in the range recorded and creating diagrams of relative velocities of different parts of galaxies. Emission and absorption spectral lines that can be seen when analyzing the spectra of galaxies helps to determine the locations of certain features. An example of this is that a certain carbon monoxide emission line at 230 GHz can be used to determine locations of star formation in a galaxy since this particular emission line indicates the presence of molecular clouds responsible for star formation.

While these cubes are powerful tools for conducting Astronomy research, they have to be created from the raw satellite data using several codes that are typically written in Python. Analysis requires the usage of additional codes or specialized programs like CARTA, the Cube Analysis and Rendering Tool for Astronomy, as the data cubes regularly have tens of millions or more spaxels, which are the three-dimensional pixels. Codes such as pPXF and Pycprops not only serve to make analyzing these large datasets more manageable, but also further help locate features of interest within galaxies that may have been missed by the human eye during manual analysis.

The goal of this project is to analyze some galaxies from the expansive data archives available for ALMA and MUSE in addition to writing code and improving existing code to help with such analysis of data cubes made from such telescope data. Programs such as DS9 and CARTA will be used to visualize the data cubes for any manual analysis that is necessary. Since different methods can be used to analyze the same set of data, running several of them and determining the optimal settings for each can help determine which method is the best for a given type of galaxy and given goal for analysis. Some current python codes can be improved to require less manual hardcoding to be done when using the same code on different galaxies. Current code for isolating specific wavelengths in MUSE data cubes is one such code that could use this improvement, as the current method used involves manual calculations in order to determine which wavelength ranges should be observed. Expanding codes like this one to involve less manual inputs can help to speed up the analysis process, especially since this particular code might be run several times for a single data cube. Overall, this project will strive to improve the effectiveness of current work being done with Dr. Sun in analyzing ALMA and MUSE data cubes, while also offering these improvements to others doing similar work.

The project will be completed over the ten week period spanning the week of May 29, 2023 to the week of July 31, 2023. In the earlier weeks of the project, the focus will be on improving and writing code in order to conduct some of the various analysis techniques on a

small number of data cubes. Once these codes appear to be working correctly and be tuned to give the best results during analysis, more cubes can be used to ensure that they work beyond just the initial data used. Near the end of the project, all of the findings and code will be compiled together for future use for analysis of similar galaxies and data. Any additional helpful information for running the processes will also be included. Analysis of the flux and other factors of the galaxies conducted during the project and comparisons made to previously conducted research will also be included here. The proposed research is expected to have a significant contribution to Dr. Sun’s research with the ALMA and MUSE data that will generate many impactful publications.