4-22-2024

Mining the Data Cubes in Astronomy: ESO 137-002

Nathanael Arthur Pichette
University of Alabama in Huntsville

Mining the Data Cubes in Astronomy: ESO 137-002:

by

Nathanael Arthur Pichette

An Honors Capstone

submitted in partial fulfillment of the requirements

for the Honors Diploma

to

The Honors College

of

The University of Alabama in Huntsville

April 22, 2024

Honors Capstone Project Director: Professor Ming Sun

Nathanael Pichette 04/22/2024

Student Date

Ming Sun 04/22/2024

Project Director Date

JAM Digitally signed by James A Miller
Date: 2024.04.22 15:49:13 -05'00'

Department Chair Date

Sean Lane Digitally signed by Sean M Lane
Date: 2024.04.24 12:14:39 -05'00'

Honors College Dean Date

Honors College
Frank Franz Hall
+1 (256) 824-6450 (voice)
+1 (256) 824-7339 (fax)
honors@uah.edu

Honors Thesis Copyright Permission

This form must be signed by the student and submitted with the final manuscript.

In presenting this thesis in partial fulfillment of the requirements for Honors Diploma or Certificate from The University of Alabama in Huntsville, I agree that the Library of this University shall make it freely available for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by my advisor or, in his/her absence, by the Chair of the Department, Director of the Program, or the Dean of the Honors College. It is also understood that due recognition shall be given to me and to The University of Alabama in Huntsville in any scholarly use which may be made of any material in this thesis.

Nathanael Pichette
____

Student Name (printed)

Nathanael Pichette
____

Student Signature

04/22/2024
____

Date

Abstract

The project "Mining the Data Cubes in Astronomy" began in the Summer of 2022. This project initially involved analyzing a set of twenty-four Atacama Large Millimeter/submillimeter Array (ALMA) data cubes. This project shifted towards data reduction and basic spectral analysis of European Southern Observatory/ Very Large Telescope/ Multi-unit Spectroscopic Explorer (ESO/VLT/MUSE) data during the Fall 2022 and Spring 2023 semesters. At the beginning of the Summer of 2023, the research group obtained several MUSE datasets for the galaxy ESO 137-002. The "Mining the Data Cubes in Astronomy" project shifted to an in-depth study of this galaxy.

"Mining the Data Cubes in Astronomy: ESO 137-002" involved working with six observing blocks of raw MUSE data that covered three regions around the galaxy ESO 137-002 and its extended tail. Python codes and SaoImageDS9 (DS9) were used to develop each raw data set into a usable data cube. These data cubes were then combined together to create a large mosaic that contained both the galaxy and the entirety of its tail in a single data cube. Kubeviz, which is an IDL program for spectral analysis, and the TARDIS data analysis pipeline, which is a spectral analysis python program, were then used to look for interesting spectral features, mainly focusing on the [NII]-Hα spectral line triplet. While both the galaxy and its tail contained such interesting spectral features, this project focused on studying the galaxy. Since X-ray data has shown ESO 137-002 to have a fairly bright active galactic nucleus (AGN) [5], the MUSE mosaic was used to look for visual light indications of this galaxy's AGN. Averaged spectra from thirteen regions defined within the galaxy were analyzed for this purpose and to check the results of the TARDIS pipeline in the upper portion of the galaxy. While these results do give some indications as to the nature of ESO 137-002, future work is needed to fully explore this galaxy and its tail.

Background

Data Cubes

Data cubes are a common and important data format in cutting edge astronomy research for telescopes such as MUSE, ALMA, the James Webb Space Telescope (JWST), and many others. At a basic level, data cubes are essentially three dimensional tables with headers that define what the values in the table mean and the step size for the x, y, and z axes. For astronomy data, each spaxel, short for spatial pixel, in a data cube contains a flux-like value while the x, y, and z positions of the spaxel give the right ascension, declination, and wavelength, which are the location in space and the location in the spectrum that the spaxel corresponds to. For the used MUSE data cubes, moving one spaxel in the x or y directions reflects a positional shift of 0.2 arcseconds and moving one cell in the z direction is a wavelength change of 1.25Å. The value in each cell for the MUSE data is the flux per wavelength in units of 10^{-20} erg/Angstrom/s/cm^2 that the telescope measured for that particular spaxel.

Programs like DS9 allow for visualization of data cubes as an image for each z-value. For example, the final mosaic data cube for ESO 137-002 has dimensions of 698 x 961 x 3681, which would correspond to 3681 different 698 x 961 pixel images in a program like DS9. This z-axis covers a wavelength range of about 4750-9350Å. In a three dimensional sense, these data cubes represent a measured spectrum for each pixel in the two dimensional picture. The depth of data provided by such data cubes results in fairly large file sizes such as the 21GB ESO 137-002 mosaic data cube that was constructed from some nearly 100GB files.

Ram Pressure Stripping

Galaxy clusters are massive astronomical structures of many gravitationally bound galaxies and an intracluster medium that exists between them. As the galaxies in these clusters travel through this intracluster medium, they can experience drag forces similar to a person trying to run through water. These forces can result in gas being stripped away from the galaxy into tail-like structures that can be seen streaking out from the galaxy into the intracluster medium it is passing through. This process is called ram pressure stripping, and while it can improve star formation by agitating the clouds of gas in the galaxy, it can also snuff it out completely if too much of the gas needed to form stars is removed [2]. As such, ram pressure stripping has important impacts on the evolution of a galaxy beyond just the creation of one or more tails of gas streaking out behind it as it travels through its cluster.

Active Galactic Nuclei

At the center of each galaxy is a super massive black hole with a mass many orders of magnitude greater than the Sun. Some of these black holes have large discs of matter, mainly gas and some dust, called accretion disks that circle around them, which they absorb over time to grow in mass. Such black holes are called active galactic nuclei (AGNs) and their accretion disks are very luminous, particularly in x-rays. Beyond accreting large amounts of matter, AGNs also can impact their host galaxies through jets that cause winds which can either improve or disrupt star formation, depending on the galaxy [4].

Spectral Lines

While there are many spectral lines present in the visual light wavelength range that MUSE covers, the objectives of this project focus on the [NII]-Hα triplet of spectral lines. This triplet is made up of the Hα line, which involves a transition between the n=2 and n=3 energy levels in hydrogen, and the "blue" and "red" [NII] lines, which are forbidden line transitions in singly ionized nitrogen. The amplitude of the blue [NII] line is known to be one third the amplitude of the red [NII] line. The blue [NII] line has the shortest wavelength of the triplet with a rest wavelength about 6548 $\mathring{A}$, followed by the Hα line that has a rest wavelength of about 6563 $\mathring{A}$, and ending with the red [NII] line which has a rest wavelength of about 6583 $\mathring{A}$. Hα and [NII] emission lines can indicate several astronomical processes such as star formation rates or the presence of the warm ionized gas needed for star formation [1][6].

ESO 137-002

ESO 137-002 is a spiral galaxy in the Abell 3626 galaxy cluster that is edge-on from the perspective of Earth. From the MUSE data, this galaxy was determined to have a redshift of about 0.01895 and the Abell 3626 galaxy cluster is around 69.6 Mpc from Earth [5]. This galaxy, shown in **Figure 1**, has a long extended tail streaking out into the intracluster medium, which indicates that this galaxy is undergoing ram pressure striping. Additionally, x-ray data of ESO 137-002 show it to have an AGN [5]. Since both ram pressure striping and the winds from an AGN can have serious impacts on a galaxy and its star formation, ESO 137-002 is a particularly interesting galaxy because it is experiencing both of these effects simultaneously.

Figure 1: 6706Å image From MUSE mosaic data cube of ESO 137-002. Note the tail extending out of the galaxy towards the south-west of the image.

Methods

Data Reduction to Mosaic Creation

Raw MUSE data like the six observing blocks used to study ESO 137-002 has to be put through a variety of processes to eventually be made into a mosaic data cube that can be used for studying the galaxy. The first step of this process is running the MUSE data reduction pipeline. This pipeline consists of python code that puts together the raw telescope data and a collection of calibration files to create the basic images and tables that will be used to create a data cube for that observing block. Two sets of images and tables were created during data reduction for each observing block used during this project, as the telescope is rotated 90 degrees after half of the observing time to help mitigate any errors from the detector.

After the data for an observing block is reduced, the esocombine python code is used to combine the exposures together into a single data cube and an image that averages the spectra. During this step, the images from the data reduction are used to correct the current world coordinate system (wcs) information that is contained in the headers of the files. By comparing the coordinates of stars in these images to star survey information like the Sloan Digital Sky Survey (SDSS) or the Two Micron All Sky Survey (2MASS), any small errors in the coordinates of the MUSE data can be corrected. An average offset is calculated for each exposure by manually comparing the positions of at least five stars in the MUSE data to the locations of those stars in the survey. After the offset of each image in an observing block was calculated using these star surveys, the esocombine code was run to simultaneously combine the two exposures and correct the location. The coordinate correction is important both so that the exposures are properly aligned with each other, and so that the MUSE data cube has the same coordinates as other telescope data for comparing results.

After the data is combined and the wcs is correctly set, the Zurich Atmospheric Purge (ZAP) code is run to further refine the data cube to a point where data analysis can begin for individual observing blocks. The purpose of running the data cubes through ZAP is to remove the noise light from the atmosphere and night sky. This makes spectral analysis easier and more accurate by leaving only the spectral features of the stars and galaxies in the data cube. Each of the six observing blocks available for ESO 137-002 were run through the process of data reduction, esocombine and alignment, and ZAP to get them research ready. **Appendix A** contains images and further description of these six datacubes.

Since the ESO 137-002 data was made up of several observing blocks, the data cubes were combined into a single mosaic for data analysis. This is done to account for overlapping between the various observing blocks and to allow for data analysis to be conducted on all the data at once. This process begins by running esocombine with the reduced data from all six observing blocks at once to create a reference mosaic. While this reference mosaic has some oddities from the overlapping data cubes, its main purpose is to define the table size needed for the final mosaic and the proper coordinates for that mosaic. Python code is then used to place all of the ZAPed data cubes into the reference mosaic using libraries like glob to properly account for overlapping regions in the data. This results in the data cubes from all of the various observing blocks being combined together into a single data cube file.

Spectral Analysis

Both Kubeviz and the TARDIS Pipeline programs were used to do basic fitting and spectral analysis of the completed mosaic data cube for ESO 137-002. Since the TARDIS Pipeline gives superior results for studying the galaxy itself as opposed to its extended tail, this project primarily used the maps file output from TARDIS for analysis. This file contains flux maps, velocity maps, velocity dispersion maps, and associated error maps for all of the emission lines within the spectral range of the data cube. The flux and velocity maps for the red [NII] line were manually analyzed to look for interesting regions to extract averaged spectra from. **Figure 2** shows these two maps with the thirteen regions of interest that were defined. The Astropy Spectral Cube Python library was used to extract spectra from all thirteen regions. This was conducted for both the base ESO 137-002 mosaic data cube and a “residual” cube that subtracts the stellar continuum from the base mosaic.

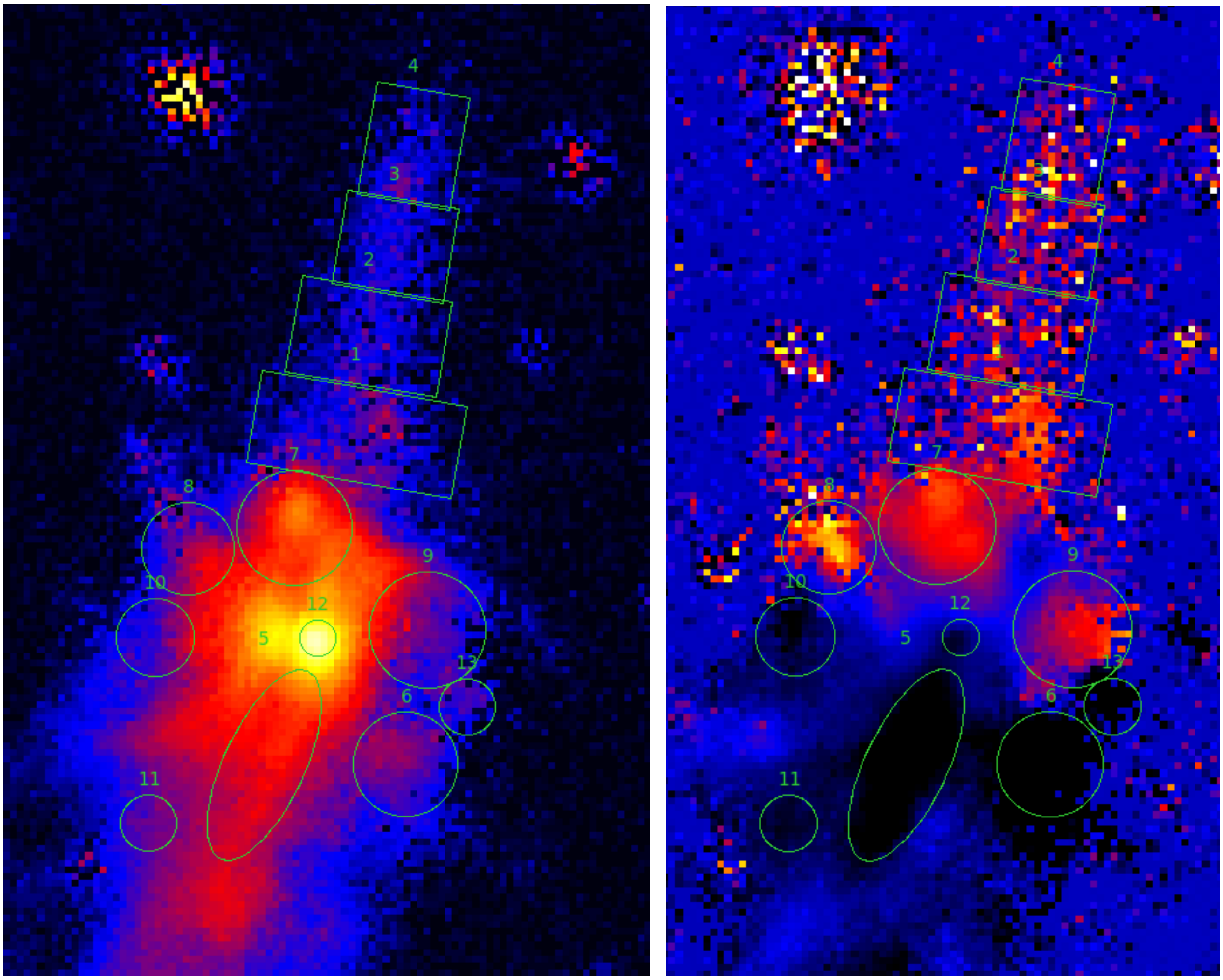
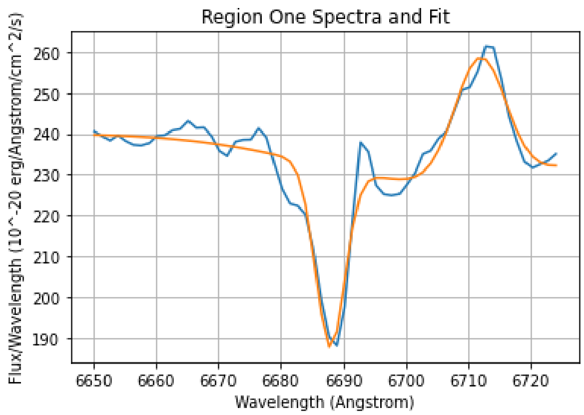


Figure 2a (left): The [NII]6583 (red [NII]) flux map with the thirteen regions of interest displayed on it. The number is above each region and follows the same numbering as the following spectra.

Figure 2b (right): The [NII]6583 (red [NII]) velocity map with the thirteen regions of interest displayed on it. The numbering scheme is the same as for **Figure 2a**.

Results

The extracted spectra for the first four regions of interest from the ESO 137-002 mosaic data cube are fairly similar to each other. In all four spectra, there is a fairly strong Hα absorption line, although a possible small Hα emission line can be seen directly after this absorption line in the first region. Also, all four spectra have a fairly short and broad red [NII] emission line and a possible small blue [NII] emission line that is hard to distinguish from the noise. Since these trends can be seen with just the spectra for regions one and two, spectra for regions three and four are included in **Appendix B**.



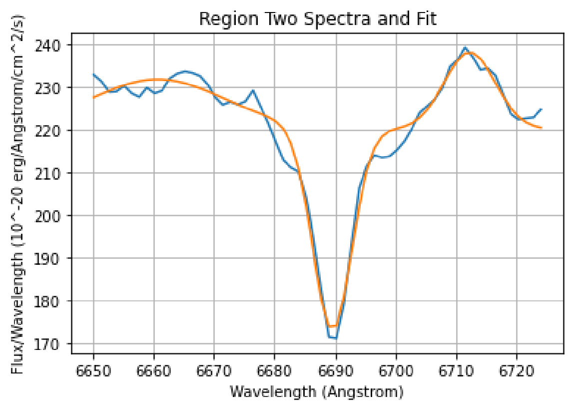
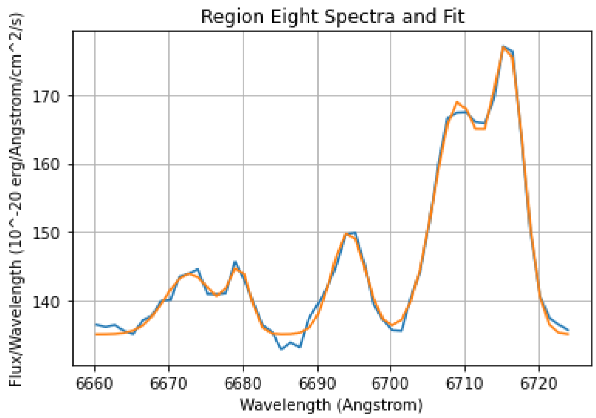


Figure 3a (left): A specutilis line fit (orange) for the extracted spectra (blue) for region one in the mosaic. **Figure 3b (right):** A specutilis line fit (orange) for the extracted spectra (blue) for region two in the mosaic.

The spectra for the rest of the regions have the two [NII] emission lines and the Hα emission line clearly present. While the amount which the lines are broadened varies between the regions, all of the spectra also have broadened emission lines, with the broadest lines being in region twelve – the closest region to the galactic nucleus. Region eight has clearly defined splitting in both [NII] emission lines. Region thirteen has messy, but still visible, splitting in all three emission lines. Region nine and region ten have small shoulder features next to the [NII] emission lines that indicate the possibility of some small emission line splitting in those regions. These trends are best shown with the spectra from regions eight, nine, ten, twelve, and thirteen; the rest of the spectra extracted from the mosaic data cube can be found in **Appendix B**.



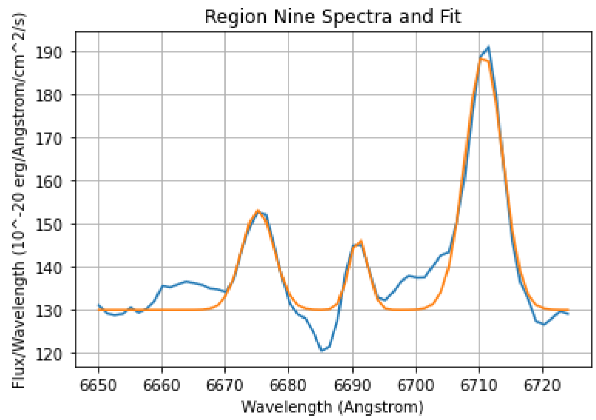
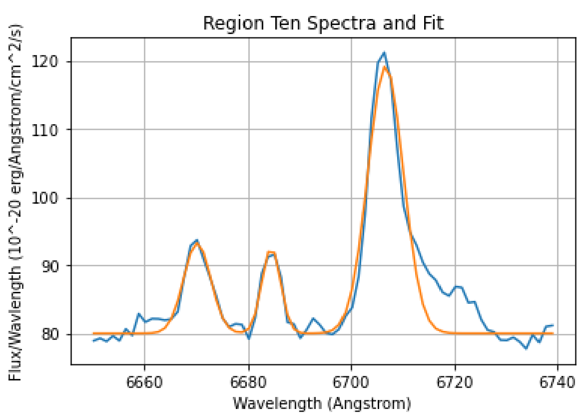


Figure 4a (left): A specutilis line fit (orange) for the extracted spectra (blue) for region eight in the mosaic. Note the split peaks of the two [NII] emission lines.

Figure 4b (right): A specutilis line fit (orange) for the extracted spectra (blue) for region nine in the mosaic. Note the shoulders to the left of the two [NII] emission lines.



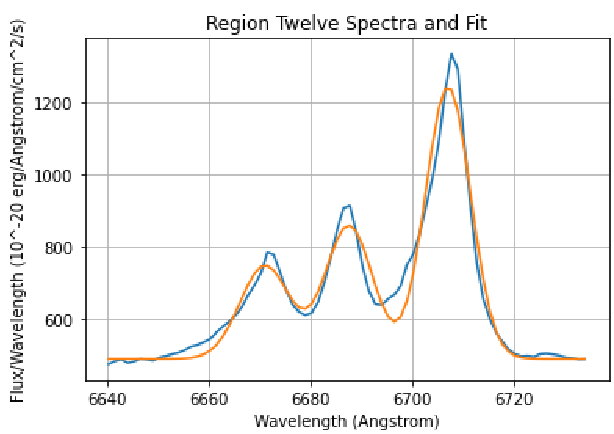


Figure 5a (left): A specutilis line fit (orange) for the extracted spectra (blue) for region ten in the mosaic. Note the shoulder-like feature to the right of the red [NII] emission line.

Figure 5b (right): A specutilis line fit (orange) for the extracted spectra (blue) for region twelve in the mosaic. Note the broad and connected nature of the emission lines.

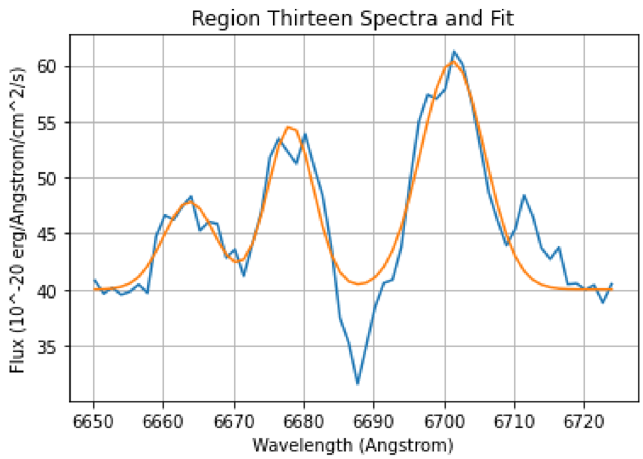


Figure 6: A specutilis line fit (orange) for the extracted spectra (blue) for region thirteen in the mosaic. Note the splitting and noisy appearance of the emission lines.

The spectra for the residual data cube have similar visual appearances to those of the mosaic data cube. However, the wavelength values for spectra from the residual cube are about 500Å to 600Å greater than those from the mosaic cube. While this problem is likely related to the residual cube having a longer wavelength axis than the mosaic cube with a step size of 1Å instead of 1.25Å, this problem has not yet been resolved. As such, the data from the residual cube is not particularly useful for analysis purposes at the current time. The residual cube spectra from region ten is included to show this issue.

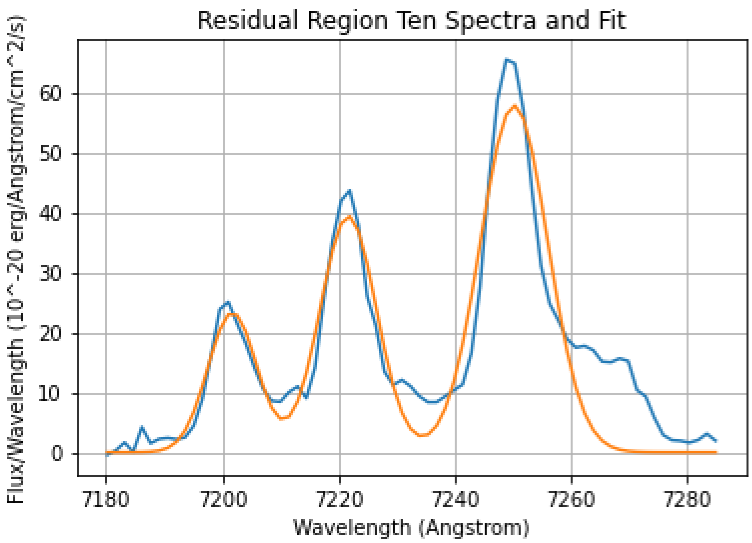


Figure 7: A specutilis line fit (orange) for the extracted spectra (blue) for region ten in the residual cube. Note that while the general appearance of the spectral lines are similar to that of **Figure 5a** the wavelength values are from around 7180Å-7280 Å rather than the correct values around 6650Å-6750Å.

Conclusions

The presence of the red [NII] emission line and faint presence of the blue [NII] emission line, and the Hα emission line in the case of region one, indicate that there is still warm gas present in the northern portion of the galaxy, which is encompassed by regions one through four. This reveals that the edge-on ram pressure stripping ESO 137-002 is experiencing from its north-east [3] has not completely stripped the gas from these arm regions of the galaxy. However, the effects of the ram pressure stripping are evident on this portion of the galaxy in that the magnitudes of the emission lines are lower than those in the other regions, with the exception of region thirteen which is the furthest from the nucleus.

The regions closer to ESO 137-002's nucleus give slight support to the x-ray data that this galaxy likely houses an AGN in two ways. Firstly, the broadened emission lines indicate a velocity dispersion in these regions. AGN winds are a possible cause of this broadening effect, especially since the broadening is most severe around the galactic nucleus encompassed by region twelve shown in **Figure 5b**. Secondly, regions eight, nine, ten, and thirteen show velocity dispersion through curve splitting or the creation of shoulders on the [NII] emission lines, and even possible splitting of the Hα line in region thirteen. These regions form an x-like shape about the nucleus that would be expected from outflow caused by an AGN. Beyond the curve splitting in the spectra, the red [NII] velocity map from the TARDIS Pipeline shows regions eight and nine to have a relatively high velocity compared to the galaxy and its surroundings, while regions ten and thirteen have a relatively low velocity as shown by the bright and dark spots in **Figure 2b**. Together, these velocity related impacts on the emission lines emitted by the warm gas in ESO 13-002 give visual light indications of the possibility of an AGN at the center of the galaxy.

Future work in studying the MUSE data for ESO 137-002 will focus on improving the results for the galaxy region and fully analyzing the extended tail. While the emission line plotting and fitting of [NII]-Hα triplet that was conducted with the specutilis python library was enough to indicate the presence of the warm gas and show some effects of velocity dispersion, more robust code will be used in the future to improve the curve fits and to give additional information about the velocity in these regions. In addition to these improvements with the lines analyzed for this project, further analysis of spectral lines other than the [NII]-Hα triplet will also be conducted on the MUSE mosaic data cube. Particularly, studying the [OIII]/Hβ line ratio will help give a better optical indication of an AGN in ESO 137-002. In addition to these future improvements to the spectral analysis of the mosaic data cube, correcting the code that creates the residual cube would yield a data cube with spectra that have the continuum removed. This would make it much easier to examine effects on the emission lines by removing noise and any remaining sky lines from the data. Although this project has focused predominantly on the galaxy itself, the extended tail can also be studied using this MUSE data. Study of this portion of the galaxy would yield more information on the stripping that ESO 137-002 is experiencing. As such, while several discoveries involving the nature of ESO 137-002 have already been made with the MUSE data cubes collected around this galaxy, there are still many discoveries to be made involving this galaxy from the MUSE data alone.

Acknowledgements

Thank you to ESO/VLT/MUSE for the cutting edge data used for this project. Thank you to Dr. Sun for the guidance throughout this project and throughout all of my research. Thank you to the UAH Honors College and UAH College of Science for starting and encouraging my

research through the Research and Creative Experience for Undergraduates and Honors Capstone Research summer programs.

References

[1] Dojčinović, et al. “The Flux Ratio of the [N II]λλ 6548, 6583 Å Lines in Sample of Active Galactic Nuclei Type 2.” January 2023, Advances in Space Research, vol 11, p. 1219-1226. https://doi.org/10.1016/j.asr.2022.04.041.

[2] Kolcu et al. “Quantifying the Role of Ram Pressure Stripping of Galaxies within Galaxy Groups.” August 2022, Monthly Notices of the Royal Astronomical Society, vol. 515, p. 5877-5893. https://doi.org/10.1093/mnras/stac2177.

[3] Laudari et al. “ESO 137-002: A Large Spiral Undergoing Edge-On Ram-Pressure Stripping with Little Star Formation in the Tail.” Monthly Notices of the Royal Astronomical Society, vol. 509, p. 3938–3956. https://doi.org/10.1093/mnras/stab3280.

[4] Rao et al. “AGN Feedback Through Multiple Jet Cycles in the Seyfert Galaxy NGC 2639.” January 2023, Monthly Notices of the Royal Astronomical Society, vol. 524, p. 1615-1624. https://doi.org/10.1093/mnras/stad1901.

[5] Sun et al. “Spectacular X-ray Tails, Intracluster Star Formation, and ULXs in A3627.” December 2009, The Astrophysical Journal, vol. 708, p. 946. https://iopscience.iop.org/article/10.1088/0004-637X/708/2/946.

[6] Tacchella et al. “H-alpha Emission in Local Galaxies: Star Formation, Time Variability and the Diffuse Ionized Gas.” March 2022, Monthly Notices of the Royal Astronomical Society, Vol. 513, p. 2904–2929. https://doi.org/10.1093/mnras/stac818.

Appendix A: Individual Datacubes

The six individual data cubes that make up the ESO 137-002 mosaic data cube are shown in **Figure A1, Figure A2,** and **Figure A3**. Field one is the cube in the north eastern corner of the mosaic that contains the galaxy itself, field two is the cube in the middle of the mosaic, and field three is in the cube in the south western corner of the mosaic. MUSE collected one observing block of good data for field one, two observing blocks of good data for field two, and three observing blocks of good data for field three. While ESO137-002's extended tail is not readily visible in the field three data cubes themselves as seen in **Figure A3**, the spectral maps indicate that the tail does indeed reach this field.

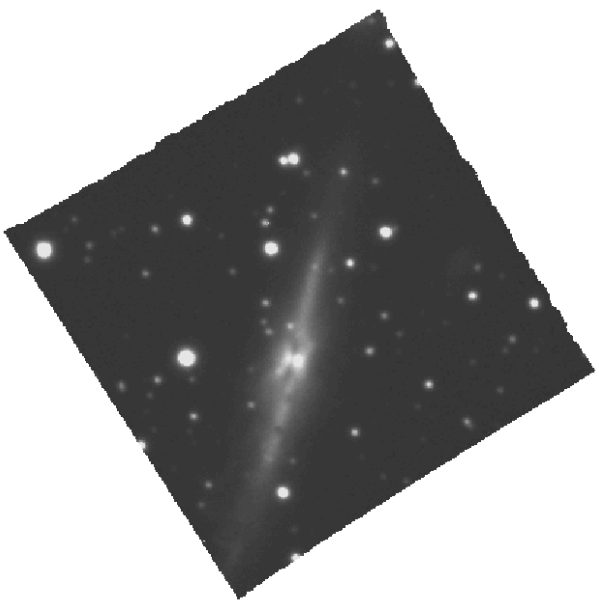


Figure A1: 6706Å image from the field one MUSE data cube for ESO 137-002.

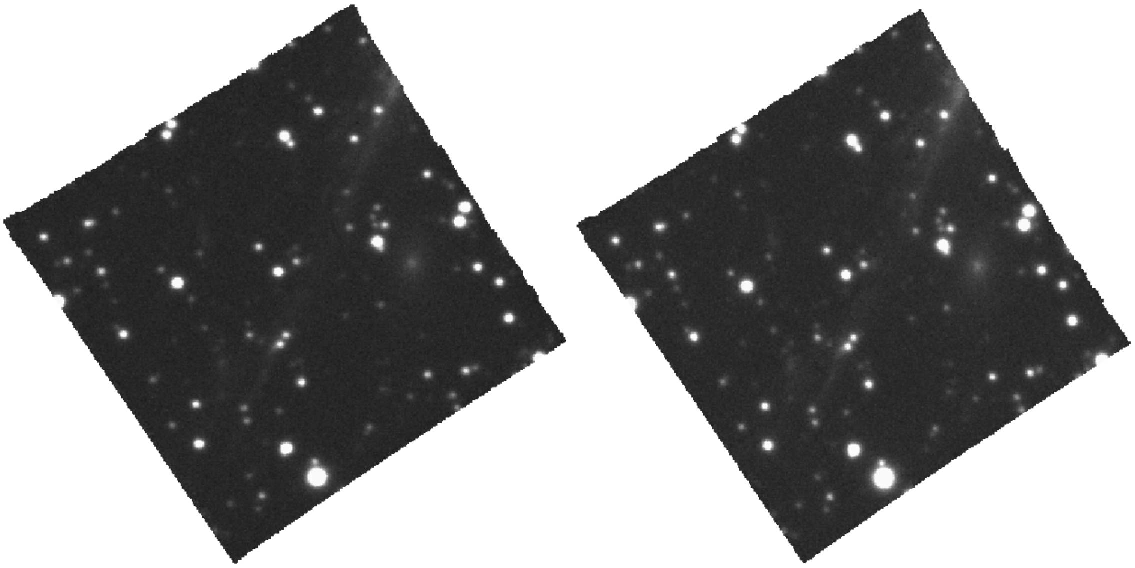


Figure A2: 6706Å images from the two field two MUSE data cubes for ESO 137-002. The left image is from the first dataset for this region and the right image is from the second dataset for this region.

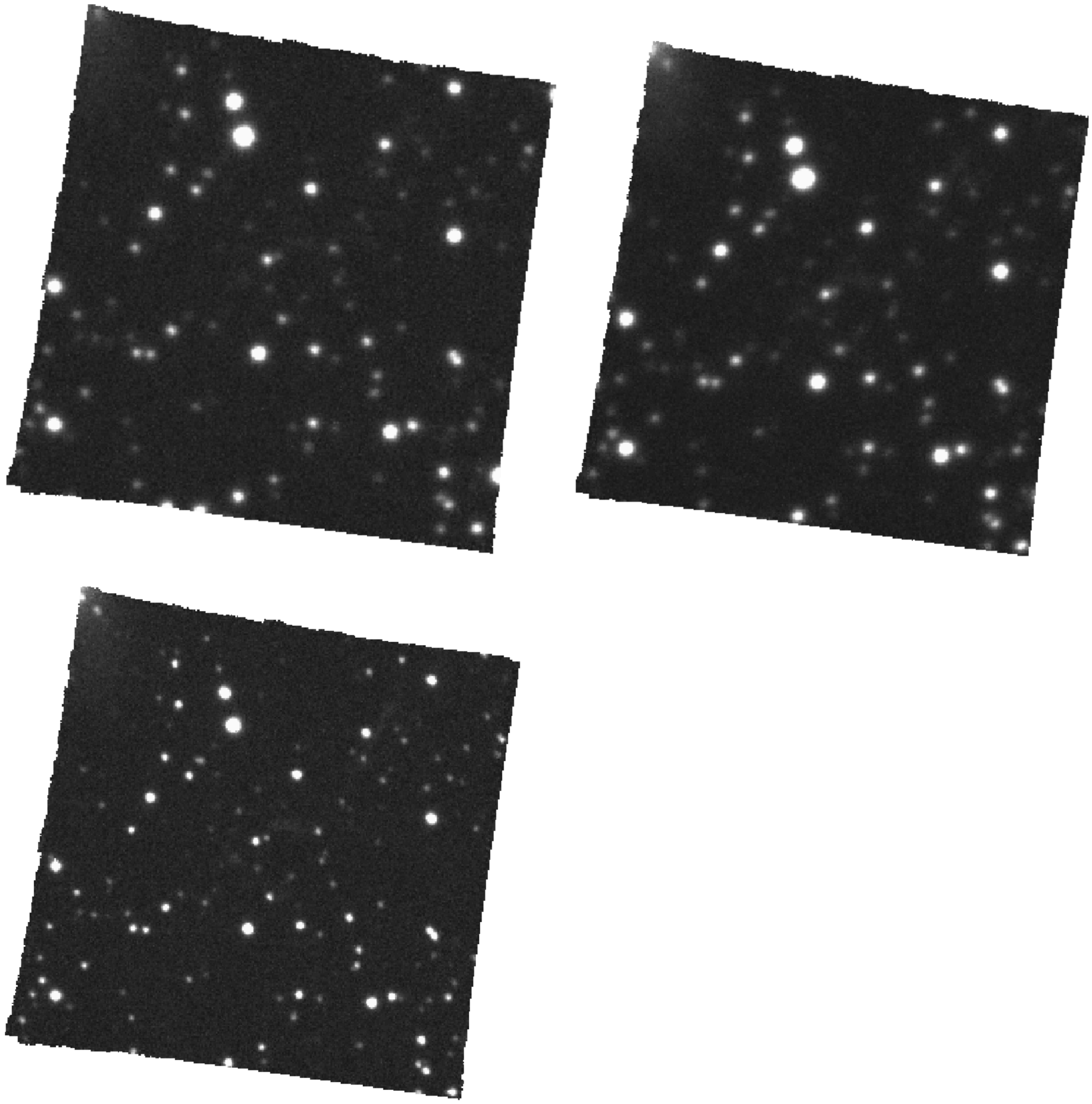


Figure A3: 6706Å images from the three field three MUSE data cubes for ESO 137-002. The top left image is from the first observing block for this field, the top right image is from the second observing block for this field, and the bottom left image is from the third observing block for this field.

Appendix B: Additional Spectra

The spectra for region three and region four are very similar to spectra for region one and region two, so they were not included in the **Results** to avoid redundancy. These spectra are included in **Figure B1**. The spectra for region five, region six, region seven, and region eleven did not have much of any notable features other than broadened emission lines, so they were not included in the **Results** since they are not used in the discussion of the AGN wind. While these regions may have possible small shoulder features on some of the emission lines, this effect was not very notable in these regions. These regions are included in **Figure B2** and **Figure B3**.

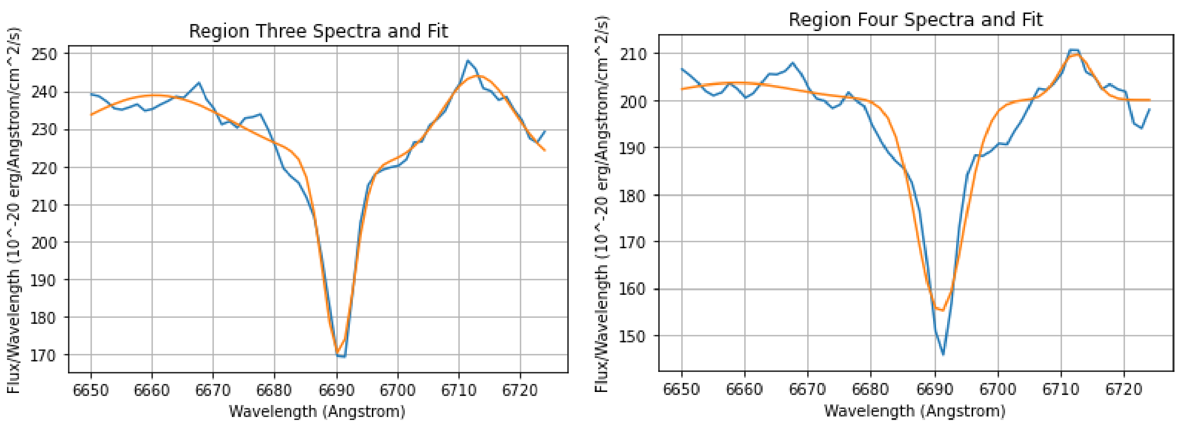


Figure B1: A specutilis line fit (orange) for the extracted spectra (blue) for region three in the mosaic (left) and region four in the mosaic (right).

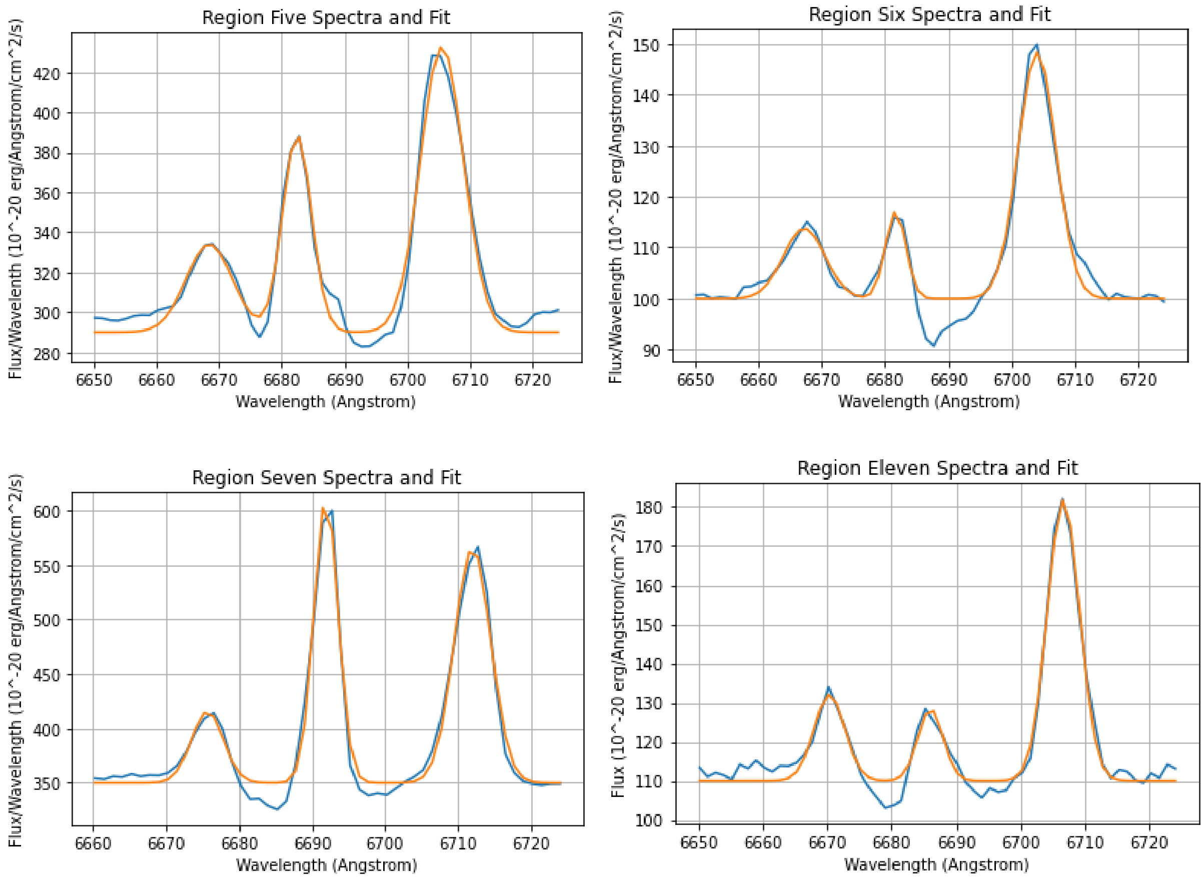


Figure B2: A specutilis line fit (orange) for the extracted spectra (blue) for region five in the mosaic (top left), region six in the mosaic (top right), region seven in the mosaic (bottom left), and region eleven in the mosaic (bottom right).