



SAPIENZA
UNIVERSITÀ DI ROMA

Classifying Seyfert galaxies with deep learning

Scuola di dottorato Vito Volterra

Dottorato di Ricerca in Fisica – XXXII Ciclo

Candidate

Yen Chen Chen

ID number 1780440

Thesis Advisor

Dr. Paolo Giommi

Co-Advisor

Prof. Remo Ruffini

A thesis submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Physics

March 2021

Thesis defended on 24 Sep. 2021
in front of a Board of Examiners composed by:
Prof.ssa Silvia Masi (chairman)
Prof.ssa Simonetta Filippi
Prof. Piero Rosati

Classifying Seyfert galaxies with deep learning

Ph.D. thesis. Sapienza – University of Rome

ISBN: 000000000-0

© 2021 Yen Chen Chen. All rights reserved

This thesis has been typeset by L^AT_EX and the Sapthesis class.

Version: December 25, 2021

Author's email: yen-chen.chen@icranet.org

Publication list

First author

- **Yen-Chen Chen**. Classifying Seyfert galaxies with deep learning. *The Astrophysical Journal Supplement Series*, 256(2):34, October 2021.
- **Yen-Chen Chen** and Chorng-Yuan Hwang. Emission line luminosity distributions of Seyfert 2 galaxies. *Monthly Notices of the Royal Astronomical Society*, 485(3):3402-3408, Mar 2019.
- **Yen-Chen Chen** and Chorng-Yuan Hwang. Morphology of Seyfert galaxies. *Astrophysics and Space Science*, 362(12):230, Nov 2017.

Co-author

- R. Ruffini, R. Moradi, J. A. Rueda, L. Li, N. Sahakyan, **Y.-C. Chen**, Y. Wang, Y. Aimuratov, L. Becerra, C. L. Bianco, C. Cherubini, S. Filippi, M. Karlica, G. J. Mathews, M. Muccino, G. B. Pisani, S. S. Xue. The morphology of the X-ray afterglows and of the jetted GeV emission in long GRBs. *Monthly Notices of the Royal Astronomical Society*, 504(4): 5301-5326, March 2021
- R. Ruffini, R. Moradi, J. A. Rueda, L. Becerra, C. L. Bianco, C. Cherubini, S. Filippi, **Y. C. Chen**, M. Karlica, N. Sahakyan, Y. Wang, and S. S. Xue. On the GeV emission of the type I BdHN GRB 130427A. *The Astrophysical Journal*, 886(2):82, Nov 2019.
- R. Ruffini, R. Moradi, Y. Aimuratov, U. Barres, V. A. Belinski, C. L. Bianco, **Chen, Y. C.**, C. Cherubini, S. Filippi, D. M. Fuksman, M. Karlica, L. Li, D. Primorac, J. A. Rueda, N. Sahakyan, Y. Wang, and S. S. Xue. GRB 190114C: A type 1 BdHN with TeV emission. *GRB Coordinates Network*, 23715:1, Jan 2019.
- R. Moradi, R. Ruffini, C. L. Bianco, **Chen, Y. -C.**, M. Karlica, J. D. Melon Fuksman, D. Primorac, J. A. Rueda, S. Shakeri, Y. Wang, and S. S. Xue. Relativistic Behavior and Equitemporal Surfaces in Ultra-Relativistic Prompt Emission Phase of Gamma-Ray Bursts. *Astronomy Reports*, 62(12):905-910, Dec 2018.
- R. Ruffini, L. Becerra, C. L. Bianco, **Chen, Y. C.**, M. Karlica, M. Kovačević, J. D. Melon Fuksman, R. Moradi, M. Muccino, G. B. Pisani, D. Primorac, J. A. Rueda, G. V. Vereshchagin, Y. Wang, and S. S. Xue. On the Ultra-relativistic Prompt Emission, the Hard and Soft X-Ray Flares, and the Extended Thermal Emission in GRB 151027A. *The Astrophysical Journal*, 869(2):151, Dec 2018.
- R. Ruffini, R. Moradi, J. A. Rueda, L. Becerra, C. L. Bianco, C. Cherubini, S. Filippi, **Chen, Y. C.**, M. Karlica, N. Sahakyan, Y. Wang, and S. S.

Xue. On the GeV emission of the type I BdHN GRB 130427A. *arXiv e-prints*, page arXiv:1812.00354, December 2018.

- R. Ruffini, J. A. Rueda, R. Moradi, Y. Wang, S. S. Xue, L. Becerra, C. L. Bianco, **Chen, Y. C.**, C. Cherubini, S. Filippi, M. Karlica, J. D. Melon Fuksman, D. Primorac, N. Sahakyan, and G. V. Vereshchagin. The inner engine of GeV-radiation-emitting gamma-ray bursts. *arXiv e-prints*, page arXiv:1811.01839, Nov 2018.
- J. A. Rueda, R. Ruffini, Y. Wang, Y. Aimuratov, U. Barres de Almeida, C. L. Bianco, **Chen, Y. C.**, R. V. Lobato, C. Maia, D. Primorac, R. Moradi, and J. F. Rodriguez. GRB 170817A-GW170817-AT 2017gfo and the observations of NS-NS, NS-WD and WD-WD mergers. *Journal of Cosmology and Astroparticle Physics*, 2018(10):006, Oct 2018.
- R. Ruffini, R. Moradi, J. A. Rueda, Y. Wang, Y. Aimuratov, L. Becerra, C. L. Bianco, **Chen, Y. -C.**, C. Cherubini, S. Filippi, M. Karlica, G. J. Mathews, M. Muccino, G. B. Pisani, D. Primorac, and S. S. Xue. On the role of the Kerr-Newman black hole in the GeV emission of long gamma-ray bursts. *arXiv e-prints*, page arXiv:1803.05476, Mar 2018.
- R. Ruffini, M. Muccino, Y. Aimuratov, M. Amiri, C. L. Bianco, **Chen, Y. - C.**, B. Eslam Panah, G. J. Mathews, R. Moradi, G. B. Pisani, D. Primorac, J. A. Rueda, and Y. Wang. On the Short GRB GeV emission from a Kerr Black hole. *arXiv e-prints*, page arXiv:1802.07552, Feb 2018.
- R. Ruffini, Y. Aimuratov, C. L. Bianco, **Chen, Y. C.**, D. M. Fuksman, M. Karlica, R. Moradi, D. Primorac, J. A. Rueda, N. Sahakyan, and Y. Wang. GRB 180728A: A long GRB of the X-ray flash (XRF) subclass, expecting supernova appearance. *GRB Coordinates Network*, 23066:1, Jan 2018.
- R. Ruffini, Y. Aimuratov, C. L. Bianco, **Chen, Y. C.**, D. M. Fuksman, M. Karlica, R. Moradi, D. Primorac, J. A. Rueda, N. Sahakyan, and Y. Wang. GRB 180720B: Testing the universality of the newly born neutron star in BdHNe. *GRB Coordinates Network*, 23019:1, Jan 2018.
- R. Ruffini, Y. Wang, Y. Aimuratov, U. Barres de Almeida, L. Becerra, C. L. Bianco, **Chen, Y. C.**, M. Karlica, M. Kovacevic, L. Li, J. D. Melon Fuksman, R. Moradi, M. Muccino, A. V. Penacchioni, G. B. Pisani, D. Primorac, J. A. Rueda, S. Shakeri, G. V. Vereshchagin, and S. S. Xue. Early X-Ray Flares in GRBs. *The Astrophysical Journal*, 852(1):53, Jan 2018.

*Dedicated to
my sunshine, Song*

Abstract

Seyfert galaxies have several subclasses according to observation features of their optical spectra. Traditional classification is usually visual inspection or using a quantity that is defined as a flux ratio between the Balmer line and forbidden line. However, visual inspection is time-consuming and a quantity derived from flux ratio does not reflect the information of emission line shape which is the original classification feature. One algorithm of deep learning is called Convolution Neural Network (CNN) and has shown several successful classification results. We use a CNN model to classify spectra of the Seyfert galaxies by building a 1-dimension CNN model as a classifier to distinguish Seyfert 1.9 galaxies from Seyfert 2 galaxies. We find our model can recognize Seyfert 1.9 and Seyfert 2 galaxies correctly with an accuracy of over 80%. Our model picks out an additional Seyfert 1.9 sample which was missed by visual inspection among Seyfert 2 samples in a short time. We also use the new Seyfert 1.9 sample to improve the performance of our model and obtain 91% precision of the Seyfert 1.9. This result indicates that our model can act as a classifier for picking out Seyfert 1.9 spectra among Seyfert 2 spectra in an efficient way for the future.

The second part of the thesis is about the emission line of our Seyfert 1.9 galaxies. We decompose the $H\alpha$ emission line into narrow component and broad component by fitting 2 Gaussian components. We use the fitting results to derive line flux and line width to investigate the properties of emission lines of our Seyfert 1.9 galaxies. We find the 641 Seyfert 1.9 sample has an averagely luminous broad $H\alpha$ component than the 656 Seyfert 1.9 sample. This result indicates that our model can pick out sources with relatively weak broad $H\alpha$ component and these sources are sometimes missed by visual inspection. We investigate the black hole mass, host galaxy morphology, BPT diagram of the two Seyfert 1.9 samples. We find that the distributions of supermassive black hole (SMBH) mass of our Seyfert 1.9 galaxies are dominated by $10^7 M_\odot$. We find the two Seyfert 1.9 samples have slightly different black hole mass distributions but the t-test results indicate that the two distributions have similar mean values. From host galaxy results, we find the distribution of the host galaxy morphology of the Seyfert 1.9 galaxies is dominant by large bulge galaxy. We find the BPT distributions of our Seyfert 1.9 samples are similar to Seyfert 2 galaxies. This result indicates that Seyfert 1.9 and Seyfert 2 galaxies have similar ionization levels in their narrow-line region. Besides, we cross-search the Strip 82 survey with our Seyfert 1.9 sources for multiple spectra to investigate the possible variability. We only have 13 Seyfert 1.9 sources that have multiple spectra and find 13 Seyfert 1.9 sources show variation in continuum and broad $H\alpha$ component.

Ringraziamenti

I thank the Secretaria in ICRANet for helping me in adapting to life in Italy when I arrived in Italy. I thank Cinzia for helping me in every moment and your kindness always warms me up.

I thank Yu-Ling Chang for taking care of me when I was first in Italy. It is so great to have a person that is from Taiwan when I was first in Italy.

I thank Wang Yu and Rahim for their guidance in the physics discussion. I always learned a lot from you. I enjoyed our coffee time in the afternoon and the strolling time in Pescara.

I thank my friends in Taiwan. Thank you for coming to Italy. Your visit always gave me a lot of power and made me feel I was back in Taiwan. This always encourages me to finish my degree. I also thank the friends who always gave me a call when I was feeling down about my work or life.

The last one whom I want to thank is my dear Song. Thank you for accompanying me. Your warmness gives me the courage to overcome the hardest time and finish the dissertation.

Contents

1 Introduction	1
1.1 Active Galactic Nuclei	1
1.2 Classification of Seyfert galaxies	1
1.3 Unification model of Seyfert galaxies	4
1.4 Applying Deep learning on astrophysics	6
2 Data Selection and Analysis	9
2.1 Data Selection	9
2.2 Convolution Neural Network model	10
2.3 Fitting H α emission line	15
3 Results	17
3.1 Convolution Neural Network Result	17
3.2 Emission line of Seyfert 1.9	22
3.3 Variability of Seyfert 1.9 galaxies	30
4 Discussion	49
4.1 Test result	49
4.2 compared to intermediate Seyfert	52
4.3 Supermassive Black Hole Mass	53
4.4 Fitting H α with consideration of [NII] λ 6548 and [NII] λ 6584	55
4.5 Host galaxy morphology	60
4.6 compare to selection method	61
5 Summary	65
A Fitting results of our 13 sources	67
A.1 Do not consider [NII] λ 6548,6584 in fitting process	67
A.2 Consider [NII] λ 6548,6584 in fitting process	67
B 1D CNN model code (Pytorch)	95

Chapter 1

Introduction

1.1 Active Galactic Nuclei

A few galaxies show strong emission activity in their nucleus than most galaxies in the universe. These galaxies are called Active Galactic Nuclei (AGNs). An AGN is believed to consist of an accretion disk and a supermassive black hole (SMBH). This structure is surrounded by an optically thick torus in the galaxy center and is considered as a phenomenon that is explained by accreting material around the supermassive black hole into the galaxy center [28]. Since emission is across all electromagnetic spectrum, different wavelengths have its' main emission mechanism and corresponding to different components of the AGNs.

For example, γ -ray is the highest energy band and Blazar, which is a subclass of the AGNs, is the most AGN source in γ -ray observation [9]. Blazar has a characteristic of a jet emerging from the vicinity of SMBH pointing to an observer at a small angle. The energy output is dominated by non-thermal radiation and explained by a bulk of particles with relativistic velocity toward an observer. The emission is regarded as Doppler boosting effect and the emission mechanism of a jet is still unclear. X-ray has the advantage of transparency and can penetrate dust and gas to investigate the inner region of the accretion disk. In optical wavelength, it usually traces emission from the accretion disk to photoionize gas surrounding the center and these sources are Seyfert galaxies and Quasar (quasi-stellar object, QSO) [32, 31]. In the infrared band (IR), the emission is from dust that is outside the scale of the Broad-line region (BLR) to re-emission radiation from the accretion disk and is related to obscuring material in the line of sight. In radio band, its emission is dominated by synchrotron emission and it is to trace magnetic field. The sources selected from the radio band are radio galaxies [8]. AGNs have various observation properties among different wavelengths. In this work, we focus on the Seyfert galaxies, which are one type of AGNs selected from the optical wavelength. The different Seyfert types are defined from their characteristic on optical spectra.

1.2 Classification of Seyfert galaxies

Seyfert galaxies were spiral galaxies with a bright nucleus and have emission lines on their optical spectra [32]. According to the characteristics of the spectra, Seyfert

galaxies have two major classes; one is Seyfert 1 galaxies and the other one is Seyfert 2 galaxies [17]. The difference between these two types of the Seyfert galaxies is related to the line width of emission lines on their optical spectra; Seyfert 1 galaxies have the characteristic of broad Balmer emission lines and the line width is usually greater than 1000 km/s whereas Seyfert 2 galaxies have narrow Balmer emission line and the line width is usually less than 1000 km/s [17]. Besides the two main classes, a few Seyfert galaxies share emission line features between type 1 and type 2 Seyfert; the Balmer line has a narrow component superimposed on a broad component. These sources are intermediate Seyfert galaxies [26]. Intermediate Seyfert galaxies are further classified into Seyfert 1.2, Seyfert 1.5, Seyfert 1.8, and Seyfert 1.9 depending on strength of the broad $H\beta$ component [23, 24]. We show example spectra of the Seyfert types in Fig. 1.1. The strength of the $H\beta$ component becomes weak from Seyfert 1.2 to Seyfert 1.9 galaxies. Seyfert 1.2 galaxies have a strong broad $H\alpha$ and $H\beta$ component with a small peak of the narrow component. Seyfert 1.5 galaxies have comparable broad and narrow components of $H\alpha$ and $H\beta$ with a slightly strong [OIII] peak value. Seyfert 1.8 galaxies have a weak broad component superimposed a strong narrow component in $H\alpha$ and $H\beta$ emission lines. Seyfert 1.9 galaxies have only a weak broad $H\alpha$ component superimposed a strong narrow component without a broad $H\beta$ component. Besides, we notice that the authors discarded Seyfert 1.2 and only use Seyfert 1.5 for the composite spectrum [25].

Since various shapes of the emission lines of the intermediate Seyfert galaxies, it is not easy to classify and pick up spectra by an automatic pipeline among huge observation data and the classification process by vision inspection usually consumes a lot of time. Some authors used a quantity to describe different types of the intermediate Seyfert galaxies and we show summarization in table 1.1 [45, 46]. Both of the methods use a flux ratio of $H\beta$ and [OIII] to describe the intermediate Seyfert galaxies. This method is fast and simple to give an approximate quantity value of the intermediate Seyfert types. However, a flux of an emission line usually represents the intensity or strength of the photon while a line width of an emission line represents the velocity distribution of the cloud due to Doppler broadening. The two parameters, flux, and line width are not always dependable. For example, a weak broad line has a large line width but a small flux while a strong narrow line has a large flux and a small line width. A quantity, which was defined from line ratio, does not always reflect the real difference between the multi-component profile of the $H\alpha$ and $H\beta$ emission line to the forbidden emission line although it is a fast and convenient method for picking the Seyfert sample. Furthermore, the Seyfert 1.9 galaxies were not assigned any quantity and are hard to be picked up from huge observation data. Lacking selection parameters and the complicated line shape cause the Seyfert 1.9 galaxies hard to be picked out and collected to large samples for advanced analysis. On the other hand, the weak broad component of the $H\alpha$ of the Seyfert 1.9 galaxies sometimes causes contamination in a sample of the Seyfert 2 spectrum.

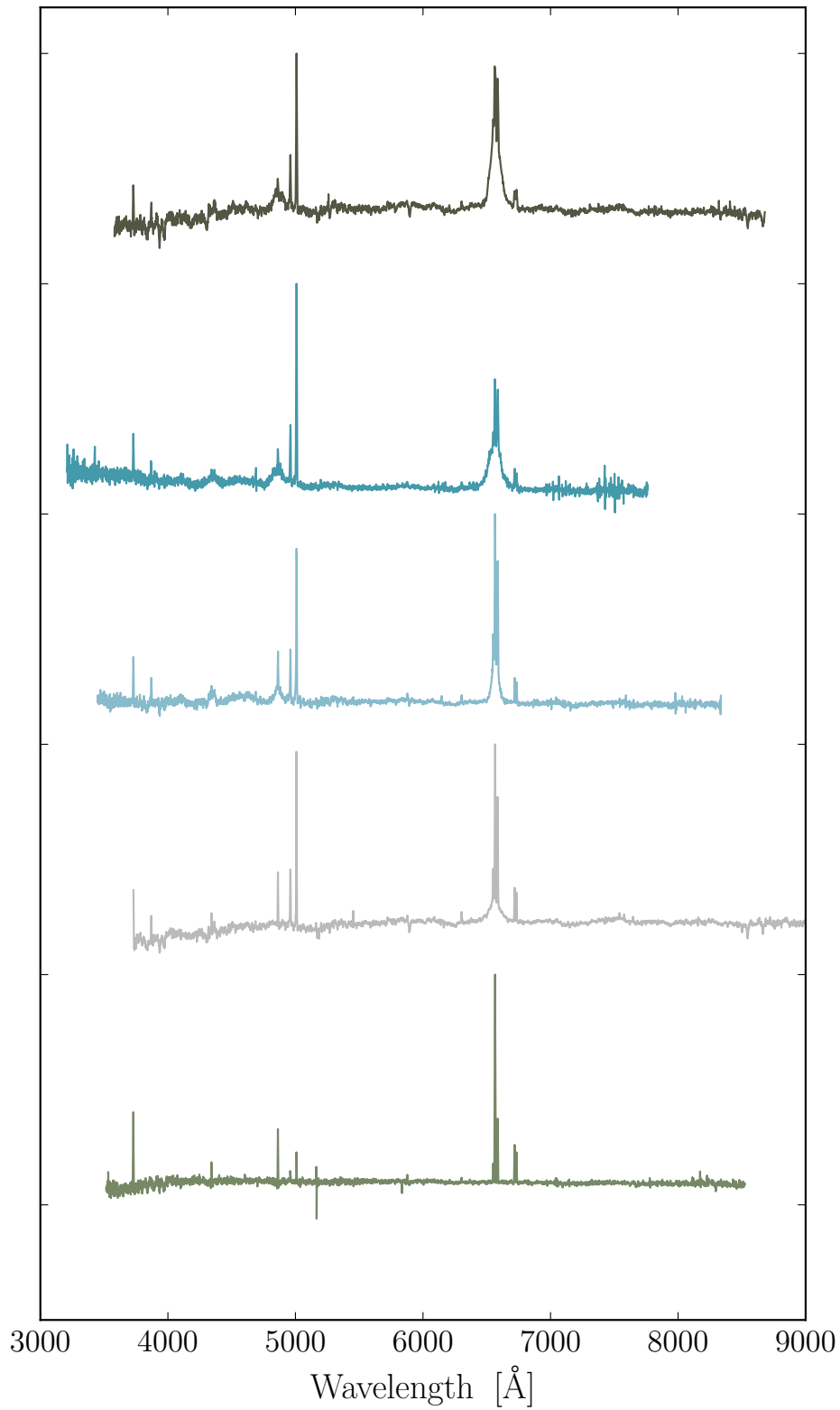


Figure 1.1. Example spectra of different Seyfert types. From top to bottom is Seyfert 1.2, Seyfert 1.5, Seyfert 1.8, Seyfert 1.9, and Seyfert 2.

Table 1.1. Comparison of the Seyfert category

	Whittle (1992)	Winkler (1992)
parameter R	$\frac{F_{[\text{OIII}]5007}}{F_{\text{H}\beta}}$	$\frac{F_{\text{H}\beta}}{F_{[\text{OIII}]5007}}$
Seyfert 1.0	$R \leq 0.3$	$R > 5$
Seyfert 1.2	$0.3 < R \leq 1$	$2 < R < 5$
Seyfert 1.5	$1 < R \leq 4$	$0.333 < R < 2$
Seyfert 1.8	$R > 4$	$R < 0.333$
Seyfert 1.9	weak broad H α	weak broad H α

1.3 Unification model of Seyfert galaxies

AGNs have diverse observation properties among different wavelengths. Therefore, it is asked that if the various observation properties of the AGNs are coming from the same physic mechanism or not. There is a model called the unification model trying to unify all AGNs into a general picture. This concept is a supermassive black hole (SMBH) with an accretion disk in the galaxy center and this structure is embedded in an opaque torus. The parameter in this model is orientation and the orientation which is blocked or not blocked by obscure material can explain the diversity observation characteristic of the AGNs [29]. A unification model indicates that the Seyfert 1 and Seyfert 2 galaxies are intrinsically the same, the only difference is due to the different viewing angle relative to torus [3, 40]. We show the conception of the unification model of AGNs in Fig. 1.2. Seyfert 1 galaxies are viewed directly to the central supermassive black hole (SMBH) whereas Seyfert 2 galaxies are viewed from the edge and the emission from the accretion disk is blocked by the torus. This model is supported by strong evidence that the polarization spectrum of a Seyfert 2 galaxy, NGC1068, showing a broad component [4].

However, the intermediate Seyfert galaxies have characteristics of type 1 and type 2 Seyfert and these intermediate sources are explained by either dust reddening or intrinsic weak ionization. In the frame of the unification model, it is explained that the line of sight passes the edge of the torus in a specified location. This idea is supported by the observation characteristic of Seyfert 1.8/1.9 galaxies, which show that the characteristic is consistent with dust reddening of the broad-line region (BLR) and continuum [24]. The other one is explained by the intrinsic ionization. NGC 2992 shows a variation on the weak broad H α component and the variation is not consistent with dust reddening; the variation is suggested to be caused by intrinsic low ionize state [37]. Infrared observation has the advantage of transparency to dust and can detect if dust exists in these sources or not. The measurement of Pa β emission line for 10 intermediate Seyferts suggests that some Seyfert 1.8/1.9 are caused by BLR reddening while the other Seyfert 1.8/1.9 have low optical depth and small ionization parameter [10]. Besides, a sample of Seyfert 1.8 and 1.9 galaxies that are analyzed by multi-wavelength shows ambiguous results; three sources are dust reddening of BLR, one is in low ionization, and four are inconclusive in a sample of eight Seyfert 1.9 galaxies while three Seyfert 1.8 sources are all in low ionization state [38]. These results indicate that mechanisms behind intermediate Seyfert galaxies

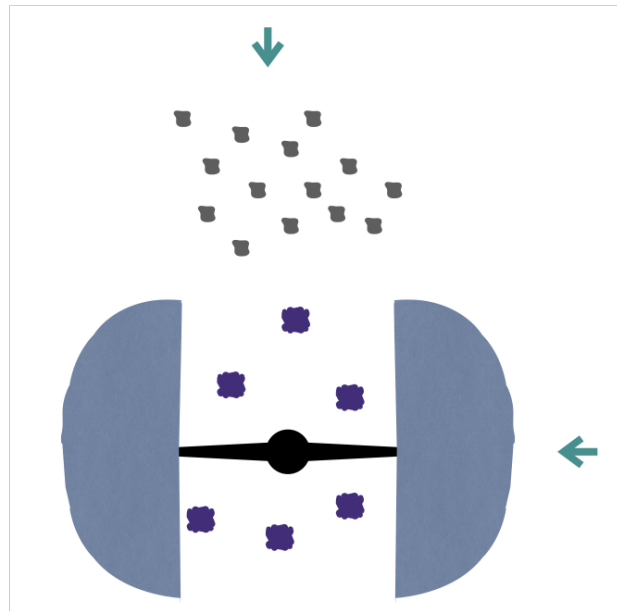


Figure 1.2. Concept of the unification model of AGNs. Dark purple represents the broad-line region (BLR). Grey represents the narrow-line region (NLR). Dark blue represents the torus. Black represents the supermassive black hole and the accretion disk.

which have two components of the $H\alpha$ and $H\beta$ emission line simultaneously are still under debate and unclear.

To investigate the emission mechanism behind intermediate Seyfert galaxies, it is necessary to collect more intermediate Seyfert sources. However, the characteristic of two components on the spectra makes the intermediate Seyfert galaxies hard to be picked up from huge amounts of observation data, especially Seyfert 1.9 galaxies, which did not be assigned any quantity. The other way to pick up Seyfert 1.9 galaxies is to stack spectra of the Seyfert 2 to check if the Seyfert 2 spectra are contaminated by Seyfert 1.9 galaxies. However, these methods shown above are inefficient when dealing with huge observation data. A traditional method for picking out Seyfert 1.9 spectra is usually visual inspection. This process is time-consuming and has a potential bias for being inspected by different people. These reasons above show the method for picking up Seyfert 1.9 among huge observation data is difficult and is still not solved well until now.

1.4 Applying Deep learning on astrophysics

The rapid development of the technology accompanies more great telescope and sky surveys to bring more information about the Universe for us, e.g., the Sloan Digital Sky Survey [48, SDSS], and the ROSAT all-sky survey [43, RASS]. This brings abundant observation data and makes the astronomy and astrophysics community into a big data era. Besides, it brings a new challenge in processing the huge amount of observation data. The development of technology benefits the field of computer science and provides more powerful computation ability than before. The development is also accompanied by large amounts of data and algorithms. Powerful computation ability, large amounts of data, and algorithms take Artificial Intelligence (AI) into the next stage. Machine learning is a subset of AI and it is an algorithm to make the machine have the ability to make a decision by learning from data. Machine learning has been noticed by the astronomy and astrophysics community and there are several application cases by astrophysics. For example, applying machine learning on distinguishing star-forming galaxies and AGNs without the dominant characteristic of $H\alpha$ and [NII] emission lines [37, 50].

One subset of machine learning is deep learning. The difference between machine learning and deep learning is the ability to fix problems. Machine learning needs a guide from the programmer when it returns an inaccurate result. Deep learning has the ability to fix the problem by itself. Besides, machine learning learns from the whole pattern of data distribution while deep learning learns from the intrinsic pattern of the different types of data. The other advantage of deep learning is that deep learning can deal with massive data well than machine learning. The basic unit of deep learning is a neural network. The characteristic of deep learning is more than one layer in the model. That is it can combine different algorithms inside a model for solving the problems. Deep learning has several algorithms depending on the structure and arrangement of the neural network and one of the famous algorithms is Convolution Neural Network (CNN). This algorithm has a specific filter to convolute with input data and recognizes features between the different types of input data. A famous and successful classification case is to classify 1.2 million

images into 1000 classes with reducing error to 15% by adding more convolution layers [18]. There is an example of applying deep learning on astrophysics and it is to determinate the stellar abundance from stellar spectra without fitting spectra with synthetic libraries, which is slow and hard to deal with low SNR spectra [19]. This example shows that deep learning provides a new and fast analysis method for the astrophysics field than before.

In early times, most observation data are usually classified by visual inspection. This classification process is slow and time-consuming when the data amount is large. However, we have another choice to deal with observation data by benefiting from the powerful computation ability now. We can apply deep learning to the huge observation data to help us process huge data fast and correctly. In this thesis, we build a 1-dimension convolution neural network to collect a large sample of the Seyfert 1.9 galaxies from Seyfert 2 galaxies. By collecting and analyzing a big sample of the Seyfert 1.9 galaxies could help us understand the physic mechanism behind Seyfert 1.9 galaxies. Besides, we decompose the $H\alpha$ emission line of the Seyfert 1.9 galaxies and compare the line properties of the Seyfert 1.9 to Seyfert 2 galaxies. From the large sample of the Seyfert 1.9 galaxies, we cross-search multiple spectra of our Seyfert 1.9 sources from the Stripe 82 survey to investigate if our Seyfert 1.9 sources have variation and the variation is related to intrinsic weak state or dust reddening. We present the results of the spectral variability of our Seyfert sources in the following. The last part is the discussion and summary. In this thesis, we used $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_m = 0.3$, $\Lambda_0 = 0.7$, $q_0 = -0.55$, $k = 0.0$.

Chapter 2

Data Selection and Analysis

In this chapter, I will explain how I select our sample first and the following is how a Convolution Neural Network (CNN) works and how I analyze our sample with our customize CNN model. The last part is to decompose $H\alpha$ emission line of our Seyfert 1.9 sources by fitting a two-Gaussian.

2.1 Data Selection

We extract our Seyfert sample from [7] and show the criteria in the following:

- $S/N_{[\text{OIII}]}$ > 5
- $S/N_{H\alpha}$ > 3
- $\log L_{[\text{OIII}]} [\text{ergs/s}] > 40.125$
- $0 < z < 0.2$

The original Seyfert 2 sample has two distributions in their $L_{[\text{OIII}]}$ [7]. The authors fit a two-Gaussian to represent the two distributions. The cut value of $L_{[\text{OIII}]}$ is an overlap of the two Gaussian distributions. We only select the luminous $[\text{OIII}]$ population because it usually has luminous $H\alpha$ emission due to the photoionization. The other reason is that the high $H\alpha$ luminosity of the Seyfert 2 might be contaminated by intermediate Seyferts. The $S/N_{[\text{OIII}]} > 5$ is to make sure sufficient accuracy of measurement. The $S/N_{H\alpha} > 3$ is to maximize the sample and include the weak sources. In the end, we have 54694 Seyfert 2 and 745 intermediate Seyfert galaxies.

We obtain spectra of our Seyfert sources from the SDSS Data Release 10 [11, DR10]. We notice 745 intermediate Seyfert galaxies include the Seyfert 1.2, 1.5, 1.8, and Seyfert 1.9 galaxies. Thus, we inspect the $H\beta$ emission line of the 745 sources and find 104 sources that have a broad $H\beta$ component. We use two training sets. Training set 1 uses 745 intermediate Seyfert and 56,494 Seyfert 2 sources. Training set 2 uses 641 Seyfert 1.9 and 56,494 Seyfert 2 sources. In order to learn the nature feature between intermediate Seyfert/Seyfert 1.9 and Seyfert 2 galaxies instead of their distribution, we choose the same number for Seyfert 1.9 and Seyfert 2 as our training data set. In training set 1, We use 300 intermediate Seyfert and 300 Seyfert 2 samples as our training data and the rest 445 intermediate Seyfert and 54394

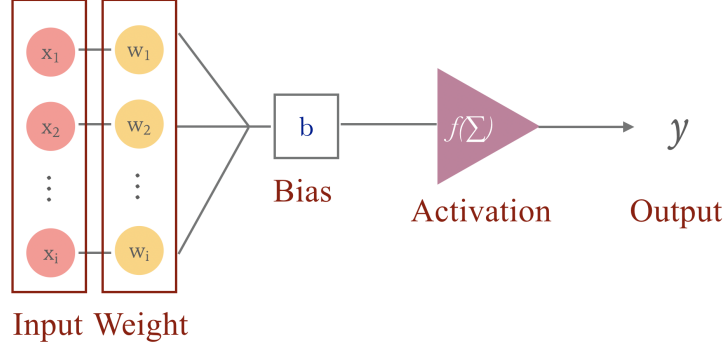


Figure 2.1. A neuron.

Seyfert 2 as our test data. In training set 2, we use 300 Seyfert 1.9 and 300 Seyfert 2 samples as our training data and 341 Seyfert 1.9 and 54394 Seyfert 2 as test data. We summarize all detail numbers used in different training sets in Table [3.1](#).

2.2 Convolution Neural Network model

A neuron is a basic unit in a neural network. It can receive information, process it, and output the information in the end. Fig. [2.1](#) shows how a neuron works and we show a mathematical form below.

$$y = f\left(\sum_i x_i w_i + b\right) \quad (2.1)$$

where x_i is input, w is weight, b is bias, f is activation function, and y is output. The output of a neuron is applying an activation function to the summation of the dot result of the input value and weight and a bias value. An activation function is a nonlinear function and can be regarded as a learning behavior of the neuron. Because the results of the dot product and bias are linear mapping, the linear result will have unlimited information. By applying the activation function, the nonlinear function makes the neuron decide if it outputs the results or not. This is regarded as a switch, on and off, in a neuron network. In this work, we use the Rectifier Linear Unit (ReLU) as our activation function and show the ReLU function in Fig. [2.2](#). This is a common activation function in a neural network and it turns all negative input to 0 and keeps positive values unchanged.

The convolution neural network (CNN) is an algorithm of deep learning. Deep usually means more than one layer in a model. We show the architecture of a multi-layer neural network in Fig. [2.3](#). In principle, a neural network has a input layer, several hidden layers, and an output layer in the end. For a convolution neural network, its basic structure is a convolution layer, pooling layer, and fully-connected layer. Each layer has its algorithm and function. We explain the details about different layers in a CNN in the following.

- Convolution layer: A filter with a specific size will move along a direction of input data to convolute with input data. We show an example below to explain the concept.

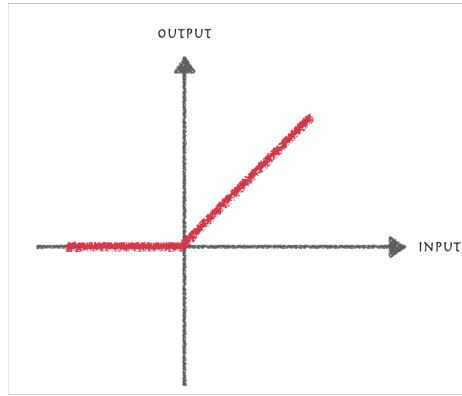


Figure 2.2. Rectifier Linear Unit (ReLU) function

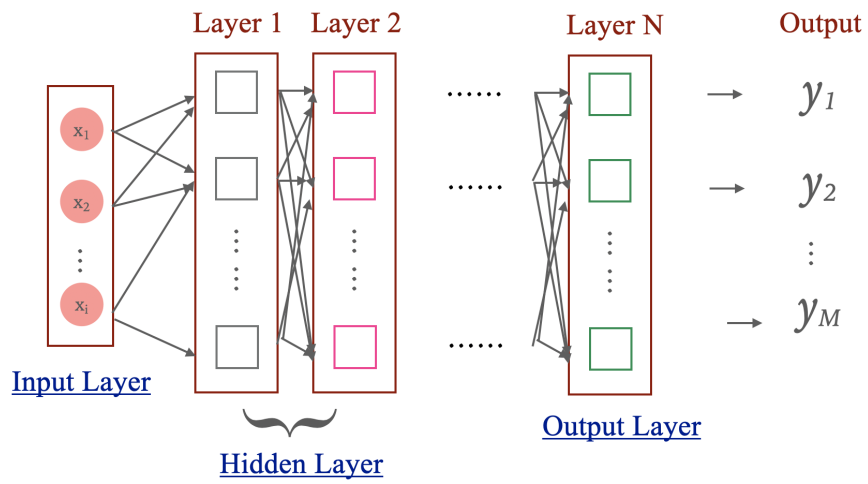
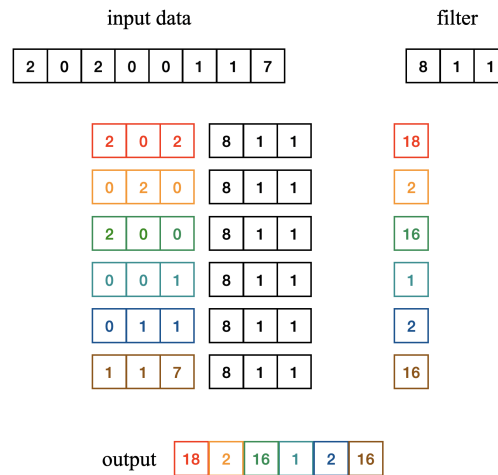
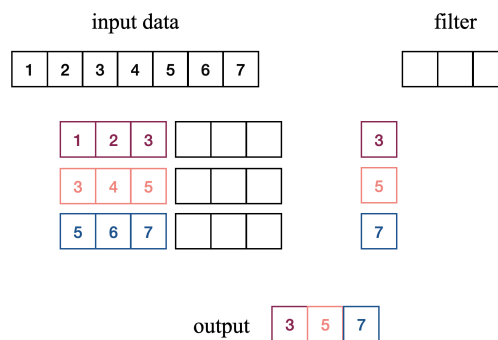


Figure 2.3. A neural network.



In this example, input data has length = 8 and the filter has length = 3. The filter has different weight values and will move from left to right to convolute with input data. The different colors represent each convolution between data and filter. The filter will move one step every time and this is called stripe = 1. The first step is red and the convoluted result is 18. In this case, the filter will move 6 times to have 6 results and the output length is 6. The convolution between input data and weight value is regarded as extracting features from input data.

- Pooling layer: This layer works to retain dominant features only and reduces data length. This is to reduce the parameters of the model and make computation fast. There are several pooling algorithms and we use max-pooling in our model. We show an example below to explain the max pooling.



In this example, input data has length = 7 and the filter size = 3. The filter will move from left to right. In this case, the stripe is 2. The different colors correspond to each step. For every step, only the maximum value will remain. For the first step, the maximum value is 3 and it will output 3 only. In this case, the output length reduces to 3 and only remains the dominant feature.

- Fully-connected layer: This layer is placed before the output layer. The layer will flatten all input neurons and make all neurons in different layers be

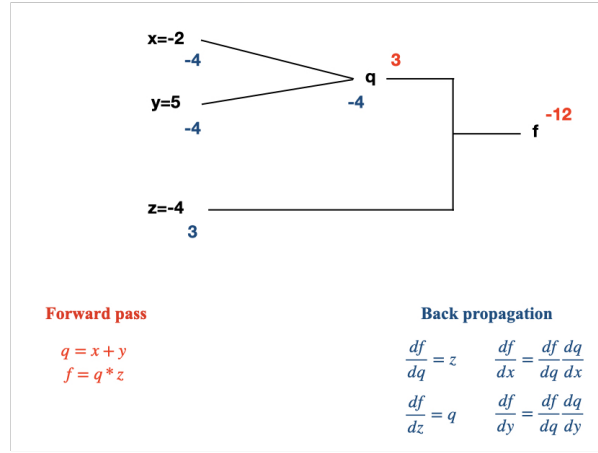


Figure 2.4. Example for explaining the forward pass and backpropagation.

connected but the neurons in the same layer are not connected. This means that every feature extracted from previous layers will be combined in this layer and a fully-connected layer can identify this information. Then, this layer will output a probability of a target class.

In a CNN model, there are two passways. One is forward pass and it means data information will be transmitted from low layer to high layer. The other one is backpropagation. It combines chain rule and gradient. Its function is to transmit gradient which is output with respect to input value in each neuron from the high layer back to the low layer. We show an example for explaining forward pass and backpropagation in Fig. 2.4

Besides data pass, we further describe parameters in a CNN model and the situation of the model during a fitting in the following.

- **Loss:** The loss function is to describe how the difference between a set of parameters of a model and output data. The lower loss represents the difference between output and the model is small and means the prediction of the model is good. The higher loss represents the difference is large and indicates that the model prediction is bad. In this work, we use the Cross-Entropy Loss because it is usually useful for a binary or multiple classification case. We show its form below:

$$L(\mathbf{p}, \mathbf{y}) = - \sum_{\mathbf{n}} \mathbf{y}_{\mathbf{n}} \log(\mathbf{p}_{\mathbf{n}}) \quad (2.2)$$

y represents label or expected output and p represents the predicted probability of the output category.

- **Optimizer:** This is a process of finding the best parameter of the model that obtains the minimize loss. To find the best parameter efficiently, we can

use *gradient*. Calculating the gradient of loss can give the right direction in parameter space and the procedure of repeating calculating the loss value to update the best parameter is called *gradient descent*.

$$w' = w - \eta \frac{\partial L}{\partial w} \quad (2.3)$$

where w' is updated weight from w , η is the learning rate, and L is loss. The learning rate is the step size for every update in the parameter space. For a small step, the progress will update slowly. For a large step, the progress will go fast but unstable. Therefore, it is important to set an appropriate step in training, and this value is usually determined by experience. We use the Stochastic Gradient Descent (SGD) in this work. This method will randomly select a sample as a start point to calculate the gradient descent instead of whole data. The gradient will be updated by a small sample selected randomly from the whole data in each iteration. This will make calculations fast when the data amount is huge.

Table 2.1. 1D CNN architecture

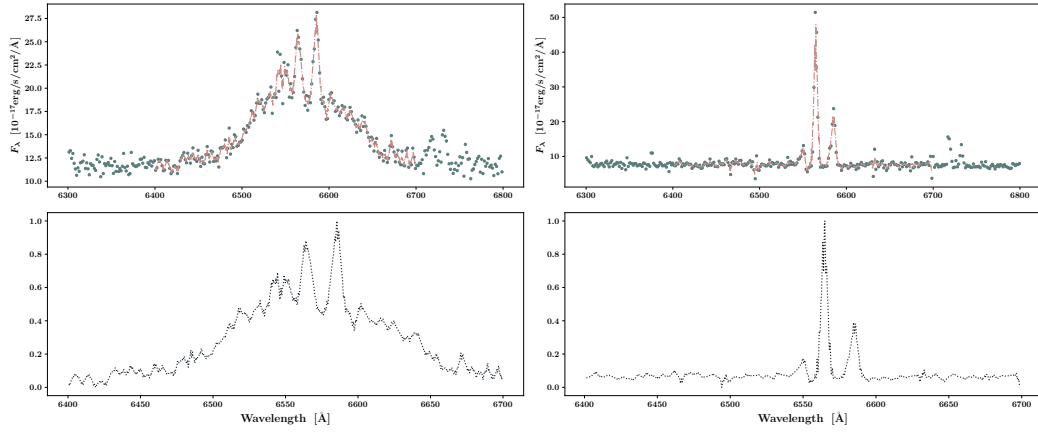
Layer	Type	Channel _{input}	Channel _{output}	Kernel size	Stride	Activation
1	Convolution	1	10	200	1	ReLU
2	Pooling			2	2	
3	Convolution	10	30	100	1	ReLU
4	Pooling			5	2	
5	Convolution	30	36	56	1	ReLU
6	Linear	3384	94			ReLU
out	Linear	94	2			

We build a 1-dimension (1D) CNN which includes 6 layers and 1 output layer and summarize all details of the model in Table 2.1. The 1D means that we only input flux values of the spectra and the data is in the form of a 1D array. The parameters in a CNN model are weight and bias and we show parameter numbers in each layer in Table 2.2. The weight number in a convolution layer is obtained by multiplying the kernel size by the channel number of output and the bias number is the same as the channel number of output. Therefore, our first convolution layer has 2000 weights and 10 bias values. For the sixth layer, weight number is estimated by multiplying the number of output channels by the number of input channels and bias is the same as the output channel number. Thus, the sixth layer has 318096 weight and 94 bias. The total trainable parameter in our model is to sum all numbers of weight and bias. Therefore, our model has 410936 parameters in total. Our model is set up and runs on Pytorch and the detail code are shown in Appendix B.

We only input a segment of the spectra. The segment ranges from 6400Å to 6700 Å because the characteristic of H α emission is inside this region. All spectra are shifted into the rest-frame. We interpolate data points to 1000 points inside the region to make the observation data more continuous. Since flux values in different sources are various, the different features have different value ranges and

Table 2.2. Parameter number in our CNN model

Layer	weight	Bias
1	2000	10
2		
3	30000	30
4		
5	60480	36
6	318096	94
out	188	2

**Figure 2.5.** Example spectrum and its input format. Left is Seyfert 1.9 galaxy. Right is Seyfert 2 galaxy. Top: steel represents the observed data and pink represents the data interpolation to 1000 data points in the range from 6400Å to 6700Å. Bottom: normalized spectrum for input.

this inconsistent range among our sources would make our model confused. In order to learn the shapes of the emission lines on the spectra, we normalize every spectrum to its peak value to make sure every spectrum scale to the same range. This normalization makes the value range of every spectrum from 0 to 1. We show an example of the observed Seyfert 1.9 and Seyfert 2 spectrum with the normalized one in Fig. [2.5](#).

2.3 Fitting H α emission line

Since the H α emission line of the Seyfert 1.9 galaxies has a weak broad component superimposed a strong narrow component, we fit two Gaussian components to represent the narrow and broad component. The range of fitting wavelength is from 6450Å to 6700Å. First, we mask all emission lines in the region and only used the region of $\lambda < 6500\text{Å}$ and $\lambda > 6650\text{Å}$ to find continuum level by fitting a linear component as the continuum part. Second, we only mask the region of [NII] λ 6548 and [NII] λ 6583 emission lines and left the H α emission line to fit 2 Gaussian components with a linear component which is derived in the previous step

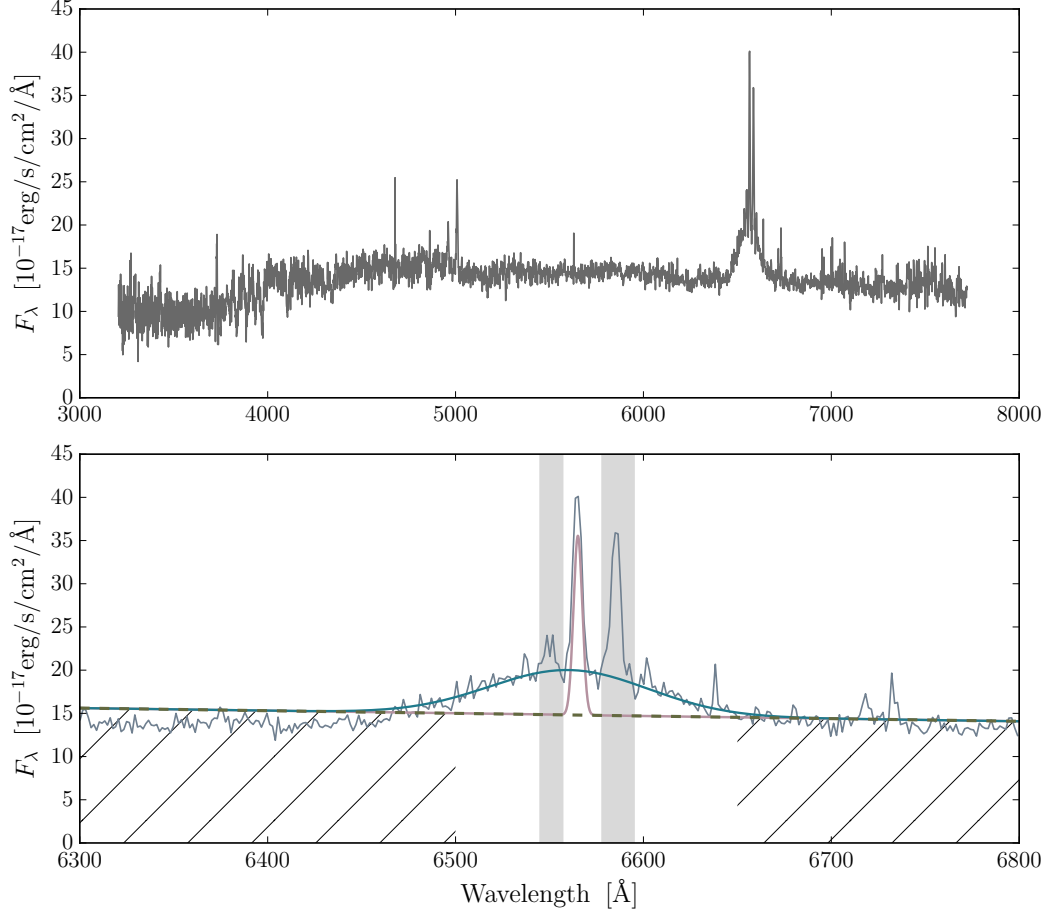


Figure 2.6. Example spectrum for fitting H α emission line. Top: whole spectrum. Bottom: spectrum of the fitting region. The hatch region is $\lambda < 6500\text{\AA}$ and $\lambda > 6650\text{\AA}$ and used for fitting continuum. Dashed-line represents the fit continuum. Pink represents the narrow component while blue represents the broad component. The grey areas are [NII] λ 6548 and [NII] λ 6583 and masked in the fitting process.

and will be fixed in the fitting to the spectrum by Levenberg-Marquardt algorithm and least squares. We show an example spectrum of our fitting in Fig. 2.6.

After obtaining parameters of Gaussian fitting, we estimate the full-width at half-maximum (FWHM) of the two components from fitting results. The FWHM is $2\sqrt{2\ln 2} \sigma$. Furthermore, we estimate the flux of two components from fitting results by integrating the area. The flux error is derived from error propagation. We also derive H α luminosity of the broad and narrow component from the flux results.

Chapter 3

Results

In this chapter, we show results from our CNN model. The following is about flux and luminosity derived from the decomposition of $H\alpha$ emission line of our Seyfert 1.9 galaxies. In the end, we present the variability results of our Seyfert 1.9 sample.

3.1 Convolution Neural Network Result

A loss value represents the difference between a predicted output from the model and the target. We show results of training set 1 of our CNN model in Fig. 3.1 and Table 3.1. We find that the loss of train and valid process decreases with epoch. The decreased loss with epochs represents the difference between the output from the model and target. The decreasing loss shows that our model finds better parameters in our model and outputs the results close to target during the fitting process. This indicates that our model learns well and obtains the best parameters. This result suggests a good fitting result. After 100 epochs, the loss of train is 1.847462 and the loss of valid is 1.361852. In order to know the classification ability of our training model, we estimate the accuracy value for our sample. The accuracy value is estimated by the classified correct sources divided by total sources. From our result, we find that all accuracy values are increased with the epoch after the 40th epoch. The valid accuracy after the 100th epoch is 98%. The value represents Seyfert 1.9 and Seyfert 2.0 during every epoch. For more detail, we check the precision of the two types of Seyferts. For Seyfert 1.9, the precision reaches 87% after 100 epochs. This represents that our model classifies correctly 389 Seyfert 1.9 samples out of 445 Seyfert 1.9 samples after iterating 100 epochs. For Seyfert 2, the precision is 98% after iterating 100 epochs and that is our model classifies correctly 53628 Seyfert 2 samples out of 54394 Seyfert 2 samples.

From Fig. 1.1, it shows that the ratio between broad and narrow $H\alpha$ components in intermediate Seyferts are different. The peak of the broad $H\alpha$ component is comparable to that of narrow $H\alpha$ in Seyfert 1.2 galaxies while the peak of broad $H\alpha$ is much smaller than that of the narrow $H\alpha$ component. A ratio of peak flux density between broad and narrow $H\alpha$ components is indicated in literature [36]. The authors indicated that a Seyfert 1.5 galaxy has the ratio = 0.5 while a Seyfert 1.9 galaxy has the ratio = 0.05. From the results of training set 1, we notice the training set is trained by intermediate Seyfert sources. This means that we have

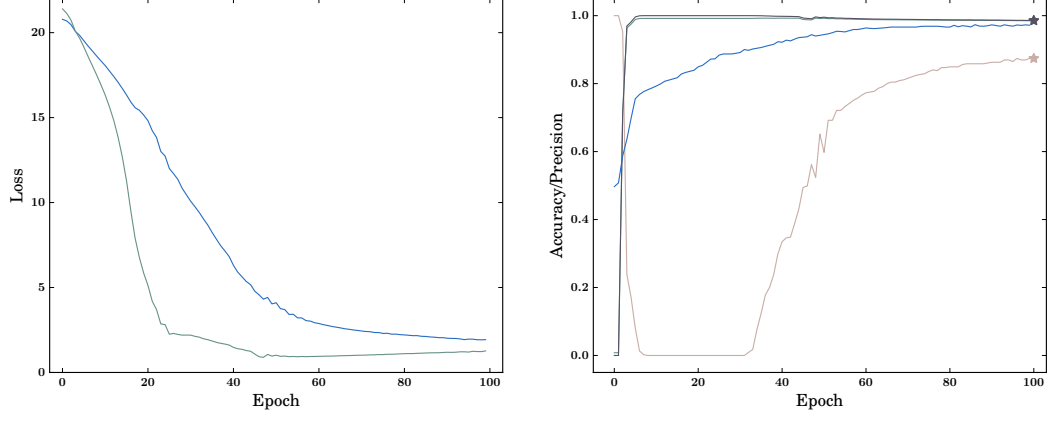


Figure 3.1. Loss value and accuracy as a function of epoch for the training set 1. **Left:** Loss versus epoch. **Right:** Accuracy/precision versus epoch. Blue represents the training process and green represents the valid process.

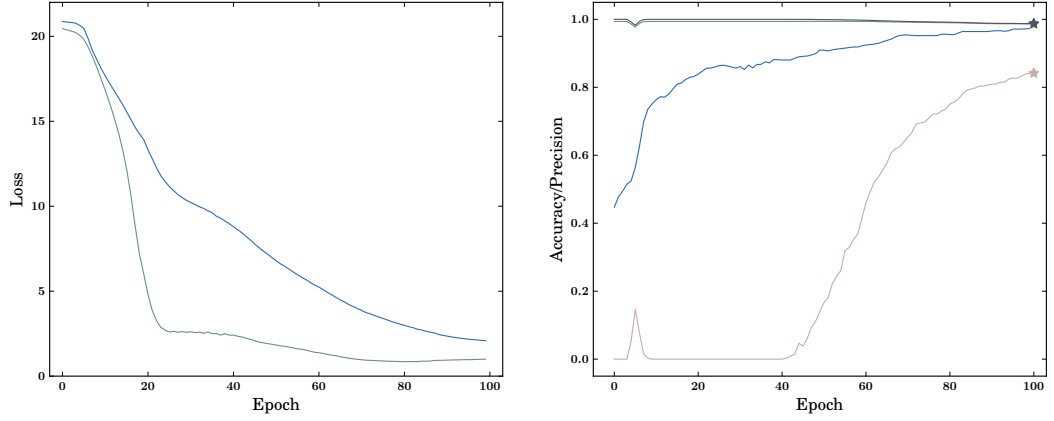


Figure 3.2. Loss value and accuracy value as a function of epoch for the training set 2. **Left:** Loss versus epoch. **Right:** Accuracy/precision versus epoch. Blue represents the training process and green represents the valid process. Light pink represents Seyfert 1.9 and dark purple represents Seyfert 2

more sources that have a slightly strong broad $H\alpha$ component in this training set. We would like to know if the model trained by intermediate Seyfert galaxies can recognize Seyfert 1.9 spectra which have a relatively weak broad $H\alpha$ component. We apply the training set 1 model to the 2nd test sample. This time, we have 641 pure Seyfert 1.9 galaxies and 54694 Seyfert 2 galaxies as our test sample. The source number and test results are summarized in Table 3.1. The total accuracy for this test is 98% and that is our model can recognize 54481 sources correctly from 55335 sources. For Seyfert 1.9, the precision is 86% and this means the model of training set 1 can classify correctly 553 Seyfert 1.9 sources from 641 Seyfert 1.9 sources. For Seyfert 2, our model can discern 53928 Seyfert 2 sources correctly from 54694 Seyfert 2 sources. This result indicates that although our training set 1 model is trained by a stronger broad $H\alpha$ sample, it still can recognize the pure Seyfert 1.9 which has a weak broad $H\alpha$ component among Seyfert 2 galaxies with an precision of 86%. In addition, we notice although the model can classify most Seyfert 2 sources correctly, it remains 766 Seyfert 2 galaxies classified as Seyfert 1.9 galaxies in this test.

We inspect the spectra of 766 Seyfert2 galaxies which are classified as Seyfert 1.9 galaxies by the training set 1 model. We find 709 spectra are intermediate Seyfert, 54 spectra are either damaged or have no emission, and 3 inclusive spectra. The results indicate that our model does have the ability to pick out a characteristic of two-component on a spectrum among Seyfert 2 galaxies. This method provides a way to pick out intermediate Seyfert galaxies fast and can be applied to future observation data. We also inspect the 709 intermediate sources and find 656 Seyfert 1.9 galaxies. In this work, we obtain a new sample of Seyfert 1.9 galaxies and provide an additional 656 Seyfert 1.9 galaxies for further analysis in the next section.

We show results of training set 2 in Fig. 3.2 and Table 3.1. In this set, we use 641 Seyfert 1.9 sources as the two-component source to train our model and test if the model still can discern the characteristic of a weak broad component of Seyfert 1.9 galaxies among the Seyfert 2 galaxies. We use 300 Seyfert 1.9 and Seyfert 2, respectively, as our training sample. The remaining Seyfert 1.9 sources reduce to 341 as the test data source. We use batch size=30 and learning rate=0.01 as the initial parameter and this set is the same in training set 1. We find the loss values are decreased with epochs steadily and converge to a value. The loss value of training is 2.086748 and that of valid is 0.997698 after 100 epochs. The accuracy values of training and test are almost steady and the final accuracy for this test is 98%, indicating that our model classifies 54000 sources correctly from 54735 sources. The precision of the Seyferts increases with the epoch after the 50th epoch. The precision of the Seyfert 2 is 98% after 100 epochs, indicating that our model can classify correctly 53713 Seyfert 2 among 54394 Seyfert 2 sources. The precision of Seyfert 1.9 after 100 epochs is 84% and that is to say, this model can classify 287 Seyfert 1.9 correctly from 341 Seyfert 1.9 galaxies. This indicates that our training model can still discern the pure Seyfert 1.9 galaxies from the Seyfert 2 galaxies.

Since we obtain additional 656 Seyfert 1.9 galaxies, we have more Seyfert 1.9 sources than the previous data set. One way to increase accuracy is to add more training sources. Therefore, we set the third training set. This time, we have 1297 Seyfert 1.9 spectra. We remove the spectra of damage or no emission from the Seyfert 2 sample and have 53928 Seyfert 2 spectra. We double the training sample number and have 600 Seyfert 1.9 and 600 Seyfert 2 galaxies respectively. The detail

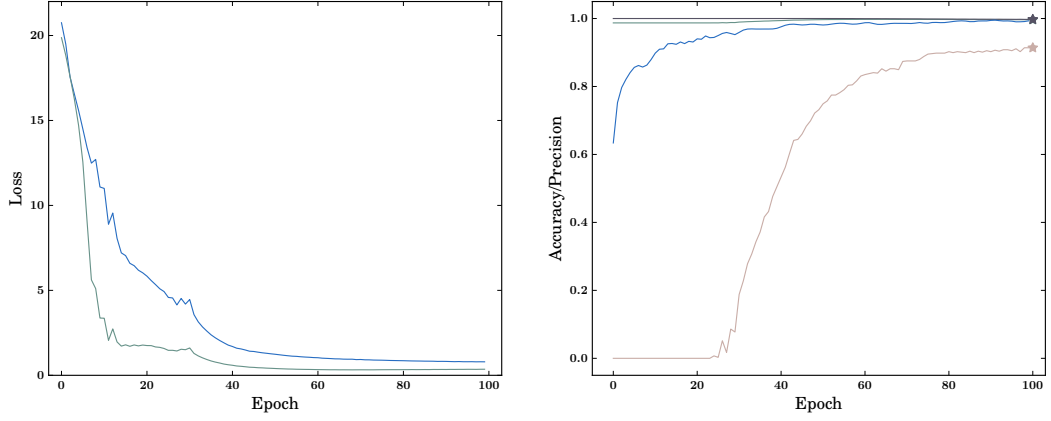


Figure 3.3. Loss value and accuracy value as a function of epoch for the training set 3. **Left:** Loss versus epoch. **Right:** Accuracy/precision versus epoch. Blue represents the training process and green represents the valid process. Light pink represents Seyfert 1.9 and dark purple represents Seyfert 2

sources number for training and test are summarized in Table. 3.1. We show results of training set 3 in Fig. 3.3 and Table. 3.1. We find that the loss values of training and valid decrease with epoch and converge to a constant value after epoch = 60. This plateau indicates our model reaches its global minimum and the fitting result is good. The accuracy of the test sample is 99% and it means our model can classify 53827 sources correctly among 54025 sources. For classifying Seyfert 1.9 galaxies, our model has a precision of 91% and it suggests our model can classify 637 Seyfert 1.9 sources correctly among 697 Seyfert 1.9 sources. For Seyfert 2, our model has a precision of 99% for classifying 53190 Seyfert 2 sources correctly among 53328 Seyfert 2 sources.

Table 3.1. Summary of training and test results

		Training sample	Test sample	Precision	Accuracy	2 nd	Test sample	Precision	Accuracy
Training 1	Sey1.9	300	445	87%	98%	54694	641*	86%	98%
	Sey2	300	54394	98%			54694	99%	
Training 2	Sey1.9	300*	341*	84%	98%				
	Sey2	300	54394	98%					
Training 3	Sey1.9	600*	697*	91%	99%				
	Sey2	600	53328	99%					

Note: * means the pure Seyfert 1.9 galaxies. That is to say these spectra do not have broad H β emission line.

We show convolution results of the spectra in Fig. 3.4. Since the length of input data is 1000, our filter size in the first layer is 200. The output length of our data after the first convolution layer reduces to 800. We set 10 output channels in this layer and this can be regarded as ten sets of weight values to convolute with our input spectra. In Fig. 3.4, we can see the weight values respond to input spectra. For example, the ninth channel in the first convolution layer is more significant in Seyfert 1.9 than that in Seyfert 2. This is how our model learns the difference between the two samples. The results show that the Seyfert 1.9 has more wide characteristics than the Seyfert 2 in the 9th channel.

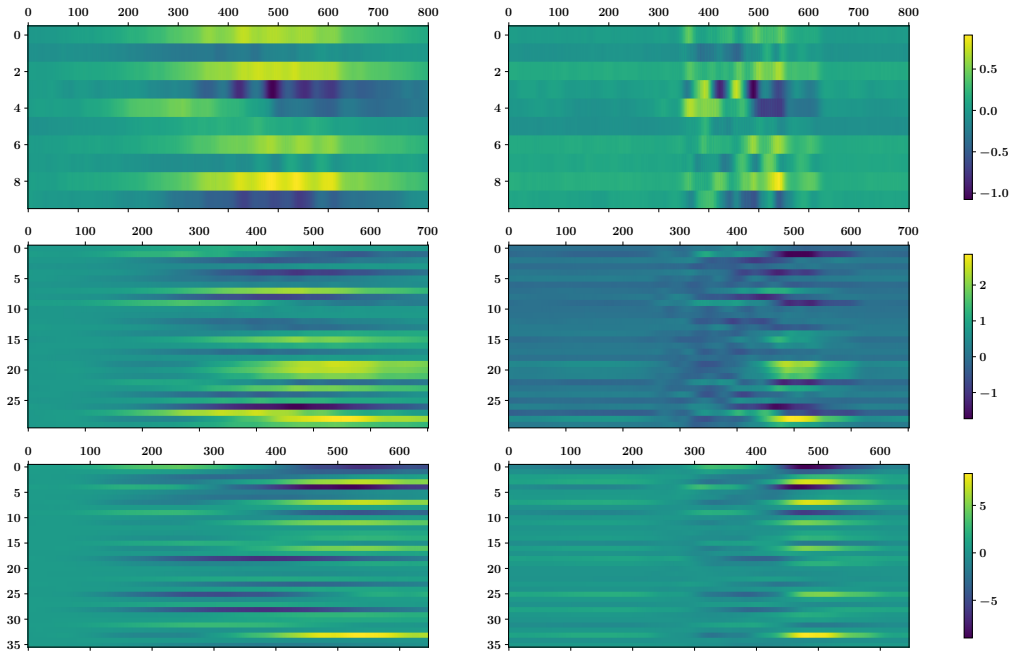


Figure 3.4. Convolution results of input spectra. Left is Seyfert 1.9 and right is Seyfert 2 galaxy. The original observation spectra of the two examples are shown in Fig. 2.5. Top: convolution results after the first convolution layer. Middle: convolution results after the third convolution layer. Bottom: convolution results after the fifth convolution layer. In each plot, the y-axis is the number of output channels and the x-axis is data length.

3.2 Emission line of Seyfert 1.9

In this work, we have original 641 Seyfert 1.9 and 656 Seyfert 1.9 that are selected by our training model. In brief, we have a total of 1297 Seyfert galaxies. We decompose the $H\alpha$ emission line of the 1297 Seyfert 1.9 galaxies to estimate the line width of the broad and narrow components. We show the distribution results of FWHM in Fig. 3.5. For narrow component distributions, the two samples of our Seyfert 1.9 show similar distribution while the 656 Seyfert 1.9 sample has a small peak at 200-300 [km/s]. For broad component distributions, the peak of the 641 Seyfert 1.9 sample is slightly higher than that of the 656 Seyfert 1.9 sample while the 656 Seyfert 1.9 sample has an extending tail toward the higher end. We further estimate the mean value and standard deviation of the FWHM distribution of the two samples and show the results in Table 3.2. We find that the 656 Seyfert 1.9 sample has a higher mean value of $L_{H\alpha, \text{broad}}$ and $L_{H\alpha, \text{narrow}}$ than the 641 Seyfert 1.9 sample. We estimate a t-test to check if the mean values of the FWHM are different in two Seyfert 1.9 samples. For the $\text{FWHM}_{H\alpha, \text{broad}}$, the test result is t-statistic = -2.46 with a p-value = 0.01. The p-value is the probability that the two samples are drawn from the same population with the same mean values. The significant p-value indicates that we can reject the null hypothesis that the two populations have the same mean values. This result suggests that the 641 Seyfert 1.9 sample has a different mean value of $\text{FWHM}_{H\alpha, \text{broad}}$ from the 656 Seyfert 1.9 sample. From the distribution of the $\text{FWHM}_{H\alpha, \text{broad}}$, the high velocity is usually from the inner part of the accretion disk in the supermassive black hole. This result indicates that the 656 Seyfert sample picked out by our model includes the sources with emissions near the central black hole.

The distribute of narrow component of 1297 Seyfert 1.9 galaxies ranges from 100 to 1000 [km/s] with a median value 340 [km/s] while the broad component distribution of 1297 Seyfert 1.9 galaxies from 1000 to 10000 [km/s] with a median value 3783 [km/s]. These results show that there is a significant boundary between the broad and narrow component and using 1000 [km/s] as a division for the narrow and broad component is adequate [17]. This result shows that there is an empty region between BLR and NLR. This region lacks any emission line and was explained by existing dust which suppresses the emission line [22].

We also estimate the $H\alpha$ luminosity of narrow and broad components and show the distribution results in Fig. 3.6. The distribution of narrow component of both Seyfert 1.9 sample spans from 10^{39} to 10^{42} [ergs/s] with a mean value $\approx 10^{40}$ [ergs/s] while the distribution of broad component spans from 10^{40} to 10^{43} [ergs/s] with a mean value $\approx 10^{41}$ [ergs/s]. We find that the luminosity of the broad component has one order larger than that of the narrow component in our Seyfert 1.9 samples. We show the mean values and standard deviations in Table 3.2. We estimate the t-test for the broad $H\alpha$ luminosity distributions of the two Seyfert 1.9 samples and the t-statistic = 4.44 with a p-value = 10^{-5} . This result suggests that the 656 Seyfert 1.9 sample has averagely weaker broad $H\alpha$ luminosity than the 641 Seyfert 1.9 sample. Additionally, we notice that although the 656 Seyfert 1.9 sample has higher mean values of $\text{FWHM}_{H\alpha}$ than the 641 Seyfert 1.9 sample, the mean value of $H\alpha$ luminosity is slightly smaller than the 641 Seyfert 1.9 sample. This result indicates that a source with a broad component is not necessary to have a high

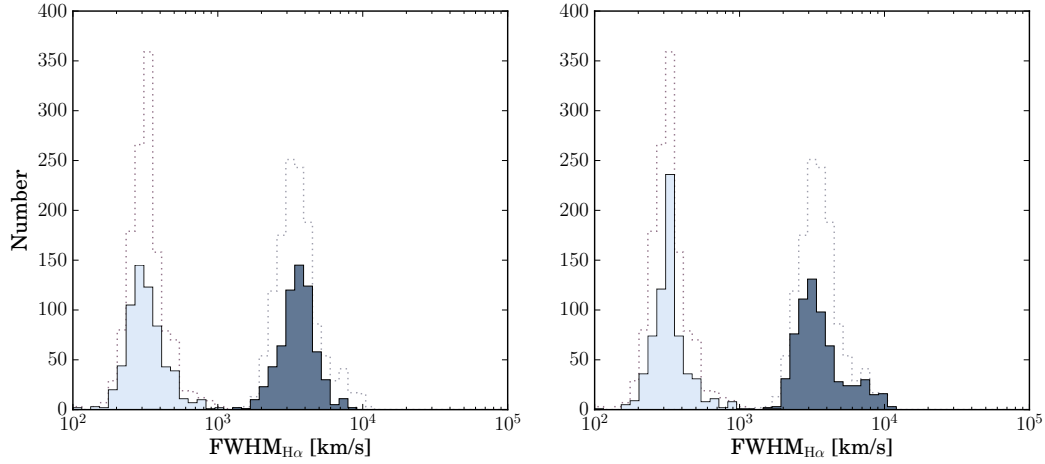


Figure 3.5. FWHM distributions of 1297 Seyfert 1.9 sources. Left: 641 Seyfert 1.9 sources. Right: 656 Seyfert 1.9 sources. Dot-line represents the total 1297 Seyfert 1.9 sources. Light blue represents the narrow component. Dark blue represents the broad component.

Table 3.2. Mean values of the Seyfert 1.9 samples

	1297 Sy19	641 Sy19	656 Sy19	
$\text{FWHM}_{\text{H}\alpha, \text{broad}}$	3783.59 (1473.92)	3682.09 (1060.36)	3882.76 (1782.30)	[km/s]
$\text{FWHM}_{\text{H}\alpha, \text{narrow}}$	340.79 (128.17)	336.09 (116.70)	345.39 (138.32)	[km/s]
$L_{\text{H}\alpha, \text{broad}}$	3.84×10^{41} (4.90×10^{41})	4.45×10^{41} (5.87×10^{41})	3.25×10^{41} (3.63×10^{41})	[ergs/s]
$L_{\text{H}\alpha, \text{narrow}}$	8.60×10^{40} (1.45×10^{41})	9.47×10^{40} (1.41×10^{41})	7.75×10^{40} (1.47×10^{41})	[ergs/s]
$L_{\text{H}\alpha}$	4.70×10^{41} (6.1×10^{41})	5.40×10^{41} (7.09×10^{41})	4.03×10^{41} (4.87×10^{41})	[ergs/s]

The values inside parenthesis are standard deviations of the distributions.

luminosity. Therefore, using flux or luminosity to pick out a sample that has a two-component characteristic on its emission line is only a rough method and might lose some sources to cause bias. In order to know the whole energy of $\text{H}\alpha$ emission line, we sum the flux of broad and narrow components to estimate the luminosity and show the result in Fig. 3.7. The mean values and standard deviations are in Table 3.2. The distribution of total $\text{H}\alpha$ luminosity of 1297 Seyfert 1.9 galaxies ranges from 10^{40} to 10^{43} [ergs/s]. We find the 641 Seyfert 1.9 sample has a slightly high mean value of total $\text{H}\alpha$ luminosity than the 656 Seyfert 1.9 sample.

There is an empirical correlation between Balmer line luminosity and continuum luminosity and this correlation suggests that the emission mechanism is photoionization [47, 33, 12]. We use $L_{5100\text{\AA}}$, which was estimated by averaging the luminosity from 5050 \AA to 5150 \AA in the rest frame, as continuum luminosity and show the $L_{\text{H}\alpha}$ of broad and narrow component versus $L_{5100\text{\AA}}$ for our Seyfert 1.9 galaxies in Fig. 3.8. We also estimate correlation coefficient between $\text{H}\alpha$ luminosity and $L_{5100\text{\AA}}$ and show the results in Table 3.3. We find both 641 Seyfert 1.9 and 656 Seyfert 1.9 samples show a correlation between $\text{H}\alpha$ luminosity and continuum luminosity for both broad and narrow component. The correlation coefficient for $L_{\text{H}\alpha, \text{narrow}}$ and $L_{5100\text{\AA}}$ is 0.26 with a p-value = 2.8×10^{-22} and for $L_{\text{H}\alpha, \text{broad}}$ and $L_{5100\text{\AA}}$ is 0.36 with a p-value = 5.07×10^{-42} for total 1297 Seyfert 1.9 galaxies. Our results show that

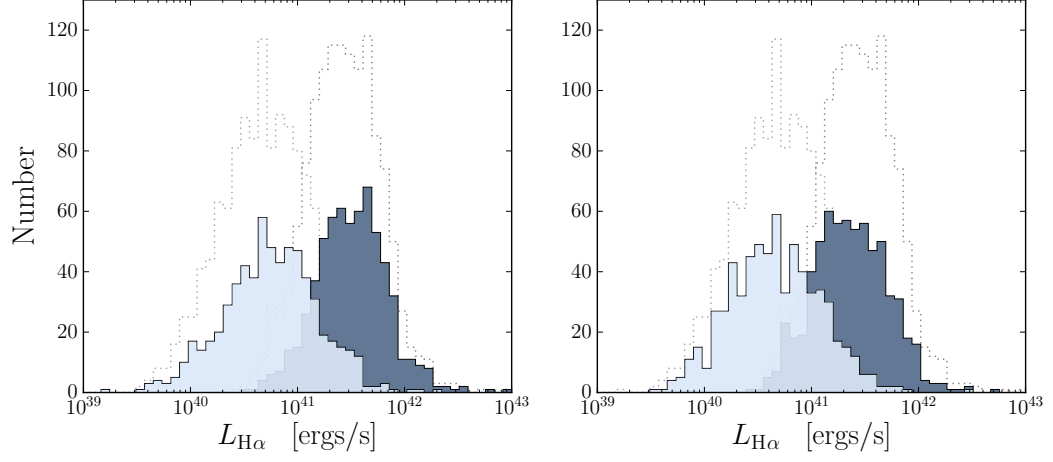


Figure 3.6. $H\alpha$ luminosity distributions of our Seyfert 1.9 sources. Left: 641 Seyfert 1.9 sources. Right: 656 Seyfert 1.9 sources. Dot-line represents the total 1297 Seyfert 1.9 sources. Light blue represents the narrow component. Dark blue represents the broad component.

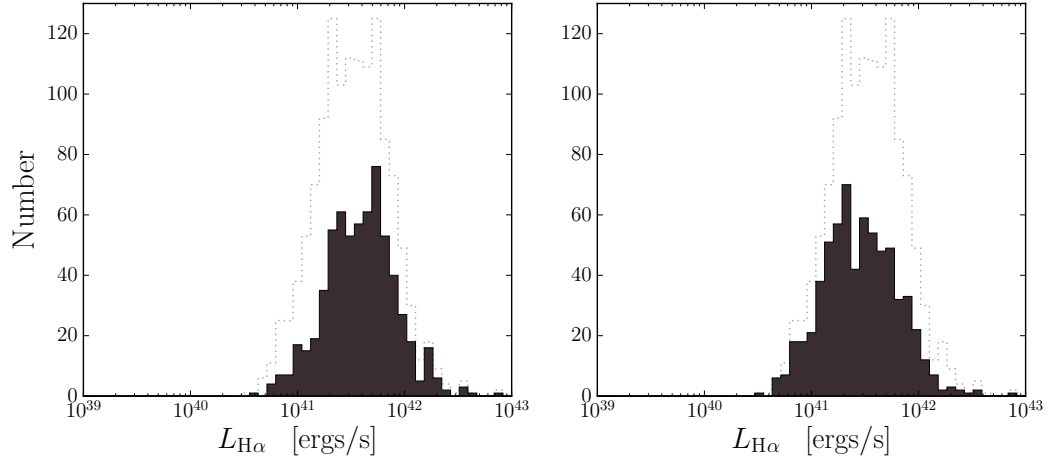


Figure 3.7. $H\alpha$ luminosity distributions of our Seyfert 1.9 sources. Left: 641 Seyfert 1.9 sources. Right: 656 Seyfert 1.9 sources. Dot-line represents the total 1297 Seyfert 1.9 sources.

Table 3.3. Correlation coefficient of the Seyfert 1.9 sample

	$L_{\text{H}\alpha, \text{broad}}, L_{5100\text{\AA}}$		$L_{\text{H}\alpha, \text{narrow}}, L_{5100\text{\AA}}$		$L_{\text{H}\alpha}, L_{5100\text{\AA}}$	
	r	p-value	r	p-value	r	p-value
1297 Sy19	0.36	5.07×10^{-42}	0.26	2.80×10^{-22}	0.36	7.20×10^{-40}
641 Sy19	0.40	1.13×10^{-25}	0.34	1.63×10^{-18}	0.40	1.65×10^{-25}
656 Sy19	0.35	4.68×10^{-20}	0.20	3.28×10^{-7}	0.32	5.84×10^{-17}

Table 3.4. Correlation coefficient of $L_{\text{H}\alpha}, L_{5100\text{\AA}}$ in different redshift ranges for 1297 Seyfert 1.9 galaxies

	$0 < z < 0.05$		$0.05 < z < 0.1$		$0.1 < z < 0.15$		$0.15 < z < 0.2$	
Number	22		297		504		474	
	r	p-value	r	p-value	r	p-value	r	p-value
$L_{\text{H}\alpha, \text{broad}}, L_{5100\text{\AA}}$	0.60	3.68×10^{-3}	0.39	3.8706×10^{-12}	0.09	5.36×10^{-2}	0.40	3.95×10^{-19}
$L_{\text{H}\alpha, \text{narrow}}, L_{5100\text{\AA}}$	0.27	0.23	0.19	1.05×10^{-2}	-0.01	0.77	0.30	1.62×10^{-11}
$L_{\text{H}\alpha}, L_{5100\text{\AA}}$	0.59	3.90×10^{-3}	0.35	3.29×10^{-10}	0.06	0.15	0.39	1.48×10^{-18}

the $\text{H}\alpha$ luminosity of both broad and narrow component have moderate correlation with $L_{5100\text{\AA}}$ and suggest that the $\text{H}\alpha$ emission of broad and narrow component are related to central continuum power, the luminous continuum have more luminous emission. We also notice that the broad component has slightly stronger correlation with $L_{5100\text{\AA}}$ than the narrow component. If the luminosity of all lines is related photoionization, the correlation between the summation luminosity of broad and narrow component and continuum would correlate better. We show the total $\text{H}\alpha$ luminosity with $L_{5100\text{\AA}}$ in Fig. 3.9 and the correlation coefficient in Table 3.3. We find that both 641 Seyfert 1.9 sample and 656 Seyfert 1.9 sample show moderate correlation in the total $\text{H}\alpha$ luminosity and $L_{5100\text{\AA}}$. We also find the correlation coefficient of total $\text{H}\alpha$ luminosity with $L_{5100\text{\AA}}$ for 1297 Seyfert 1.9 galaxies is 0.36 with a p-value= 7.2×10^{-40} . We also estimate the correlation coefficient of $L_{\text{H}\alpha}$ and $L_{5100\text{\AA}}$ for 641 Seyfert 1.9 and 656 Seyfert 1.9 sample and show the results in Table. 3.3. We find that $L_{\text{H}\alpha, \text{broad}}$ to $L_{5100\text{\AA}}$ and total $\text{H}\alpha$ luminosity to $L_{5100\text{\AA}}$ have moderate correlation coefficient.

We further check if the correlation between $\text{H}\alpha$ luminosity and $L_{5100\text{\AA}}$ is affected by redshift. We divided the redshift range by a segment = 0.05. We show the distributions in Fig. 3.10 and the source number in each redshift range and correlation coefficient in Table 3.4. We find the narrow $\text{H}\alpha$ luminosity and $L_{5100\text{\AA}}$ show a high p-value and very weak correlation in the redshift range from 0 to 0.05 and 0.1 to 0.15. For the broad $\text{H}\alpha$ luminosity versus $L_{5100\text{\AA}}$ in redshift 0.1 to 0.15 shows a very low correlation with a probability $\approx 5\%$. For total $\text{H}\alpha$ luminosity versus $L_{5100\text{\AA}}$ in redshift range from 0.1 to 0.15 shows a very weak correlation with a probability $\approx 15\%$. These results indicate that the correlation between $\text{H}\alpha$ luminosity and $L_{5100\text{\AA}}$ are affected by redshift.

We show FWHM and luminosity of $\text{H}\alpha$ emission line in Fig. 3.11. We estimate the correlation coefficient for this distribution. For the narrow $\text{H}\alpha$, the correlation coefficient $r = 0.25$ with a p-value = 2.0481×10^{-19} . For the broad $\text{H}\alpha$, the correlation

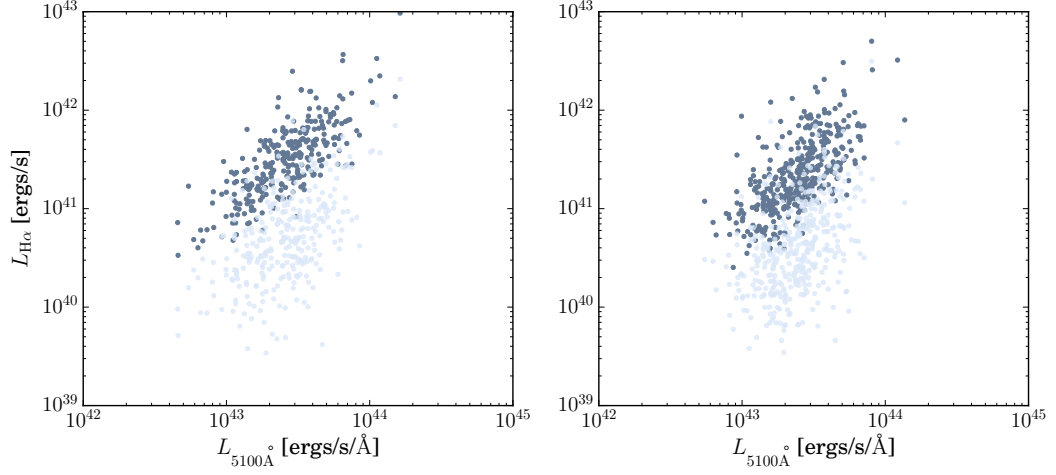


Figure 3.8. $L_{H\alpha}$ versus $L_{5100\text{\AA}}$ luminosity of our Seyfert 1.9 sources. Left: 641 Seyfert 1.9 sources. Right: 656 Seyfert 1.9 sources. Light blue represents the narrow component. Dark blue represents the broad component.

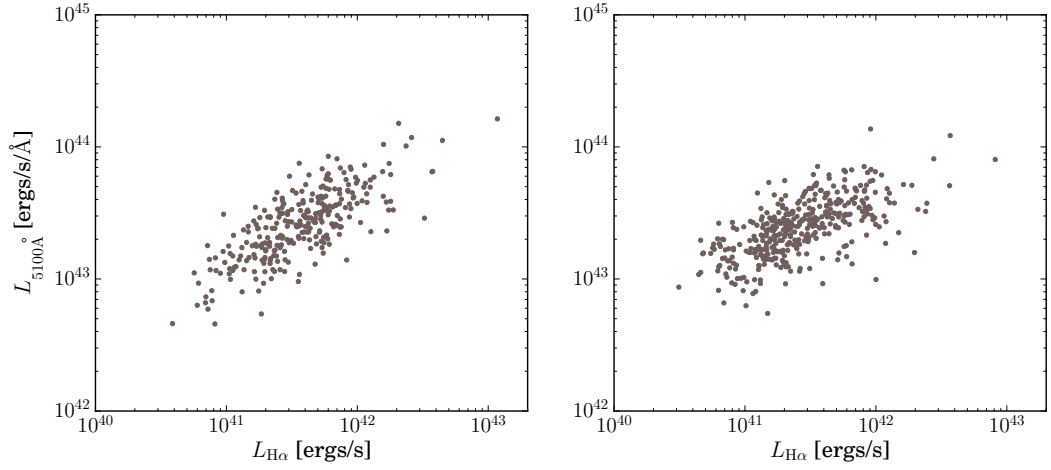


Figure 3.9. $H\alpha$ luminosity versus $L_{5100\text{\AA}}$. Left: 641 Seyfert 1.9 sources. Right: 656 Seyfert 1.9 sources.

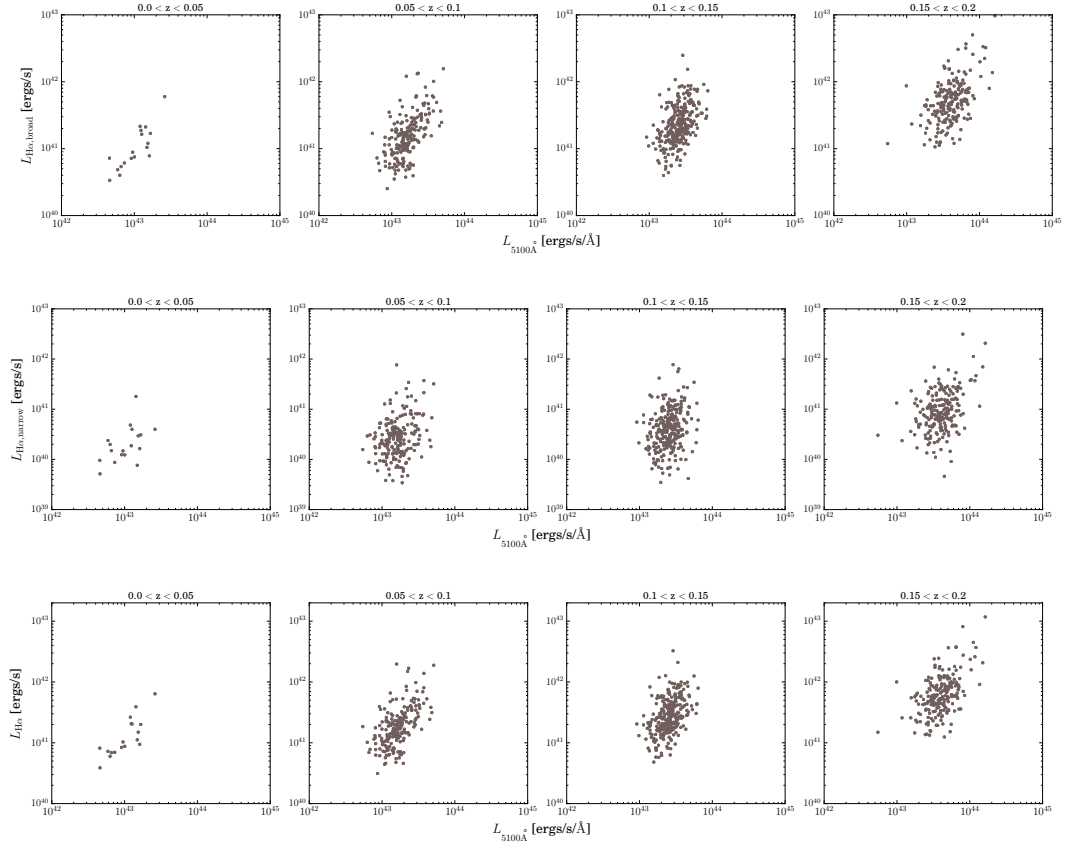


Figure 3.10. $L_{5100\text{\AA}}$ versus $\text{H}\alpha$ luminosity in different redshift ranges. Top: $L_{5100\text{\AA}}$ versus broad $\text{H}\alpha$ luminosity. Middle: $L_{5100\text{\AA}}$ versus narrow $\text{H}\alpha$ luminosity. Bottom: $L_{5100\text{\AA}}$ versus total $\text{H}\alpha$ luminosity.

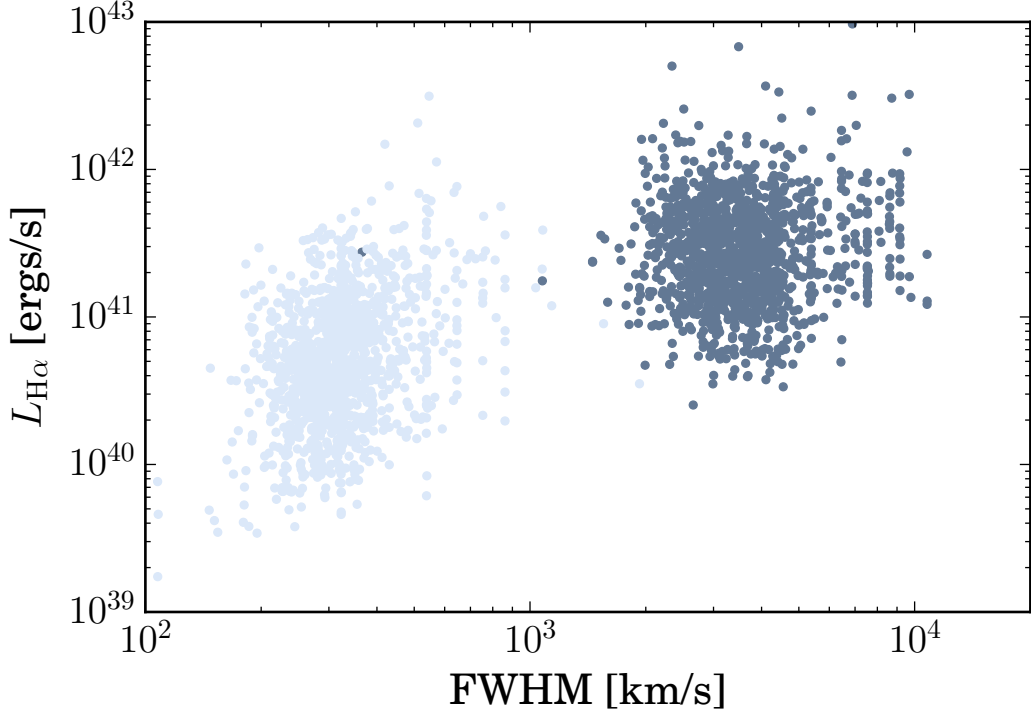


Figure 3.11. FWHM versus $H\alpha$ luminosity for our Seyfert 1.9 sources. Light blue represents the narrow component. Dark blue represents the broad component.

Table 3.5. Correlation coefficient of $L_{[\text{OIII}]}$ versus $L_{H\alpha}$ and $L_{5100\text{\AA}}$

	$L_{H\alpha, \text{broad}}$		$L_{H\alpha, \text{narrow}}$		$L_{H\alpha, \text{all}}$		$L_{5100\text{\AA}}$	
	r	p-value	r	p-value	r	p-value	r	p-value
$L_{[\text{OIII}]}$	0.66	1.92×10^{-162}	0.83	0	0.73	7.69×10^{-213}	0.37	3.83×10^{-44}

coefficient $r = 0.102$ with a $p\text{-value} = 0.0002$. The $p\text{-value}$ is an uncorrelated probability. These results indicate that the FWHM and luminosity of $H\alpha$ emission line has relatively weak correlation.

We further show the $L_{[\text{OIII}]}$ versus $L_{H\alpha}$ and $L_{5100\text{\AA}}$ in Fig. 3.12. The $[\text{OIII}]$ luminosity spans two orders while $H\alpha$ luminosity and $L_{5100\text{\AA}}$ span one order. We estimate the correlation coefficient between $L_{[\text{OIII}]}$ versus $L_{H\alpha}$ and $L_{5100\text{\AA}}$ and show the results in Table 3.5. We find that $[\text{OIII}]$ luminosity correlates strongly with $H\alpha$ luminosity and correlates weakly with $L_{5100\text{\AA}}$. The $p\text{-value}$ is the probability that the two distributions are zero correlation and we find the extremely low $p\text{-value}$ suggests that the $[\text{OIII}]$ luminosity are correlated with $L_{H\alpha}$ and $L_{5100\text{\AA}}$. Since the $[\text{OIII}]$ luminosity is isotropic and these results indicate our $H\alpha$ luminosity do not be affected by orientation. We find the $[\text{OIII}]$ luminosity versus $L_{5100\text{\AA}}$ are more dispersion than that with $H\alpha$. This suggests that our continuum luminosity might include the contribution from the host galaxy.

In addition to AGNs, there are other origins to produce a broad $H\alpha$ emission

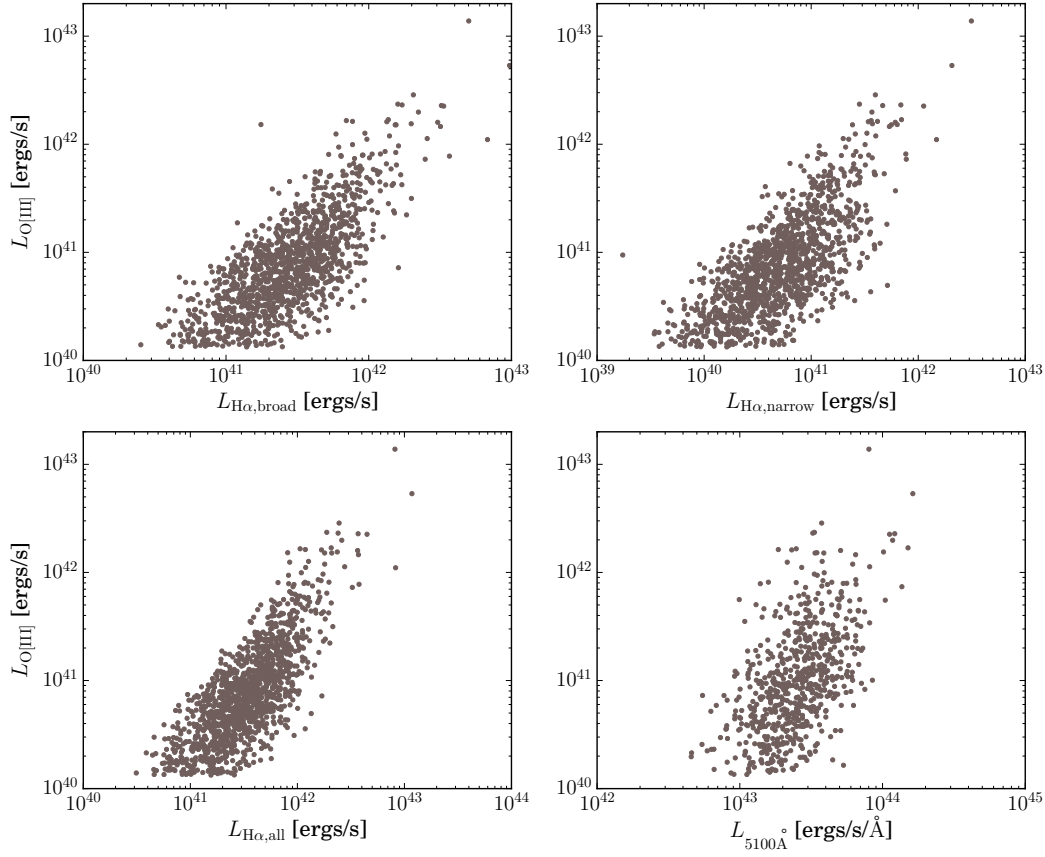


Figure 3.12. [OIII] luminosity versus $L_{H\alpha}$ and $L_{5100\text{\AA}}$ for our Seyfert 1.9 galaxies.

line. The Wolf-Rayet (WR) stars and the SN events [13]. The energy output by the stellar winds from WR or luminous blue variable (LBV) stars is 10^{36} to 10^{40} [ergs/s]. Our Seyfert 1.9 sample has luminosity of $H\alpha$ component ranges from 10^{40} to 10^{43} [ergs/s]. The different luminosity range indicates the broad emission line of our Seyfert 1.9 galaxies is caused by AGNs.

The “Baldwin, Phillips & Terlevich” (BPT) diagram is to distinguish different photoionization mechanisms in narrow-line regions of galaxies by using the emission line ratio [5]. Since the diagram is only for the narrow-line region (NLR), we use the flux of the narrow component of $H\alpha$ to plot the BPT diagram and show the results in Fig. 3.13. We find that both of the two Seyfert 1.9 samples have the most sources inside the AGNs region in the $[NII]/H\alpha$, $[SII]/H\alpha$, and $[OI]/H\alpha$ diagram. We estimate the sources outside the AGN region in $[NII]/H\alpha$, $[SII]/H\alpha$, and $[OI]/H\alpha$ diagram and show the results in Table 3.6. These results indicate that most sources of our Seyfert 1.9 samples are ionized by AGNs instead of stellar. We also find that our Seyfert 1.9 sources are dominated in the region of Seyfert in the $[SII]/H\alpha$ and $[OI]/H\alpha$ diagram. In $[NII]/H\alpha$ diagram, an experienced division line for Seyfert and LINER indicates that some of our Seyfert 1.9 sources stay in the LINER region although most sources stay in the Seyfert region of the $[SII]/H\alpha$ and $[OI]/H\alpha$ region. These results suggest that some of our Seyfert 1.9 sources have a relatively low

Table 3.6. Number of the Seyfert 1.9 outside the AGN region in BPT diagram

	[NII]/H α	[SII]/H α	[OI]/H α
641 Seyfert 1.9	32	187	107
656 Seyfert 1.9	31	131	67

ionized level.

We also compare the BPT diagram of our Seyfert 1.9 sample to the Seyfert 2 galaxies [7]. Both of our Seyfert 1.9 samples are dominated in the AGN region in the [NII]/H α , [SII]/H α , [OI]/H α diagram. These results are similar to the Seyfert 2 galaxies which show dominate in the AGN region. This result indicates that the NLR of the Seyfert 1.9 galaxies is ionized by AGNs and the mechanism is the same with the Seyfert 2 galaxies. However, we notice that in [NII]/H α diagram Seyfert 2 galaxies are dominant the region which has relatively weak [NII]/H α and higher [OIII]/H β while the Seyfert 1.9 galaxies are dominant the region which has relatively strong [NII]/H α and lower [OIII]/H β . The relatively higher [NII]/H α of the Seyfert 1.9 might relate to the host galaxy [49].

3.3 Variability of Seyfert 1.9 galaxies

Stripe 82 is a survey in the SDSS. This survey is along the Celestial Equator in the Southern Galactic Cap and has repeatable spectroscopic and photometric observations of the objects in the covering range with $-60^\circ < \text{R.A.} < 60^\circ$ and $-1.26^\circ < \text{Dec.} < 1.26^\circ$ [2]. We search for repeatable spectra of our 1297 Seyfert 1.9 sources in the Stripe 82 region. However, we only have 13 sources with multiple spectra and obtain the spectra from Catalog Archive Server (CAS) Stripe 82 database. We list coordinates and detailed observation information of our 13 sources in Table. 3.7. We show the multiple spectra of the 13 sources in the top panel in Fig. 3.14 to Fig. 3.26. The shortest time interval of our data is 1 day and the longest time interval of that is 1002 days. In order to investigate H α variations of these sources, we fit the spectra with a linear function as a pseudo continuum and two Gaussian components as the broad and narrow components. The detail fitting process is described in Section 2.3 and the detail fitting results for each source are in Appendix A. We estimate the variation flux of broad H α components by calculating the difference between observations with the first spectrum.

We show the multiple observations of spectra, light curves of broad components and continuum, and profiles of the broad component for each source in Fig. 3.14 to Fig. 3.26. From the light curve of the broad H α component and continuum, we find the flux of the broad H α component varies along with the flux density of the continuum. However, we find two sources (Fig. 3.14 and Fig. 3.25) have stronger broad H α components and weaker continuum in certain period (MJD= 51876 to 51913 and MJD=51885 to 52201). For the source in Fig. 3.14, the time scale is about 37 days and it is inside the of the light travel effect. For the source in Fig. 3.25, the variation time scale is 316 days and is hard to be explained by the light travel effect. The other source (Fig. 3.19) shows a weak broad H α component while the

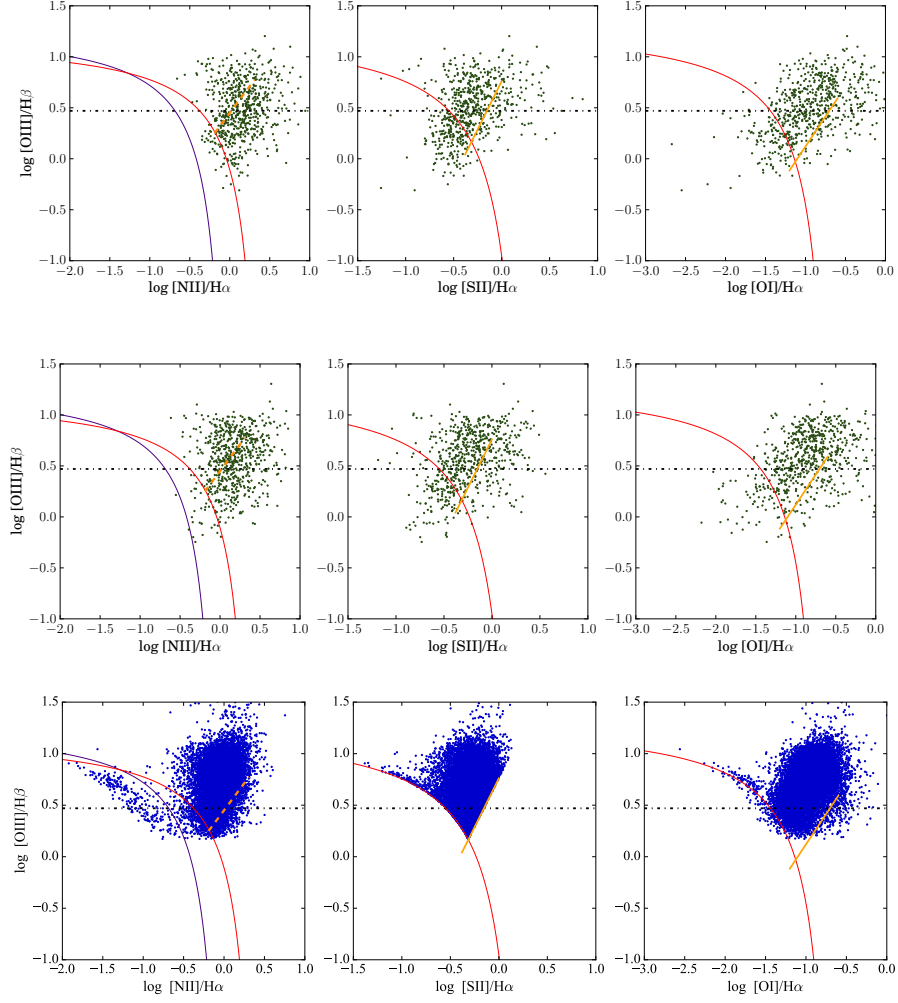


Figure 3.13. BPT diagram for our Seyfert 1.9 sample. Top: 641 Seyfert 1.9 sample. Middle: 656 Seyfert 1.9 sample. Bottom: Seyfert 2 galaxies with $L_{[\text{OIII}]}$ [ergs/s] > 40.125 [7]. The red line represents the definition of the starburst limit [15]. The dashed-orange line in [NII]/H α diagram represents the Seyfert-LINER [30]. The orange line represents the Seyfert-LINER line [16]. The purple solid line represents the AGN definition [14]. The black dashed line represents the ratio of [OIII]/H β = 3.

continuum is stronger between MJD=52174 and MJD=52178. We use the earliest spectrum in each source as a basement to estimate varying percentages of the broad H α component. The maximum variation of the broad H α component increases 94% in 1187 days (Fig. 3.26). The minimum variation of broad H α component increases 0.01% in 56 days (Fig. 3.22). The maximum variation of continuum level increases 36% in 1187 days (Fig. 3.26) and the minimum variation of the continuum level decreases 1.7% in 56 days (Fig. 3.22). We find the largest variation in broad H α and continuum level is in Fig. 3.26 while the minimize variation is the source in Fig. 3.22. The flux of broad H α and continuum level have small errors and we conclude these sources have significant variation. We find a source (Fig. 3.25) that has a significant variation in continuum level and broad H α component with a small error bar. Its broad H α component increases 45.9% and the continuum level decreases 19.3% in 390 days. We also notice its $\text{FWHM}_{\text{H}\alpha, \text{Broad}}$ changes from 4178 to 3418 [km/s]. Usually, the broad H α component responds to the variation of a continuum with a time lag due to the light travel effect. The time scale of the light travel effect is about few days to few weeks but we find this time scale is larger than the time scale of the light travel. Therefore, we rule out the possibility of the light travel effect. Another possibility might be due to the continuum flux does not correspond to ionizing continuum flux. From the profiles of the broad component, we do not find significant variation. The profiles only show the intensity variation of the amplitude. However, we notice some of our sources (Fig. 3.16, Fig. 3.17, Fig. 3.19, Fig. 3.22, Fig. 3.23, Fig. 3.25) have broader wings and the wings overlap in different observed time. The different profiles of the broad H α component might relate to the kinematic of the BLR. The authors showed the profiles of the different kinematic properties of the gas in the BLR respond to a continuum [39]. For our sources that are listed above, we conclude that these sources are mainly caused by chaotic gas motion in their BLR. The other sources (Fig. 3.18 & Fig. 3.24) that do not have overlap wings in different observed times are related to gas rotation in the BLR.

We show the results of flux of broad H α component and continuum level at 6562Å in Fig. 3.27. We find the flux of broad H α component responses to the continuum level. We estimate a correlation coefficient between the flux of broad H α component and continuum level at 6562Å to test if the two values are correlated. The correlation coefficient is 0.84 with a p-value = 9.17×10^{-11} . We also show luminosity of broad H α component and luminosity density at 6562Å in Fig. 3.28. The correlation coefficient for broad H α luminosity and continuum level luminosity is 0.73 with a p-value = 4.72×10^{-7} . The p-value is an uncorrelated probability and these results show that the broad H α component is highly correlated to the continuum level at 6562Å. This result is consistent with the literature; the strong continuum level has luminous broad H α component [47]. It suggests that the emission line is related to its continuum radiation and photoionized by center radiation [27].

Table 3.7. Coordinates, redshifts, observation information for our 13 sources

R.A.	Dec.	z	plate	MJD	fiberID
52.935422	-1.0879826	0.0851696	415	51810	246
		0.085141	415	51879	248
30.385701	0.39812053	0.078188	404	51877	347
		0.078162	404	51812	343
13.262922	-0.17965537	0.13831	394	51913	161
		0.138298	394	51812	176
		0.138837	394	51876	166
5.2221571	0.63686113	0.144553	390	51900	456
		0.144353	390	51816	454
		0.144547	688	52203	422
42.700671	0.31673936	0.18769	707	52177	631
		0.187389	410	51816	360
		0.187546	410	51877	351
		0.187462	708	52175	393
55.166922	0.088592222	0.130904	714	52201	496
		0.130917	416	51811	479
		0.130941	416	51885	463
343.38084	0.80708478	0.0724755	676	52178	453
		0.072455	379	51789	507
		0.072478	676	52174	458
351.10088	0.14853945	0.148683	383	51818	493
		0.148586	680	52200	514
14.282579	0.088951841	0.195946	395	51783	393
		0.19587	693	52254	347
45.505091	-1.0216328	0.166582	411	51817	260
		0.166661	411	51873	248
		0.166662	411	51914	253
330.36646	-0.80064708	0.110172	1032	53175	54
		0.110156	372	52173	56
24.6492	0.20733	0.124992	1078	52643	571
		0.125044	1077	52644	358
56.497262	0.303874	0.161065	1632	52996	588
		0.16106	416	51811	624
		0.161073	416	51885	624
		0.161196	714	52201	629
		0.161058	1633	52998	394

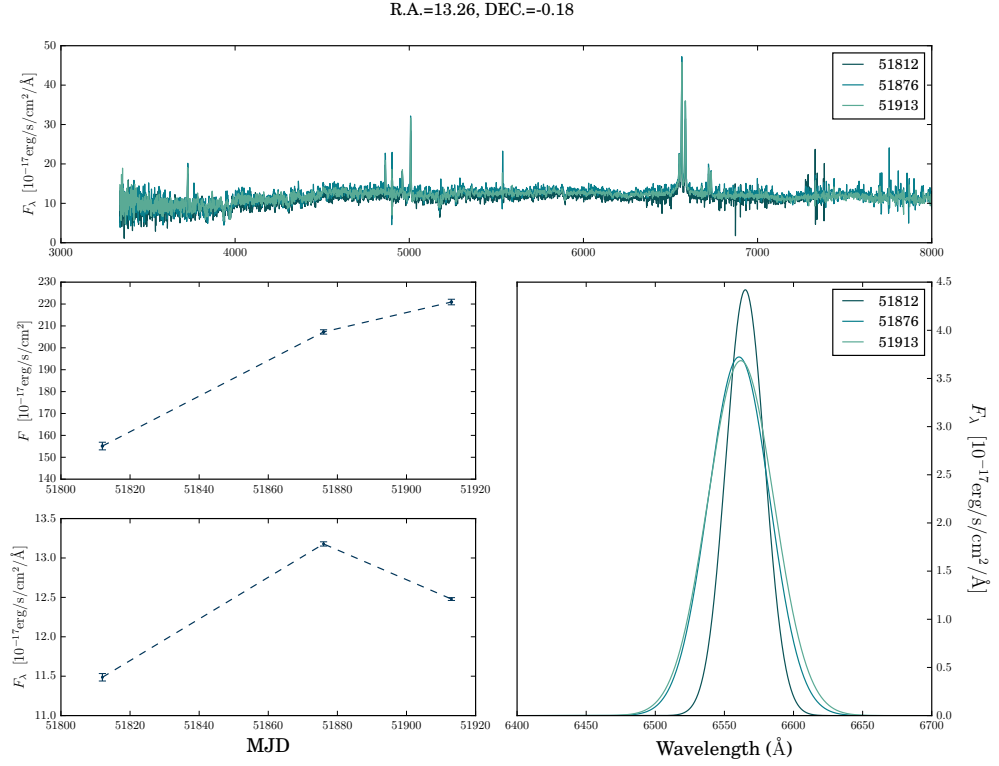


Figure 3.14. Top: repeat observation spectra. Middle left: variation of the flux of broad component with MJD. Bottom left: variation of the flux density at 6562 \AA with MJD. Bottom right: profile variation of the broad component.

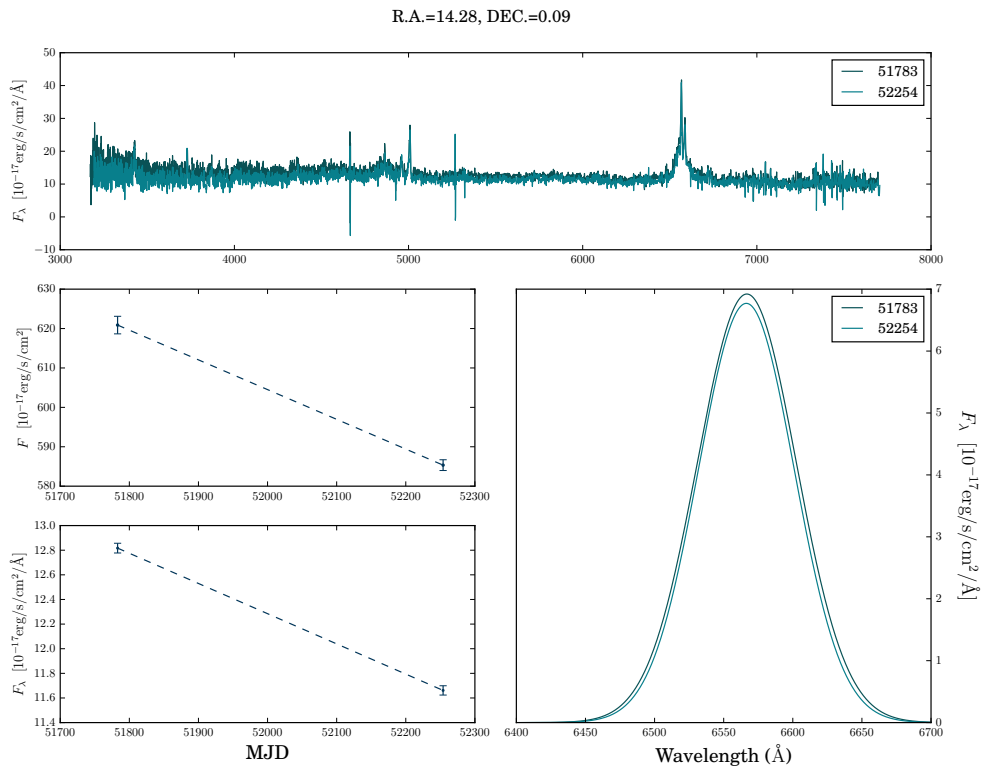


Figure 3.15. Top: repeat observation spectra. Middle left: variation of the flux of broad component with MJD. Bottom left: variation of the flux density at 6562Å with MJD. Bottom right: profile variation of the broad component.

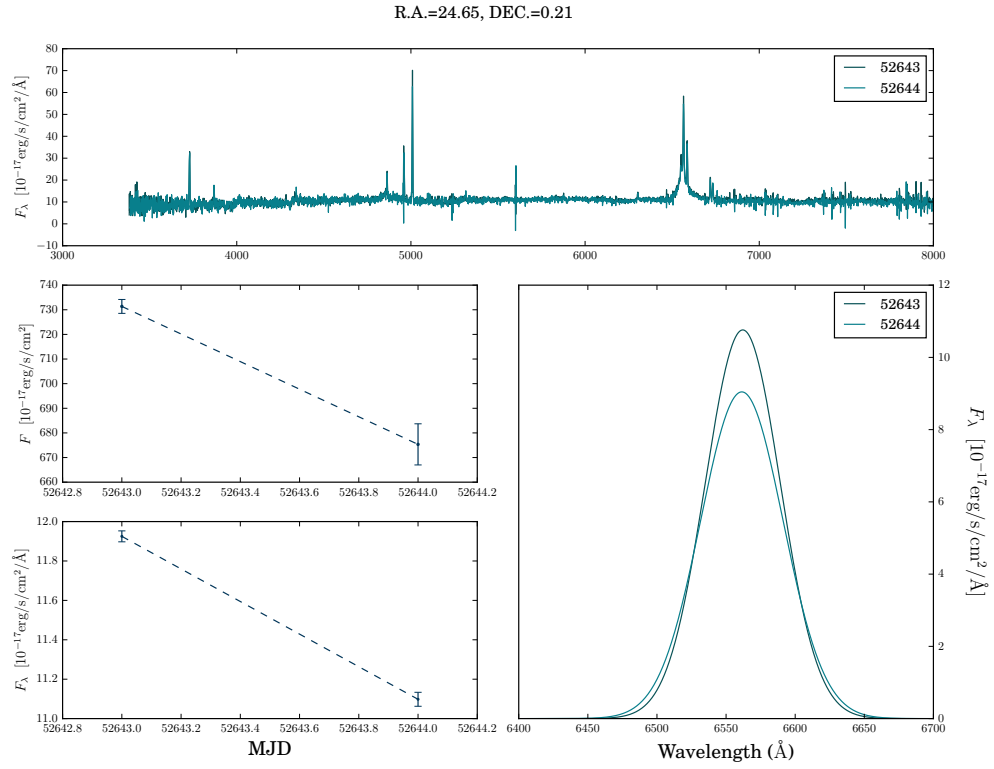


Figure 3.16. Top: repeat observation spectra. Middle left: variation of the flux of broad component with MJD. Bottom left: variation of the flux density at 6562Å with MJD. Bottom right: profile variation of the broad component.

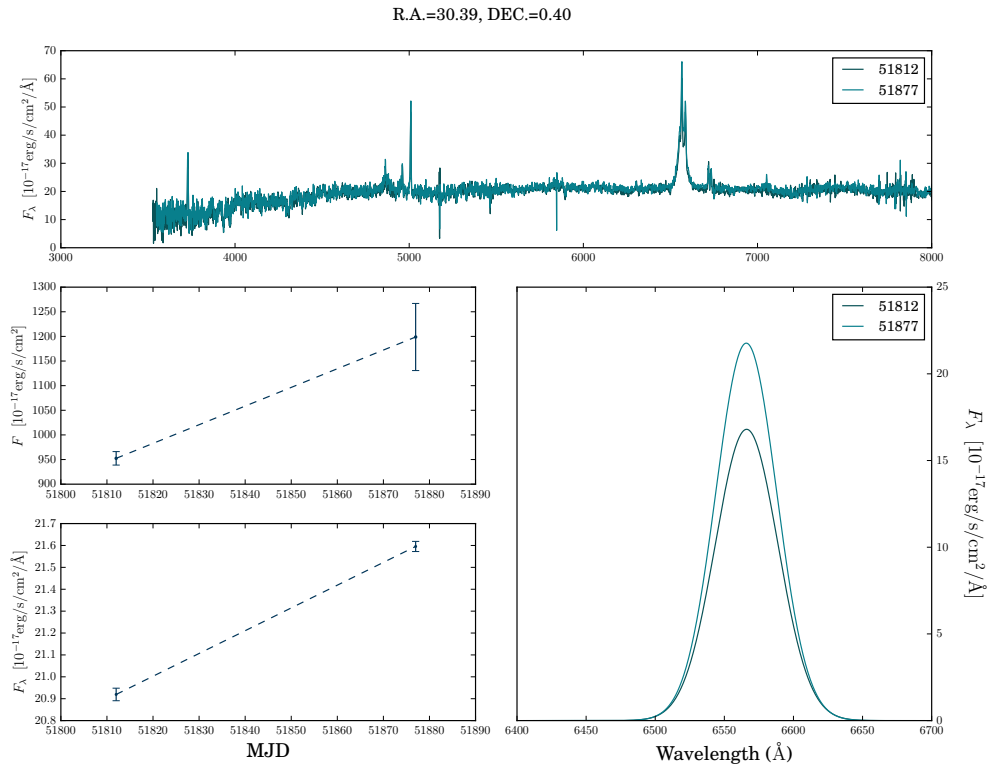


Figure 3.17. Top: repeat observation spectra. Middle left: variation of the flux of broad component with MJD. Bottom left: variation of the flux density at 6562Å with MJD. Bottom right: profile variation of the broad component.

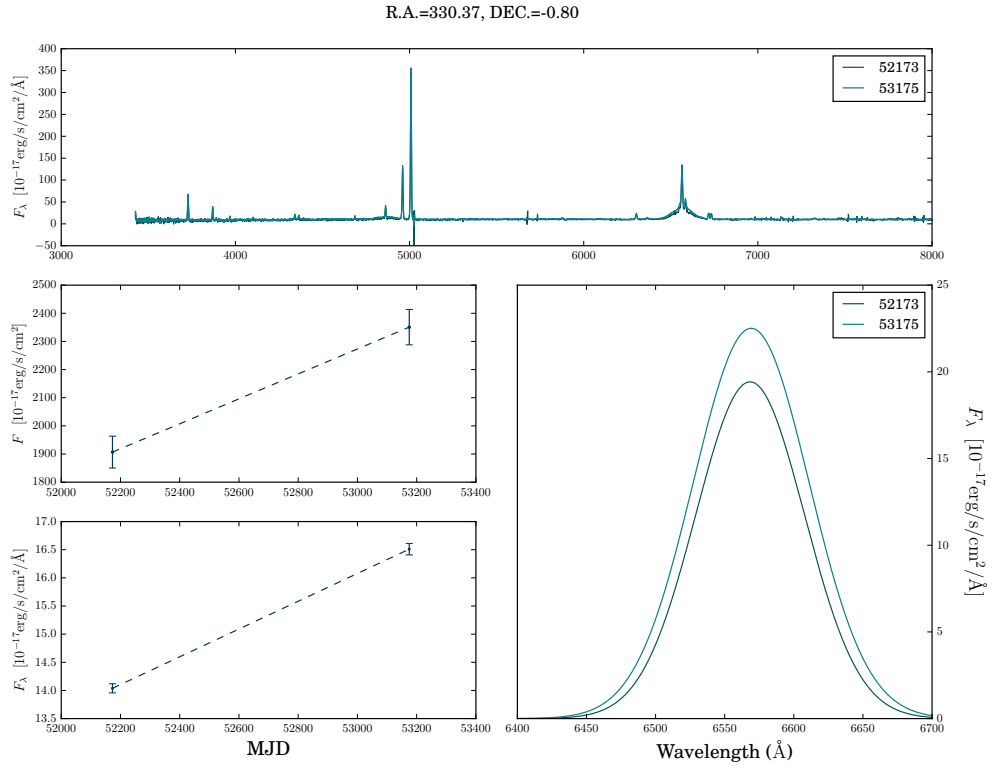


Figure 3.18. Top: repeat observation spectra. Middle left: variation of the flux of broad component with MJD. Bottom left: variation of the flux density at 6562Å with MJD. Bottom right: profile variation of the broad component.

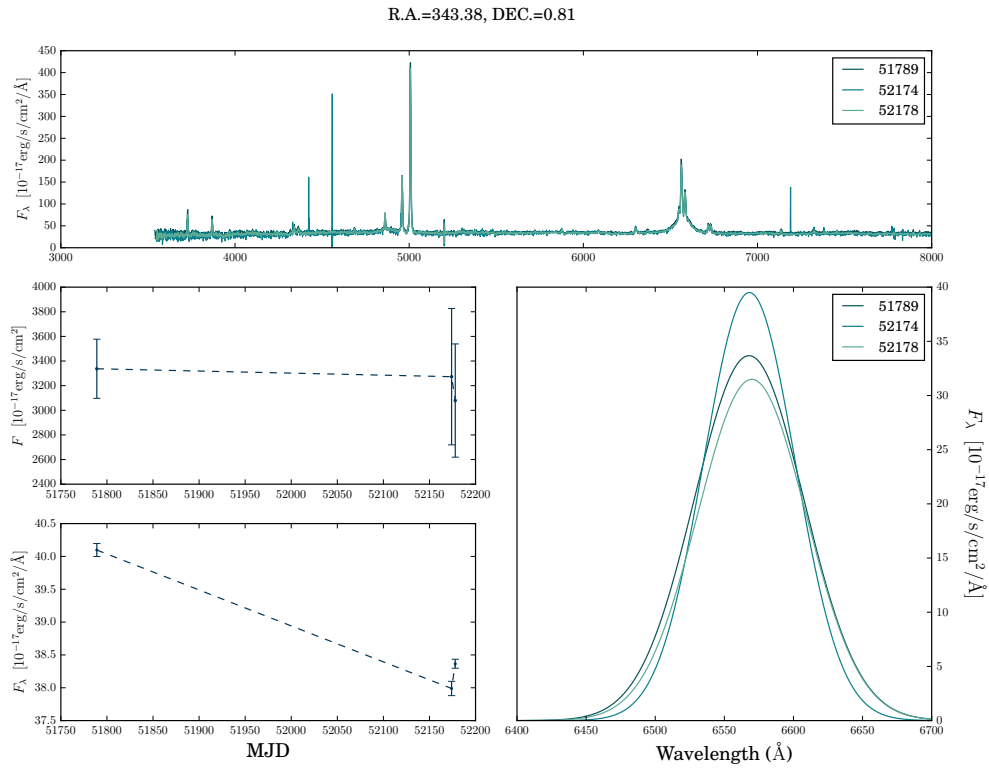


Figure 3.19. Top: repeat observation spectra. Middle left: variation of the flux of broad component with MJD. Bottom left: variation of the flux density at 6562Å with MJD. Bottom right: profile variation of the broad component.

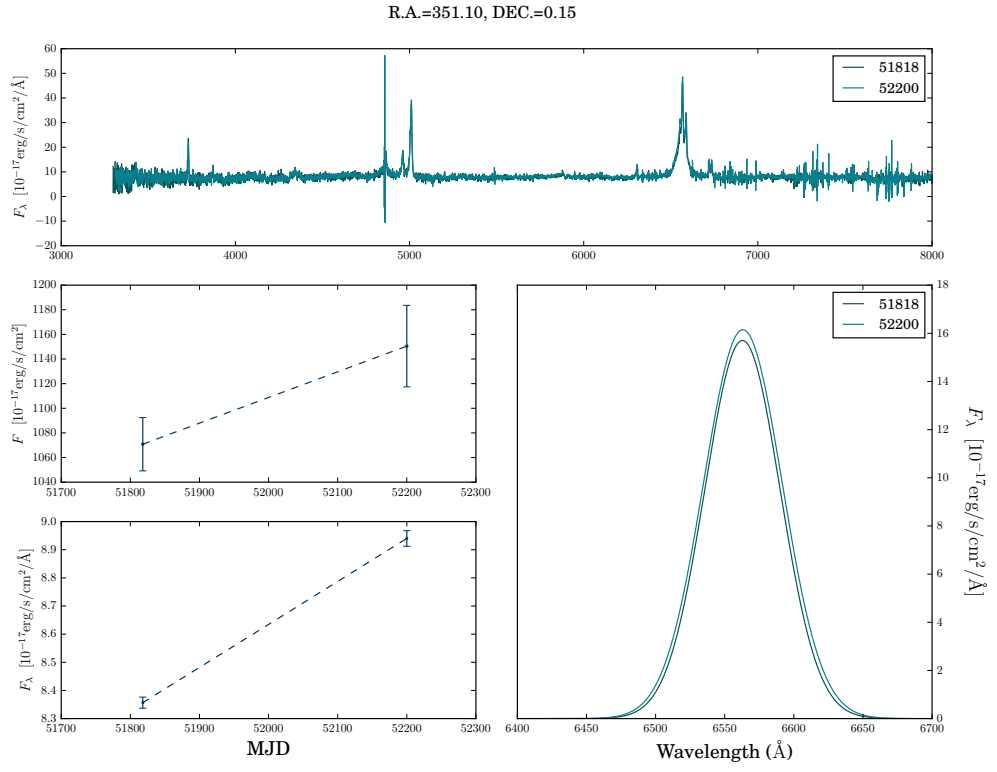


Figure 3.20. Top: repeat observation spectra. Middle left: variation of the flux of broad component with MJD. Bottom left: variation of the flux density at 6562\AA with MJD. Bottom right: profile variation of the broad component.

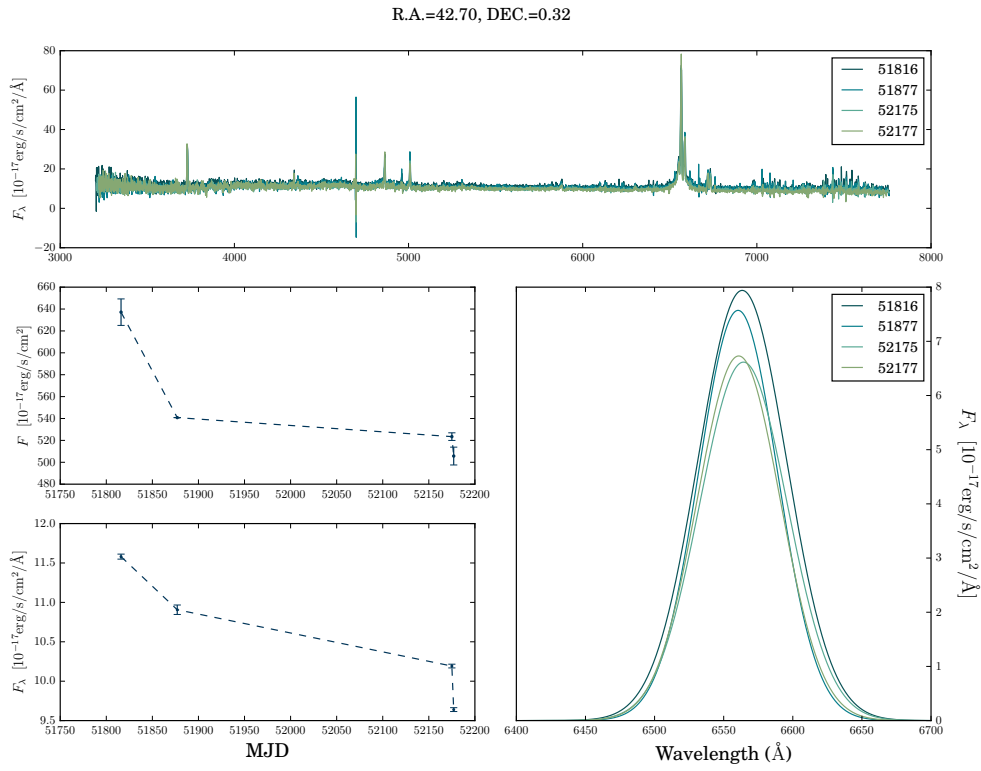


Figure 3.21. Top: repeat observation spectra. Middle left: variation of the flux of broad component with MJD. Bottom left: variation of the flux density at 6562\AA with MJD. Bottom right: profile variation of the broad component.

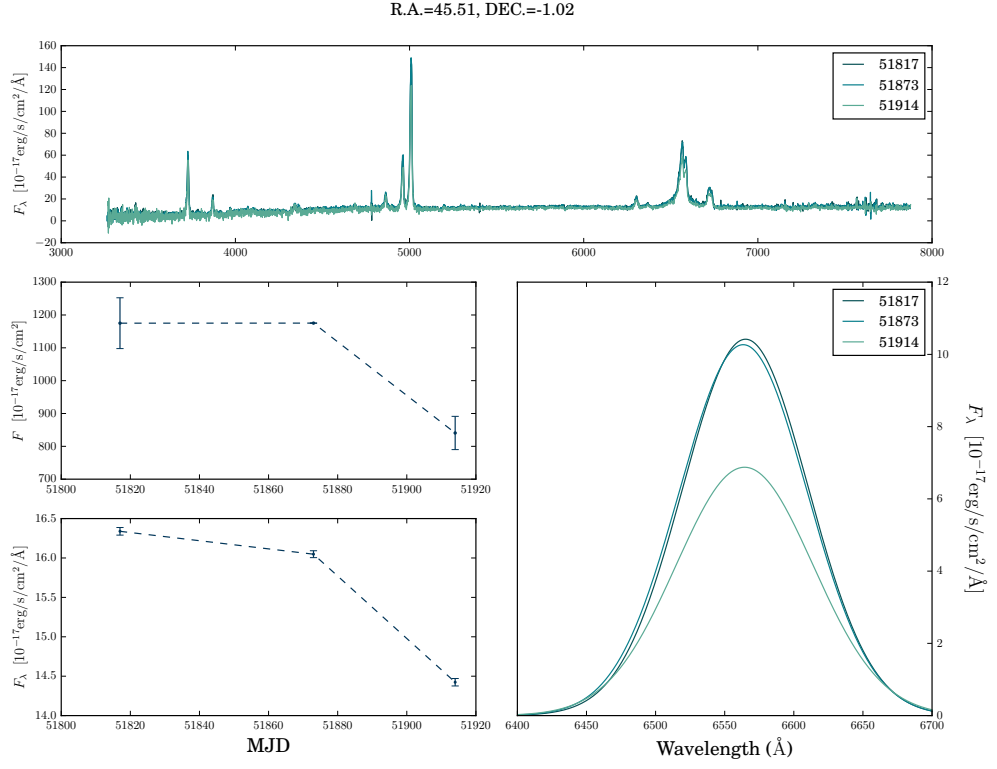


Figure 3.22. Top: repeat observation spectra. Middle left: variation of the flux of broad component with MJD. Bottom left: variation of the flux density at 6562Å with MJD. Bottom right: profile variation of the broad component.

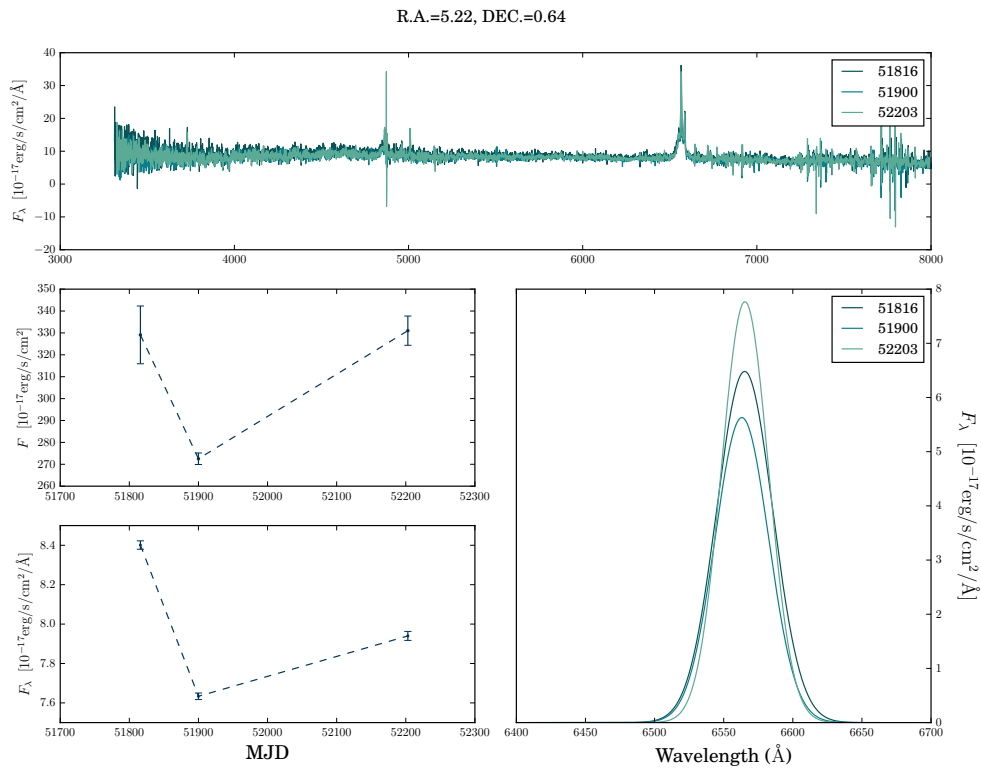


Figure 3.23. Top: repeat observation spectra. Middle left: variation of the flux of broad component with MJD. Bottom left: variation of the flux density at 6562 \AA with MJD. Bottom right: profile variation of the broad component.

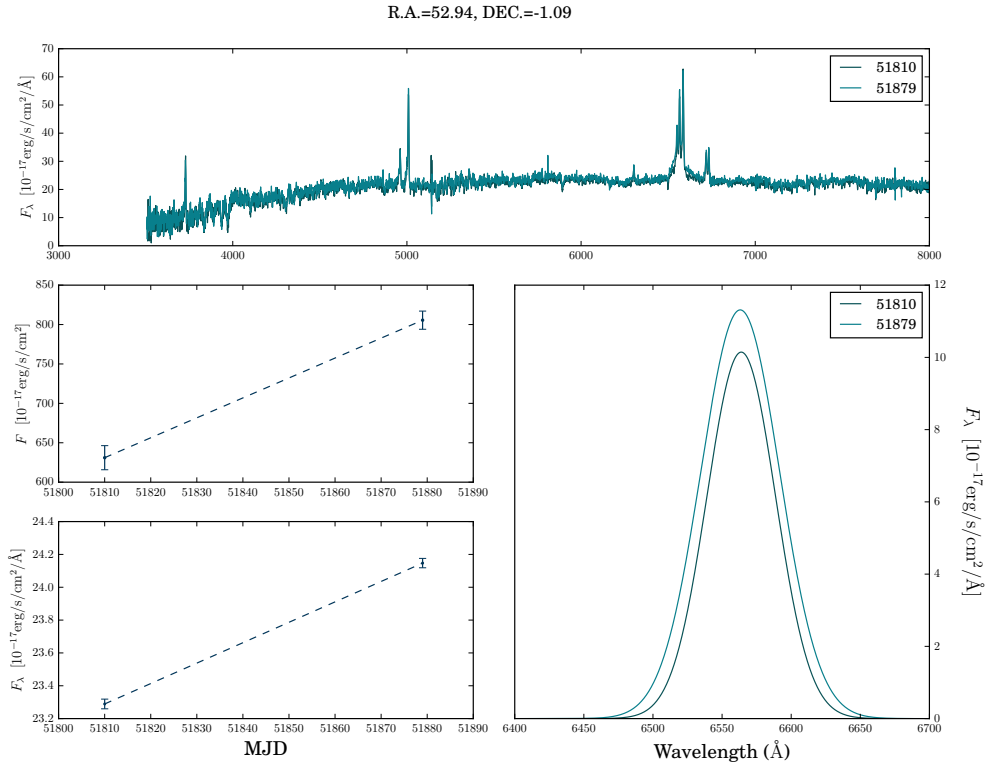


Figure 3.24. Top: repeat observation spectra. Middle left: variation of the flux of broad component with MJD. Bottom left: variation of the flux density at 6562 Å with MJD. Bottom right: profile variation of the broad component.

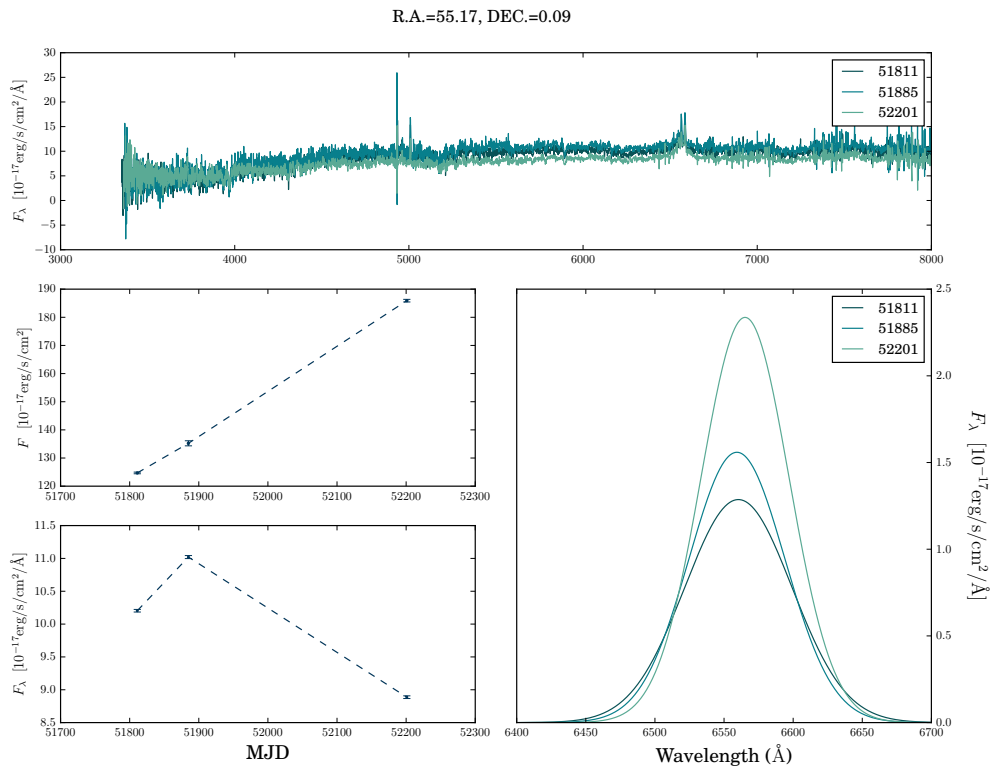


Figure 3.25. Top: repeat observation spectra. Middle left: variation of the flux of broad component with MJD. Bottom left: variation of the flux density at 6562 \AA with MJD. Bottom right: profile variation of the broad component.

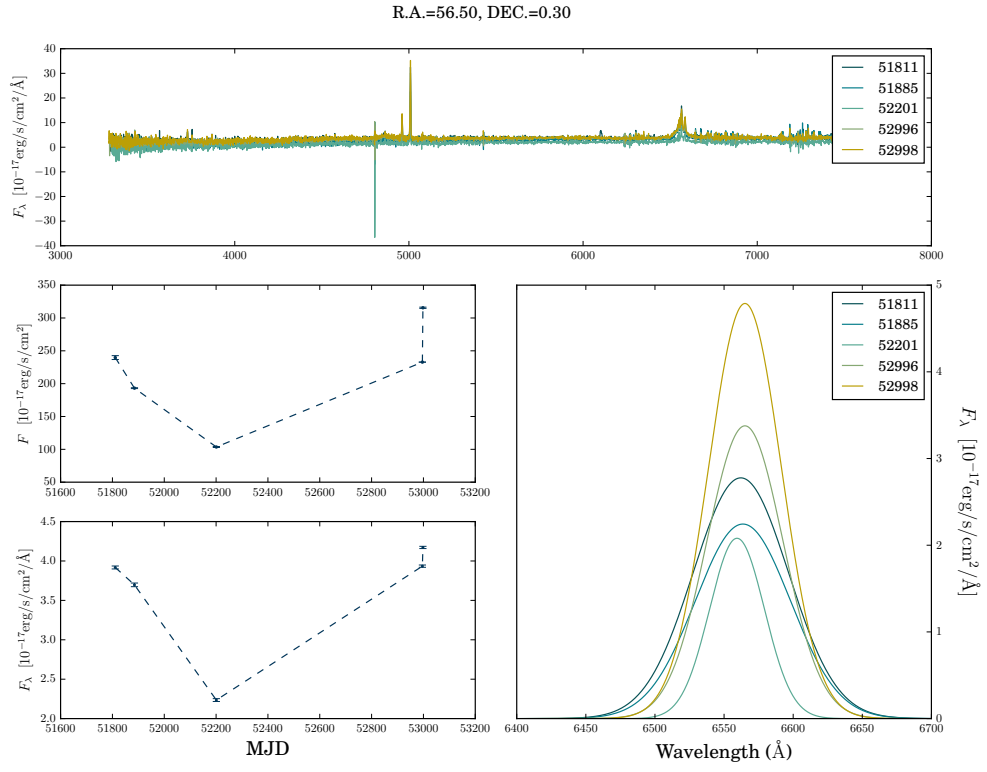


Figure 3.26. Top: repeat observation spectra. Middle left: variation of the flux of broad component with MJD. Bottom left: variation of the flux density at 6562Å with MJD. Bottom right: profile variation of the broad component.

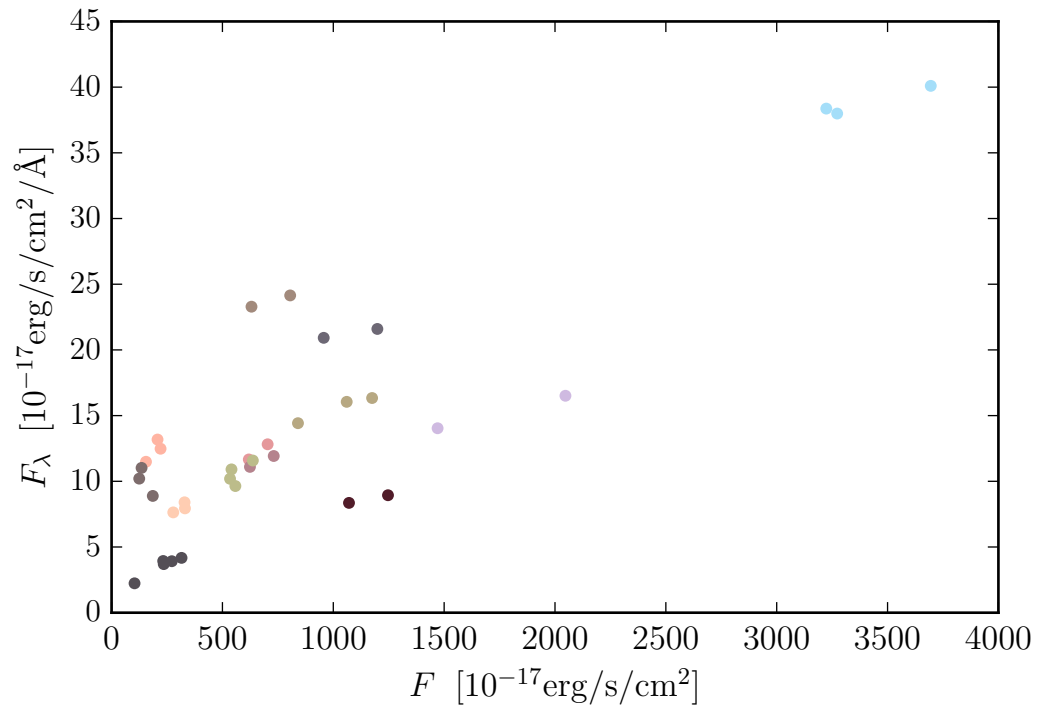


Figure 3.27. The flux of broad $\text{H}\alpha$ component versus flux density at 6562\AA . The same color represents the multiple observation of the same sources and different colors represent our 13 sources.

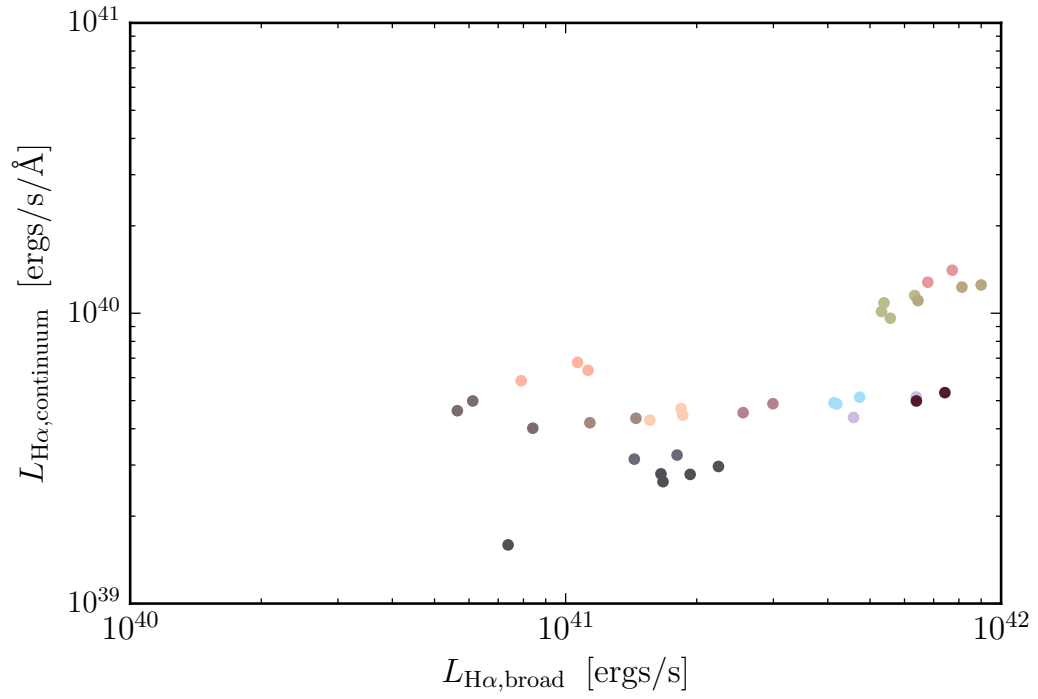


Figure 3.28. The luminosity of broad H α component versus luminosity density at 6562Å. The same color represents the multiple observation of the same sources and different colors represent our 13 sources.

Chapter 4

Discussion

4.1 Test result

We would like to know if the Seyfert 1.9 galaxies that are classified by our model as Seyfert 2 galaxies are related to the low-quality spectra and we check the S/N distributions of the training data and test data. We show the S/N distribution of the training set 1 in Fig. 4.1, the training set 2 in Fig. 4.2, and the training set 3 in Fig. 4.3. We also estimate the K-S test for the three data sets and show the results in Table 4.1. We find that the S/N distribution of the whole population is not different from that of training data and test data for the Seyfert 1.9 and Seyfert 2 sample in the three data sets. From the K-S test, we find all p-value results show that we can not reject the null hypothesis that the two distributions are drawn from the same population. These results indicate that the sources that are classified wrong by our model are not related to the low S/N and suggest that our model can recognize Seyfert 1.9 spectrum and Seyfert 2 spectrum with low S/N. The possible reasons for classifying wrong are insufficient training sources or adding more layers into our model to recognize the fine distinction between the spectra of Seyfert 1.9 and Seyfert 2 galaxies.

We notice the convolution filter is shifted from short wavelength to long wavelength and we would like to know if the classification results change when we change the input direction. We reverse the input spectra to make the filter shift from red to blue wavelength to test if the shift direction affects the classifying results. We still obtain a similar result with our training set 1 and 2. The model trained by reverse spectra still has a precision of over 80% for classifying correctly Seyfert 1.9 galaxies. This result indicates our model still can recognize the features between the Seyfert 1.9 and Seyfert 2 galaxies.

The loss curves in our three data sets have slightly different performances. A better loss curve is in data set 3; the loss curves of training and testing are horizons with a small gap after the 80th epoch. The small loss gap indicates the predicted output is very closed to the target and the horizon lines show the fitting process reaches its stability in parameter space. The loss curve in data set 1 is a moderate performance because it just begins to converge its stable point at epoch =100. The loss curve in data set 2 does not show significant stability and it can reach the stable point by extending the epoch. Due to the 2nd test result from the training set 1, we

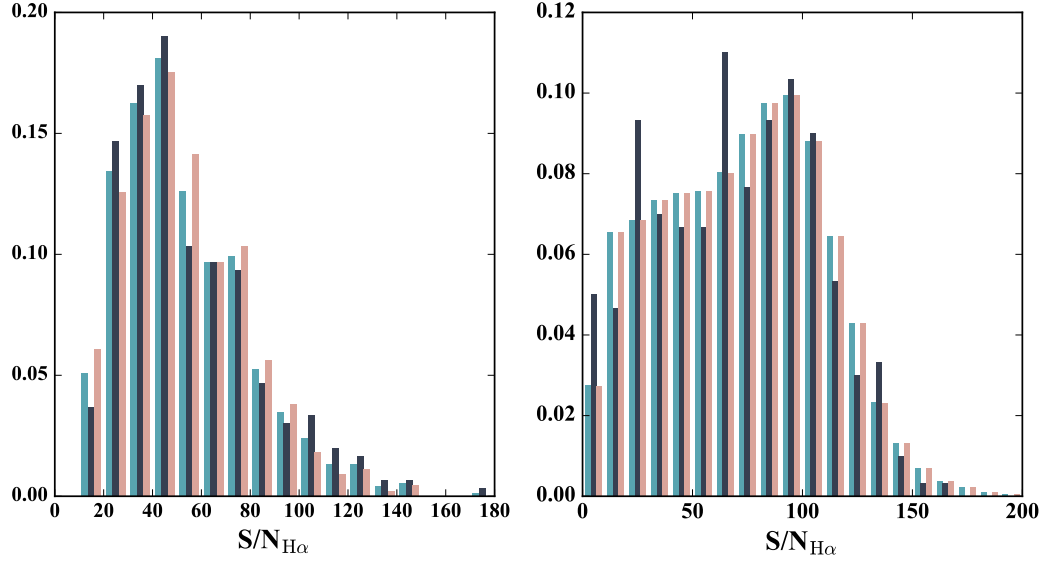


Figure 4.1. S/N of the $H\alpha$ emission line of our Seyfert sources in training set 1. Left: Intermediate Seyfert. Right: Seyfert 2. Fountain blue represents the whole population. Black represents the training data. Pink represents the test data.

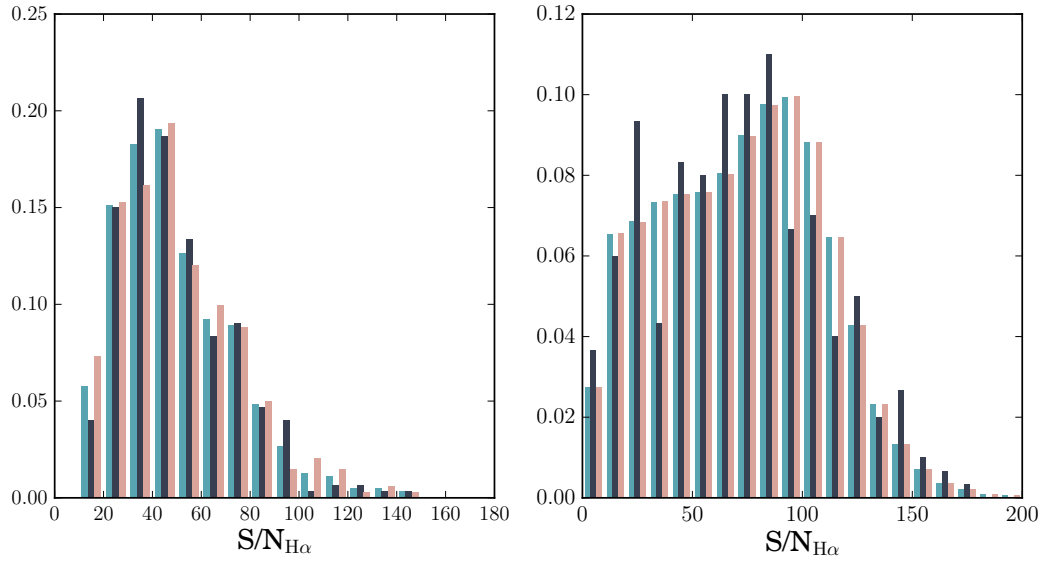


Figure 4.2. S/N of the $H\alpha$ emission line of our Seyfert sources in training set 2. Left: Seyfert 1.9. Right: Seyfert 2. Fountain blue represents the whole population. Black represents the training data. Pink represents the test data.

Table 4.1. K-S test results for S/N distributions in different training sets

Training set 1	D-statistic p-value	
whole Sy19 vs training Sy19	0.035	0.954
whole Sy19 vs test Sy19	0.024	0.998
training Sy19 vs test Sy19	0.058	0.562
whole Sy2 vs training Sy2	0.051	0.411
whole Sy2 vs test Sy2	0.0003	1.000
training Sy2 vs test Sy2	0.051	0.404
Training set 2	D-statistic p-value	
whole Sy19 vs training Sy19	0.033	0.978
whole Sy19 vs test Sy19	0.029	0.991
training Sy19 vs test Sy19	0.062	0.56
whole Sy2 vs training Sy2	0.068	0.123
whole Sy2 vs test Sy2	0.0004	1.000
training Sy2 vs test Sy2	0.068	0.119
Training set 3	D-statistic p-value	
whole Sy19 vs training Sy19	0.019	0.999
whole Sy19 vs test Sy19	0.016	1.000
training Sy19 vs test Sy19	0.035	0.829
whole Sy2 vs training Sy2	0.035	0.463
whole Sy2 vs test Sy2	0.005	0.433
training Sy2 vs test Sy2	0.030	0.647

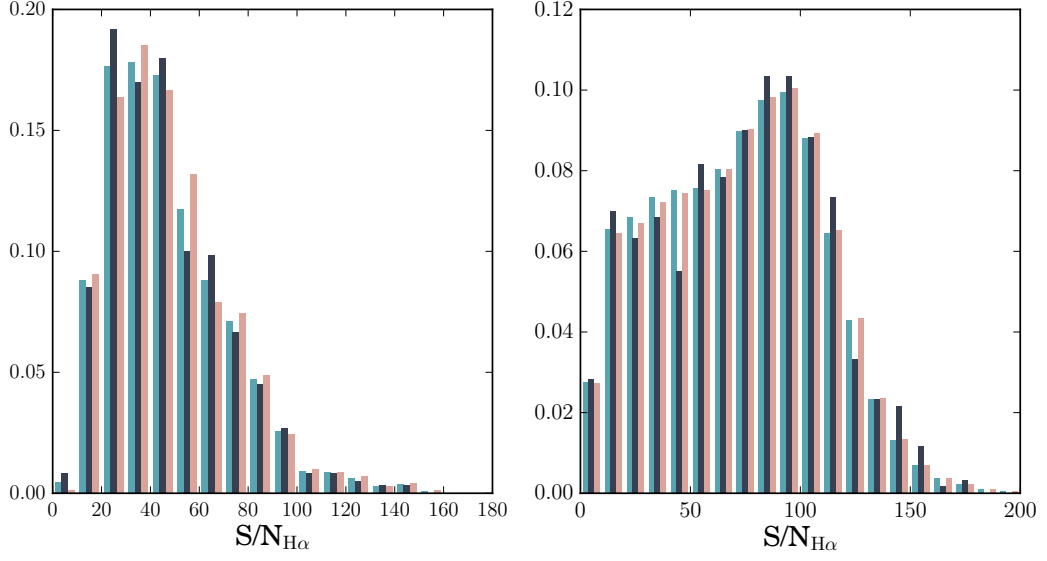


Figure 4.3. S/N of the H α emission line of our Seyfert sources in training set 3. Left: Seyfert 1.9. Right: Seyfert 2. Fountain blue represents the whole population. Black represents the training data. Pink represents the test data.

have more Seyfert 1.9 sources. We find adding these Seyfert 1.9 sources as training sources helps our model to have a lower loss value than training set 1 and 2. We also notice the precision of Seyfert 1.9 increases slightly from 88% in training set 1 to 91% in training set 3. These results indicate that adding more training sources improves the performance of our model. Besides expanding training sources, another possible improvement is to add more layers into our model or use a varied learning rate. Adding more layers into the model will have more parameters in our model and get overfitting easily. A learning rate represents a step size in the loss surface. At the early stage, a larger learning rate could help the model not be stuck in the local point. However, at the late stage, a big learning rate might jump out from the global minimum and the model can not obtain a good fitting result. A varied learning rate might cause the training process unstable if we do not have enough experience for deciding the value of the learning rate.

4.2 compared to intermediate Seyfert

In this work, we provide a total of 1297 Seyfert 1.9 galaxies. Besides, we have $104 + 53 = 157$ sources that have the broad H β component superimposed on the narrow H β component. These intermediate Seyfert galaxies usually have variability. Therefore, we present these sources to the public so that these targets can be traced by follow-up observation of optical or X-ray to have more spectrum for analysis. ^[1]

We find that our 656 Seyfert 1.9 sample has more sources with high $\text{FWHM}_{\text{H}\alpha, \text{broad}}$ than the 641 Seyfert 1.9 sample. However, the luminosity of the broad H α component in the 656 Seyfert 1.9 sample is not luminous as that in the 641 Seyfert 1.9 sample.

¹openuniverse.asi.it/syfert1.9

Usually, a high FWHM has a high luminosity. However, our 656 Seyfert sample does not have high luminosity. This is because the broad $H\alpha$ component is more smooth toward the two sides. The broad $H\alpha$ component has an extended deviation and a low amplitude. This characteristic is less significant and prominent than the narrow $H\alpha$ component on the spectrum. The less significant characteristic is the possible reason that these sources were missed when we picked sources out by visual inspection. This result indicates that our model can pick out the Seyfert sources with a relatively weak and broad $H\alpha$ component.

We compare the total $H\alpha$ luminosity of an intermediate AGN sample in literature [35]. We find the $H\alpha$ luminosity of the sample ranges from 10^{40} to 10^{44} [ergs/s] and $H\alpha$ luminosity of our Seyfert 1.9 galaxies have one order weak than the intermediate AGN sample. The intermediate AGN sample is selected from a population with a strong broad $H\alpha$ component, indicating the sample is dominated by Seyfert 1.2 or 1.5. Therefore, the sources with a strong broad $H\alpha$ component have higher luminosity than our Seyfert 1.9 sample. We also notice the author showed the mean $H\alpha$ profile spectra in bins of broad $H\alpha$ luminosity and FWHM and find our Seyfert 1.9 sources that have high $\text{FWHM}_{H\alpha, \text{broad}}$ and low $H\alpha$ luminosity are similar to the $H\alpha$ profile in a grid with $L_{H\alpha, \text{broad}} = 10^{40}$ to 10^{41} [ergs/s].

We show the $[\text{OIII}]/H\beta$ distribution of Seyfert 1.9 in Fig. 4.4. We find our Seyfert 1.9 galaxies show an extending distribution of $[\text{OIII}]/H\beta$ ranging from 0 to 15 with a peak value at 2-3. A Seyfert 2 usually has a value of $[\text{OIII}]/H\beta > 3$ and the $[\text{OIII}]/H\beta$ value usually represents the ionization level [34]. This result indicates that most Seyfert 1.9 have a similar ionization level to the Seyfert 2 galaxies. In Table. 1.1, we find the value of $[\text{OIII}]/H\beta$ of the Seyfert 1.0 to Seyfert 1.8 overlaps with our Seyfert 1.9 sources. Therefore, this might be the reason that Seyfert 1.9 was not assigned $[\text{OIII}]/H\beta$ value because there is no distinct difference between the $[\text{OIII}]/H\beta$ value in Seyfert 1.9 and other intermediate Seyferts.

4.3 Supermassive Black Hole Mass

We estimate Supermassive Black Hole (SMBH) mass of the two Seyfert 1.9 samples. We use the formula in literature and show the formula below [11].

$$M_{\text{BH}} = (3.0^{+0.6}_{-0.5}) \times 10^6 \left(\frac{L_{H\alpha}}{10^{42} \text{ ergs s}^{-1}} \right)^{0.45 \pm 0.03} \left(\frac{\text{FWHM}_{H\alpha}}{10^3 \text{ km s}^{-1}} \right)^{2.06 \pm 0.06} M_{\odot} \quad (4.1)$$

where M_{BH} is mass of SMBH, $L_{H\alpha}$ is broad $H\alpha$ luminosity, $\text{FWHM}_{H\alpha}$ is broad $H\alpha$ FWHM, and M_{\odot} is solar mass.

We show M_{BH} distributions of our two Seyfert 1.9 samples in Fig. 4.5. The M_{BH} distribution ranges from 10^6 to $10^9 M_{\odot}$. We find the two M_{BH} distributions are dominant at $\approx 10^7 M_{\odot}$ but show slightly different peak values. We estimate the mean values and standard deviations of the two distributions. The mean value for the 641 Seyfert 1.9 sample is 3.22×10^7 (3.30×10^7) M_{\odot} and for the 656 Seyfert 1.9 sample is 3.52×10^7 (5.15×10^7) M_{\odot} . We further estimate a t-test to investigate if the mean values of the SMBH mass distributions in the two Seyfert 1.9 samples are different or not. The t-statistic = -1.25 with a p-value = 0.21. The p-value is

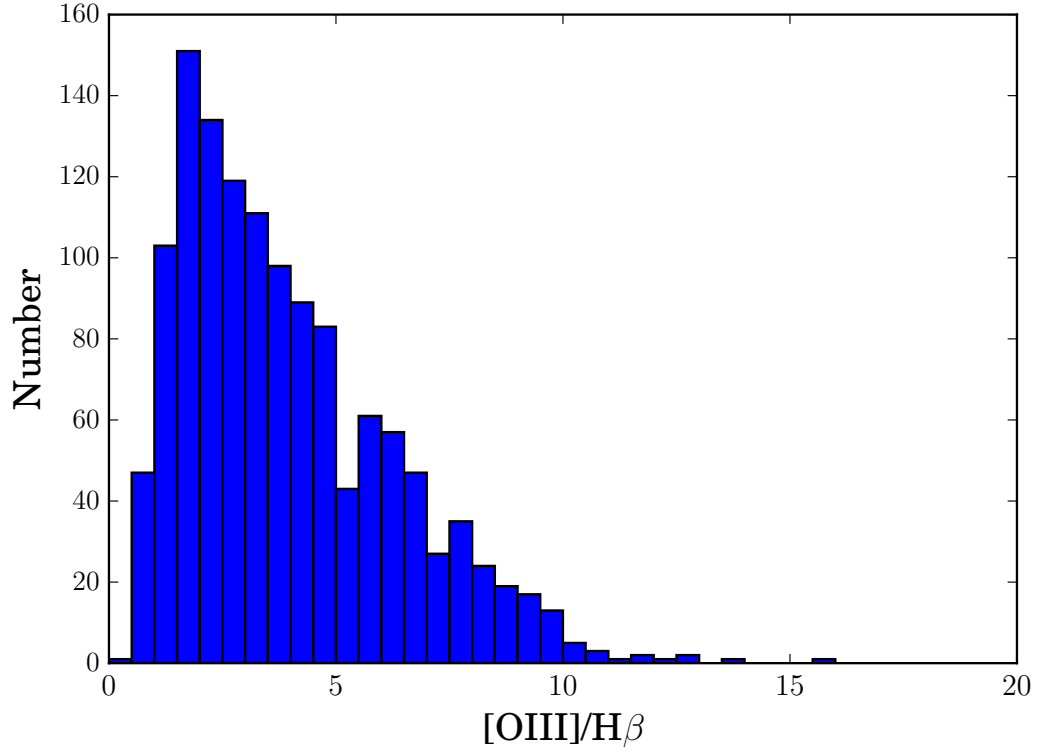


Figure 4.4. The distributions of [OIII]/H β of our Seyfert 1.9 galaxies.

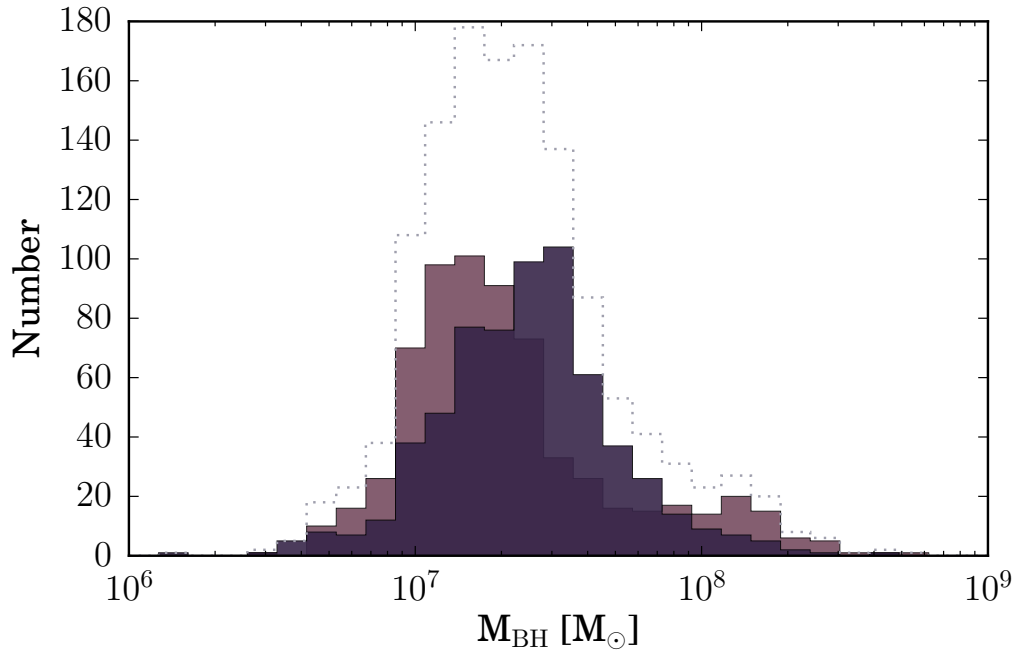


Figure 4.5. The distributions of SMBH mass of our Seyfert 1.9 samples. Dark-purple represents the 641 Seyfert 1.9 sample. Twilight-lavender represents the 656 Seyfert 1.9 sample

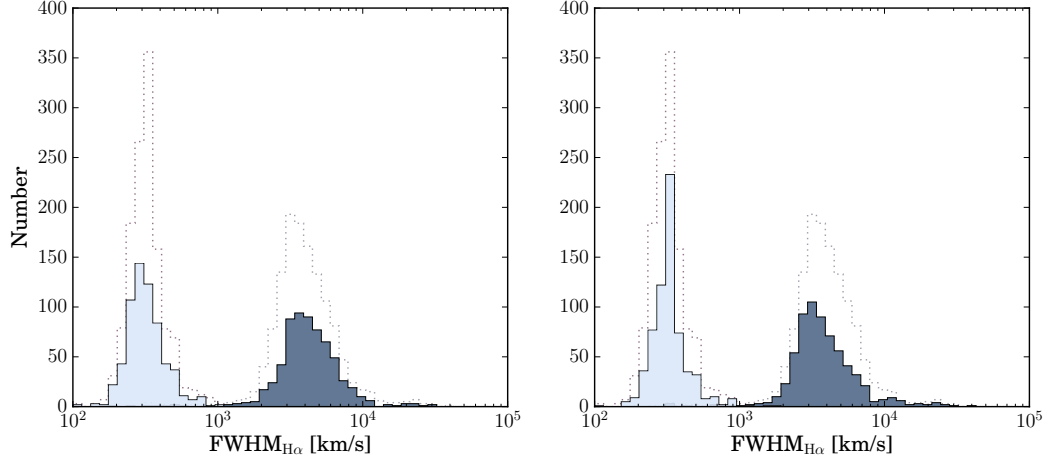


Figure 4.6. The distributions of FWHM of our Seyfert 1.9 samples. Left: 641 Seyfert 1.9 sources. Right: 656 Seyfert 1.9 sources. Dot-line represents the total 1297 Seyfert 1.9 sources. Light-blue represents the narrow component. Dark-blue represents the broad component.

the probability that the two samples are drawn from the same population and have the same population means. From this result, we can not reject the null hypothesis that the two populations have the same mean values. From the results of standard deviation, the 656 Seyfert 1.9 sample has slightly spread distribution in SMBH mass than the 641 Seyfert 1.9 sample. We notice the 656 Seyfert 1.9 sample shows a small tail toward the high mass end and estimate the amount above $10^8 M_{\odot}$ in the two Seyfert 1.9 samples. The percentage for the SMBH mass above $10^8 M_{\odot}$ in the 641 Seyfert 1.9 sample is 2.96% (19/641) and in the 656 Seyfert 1.9 sample is 8.23% (54/656).

4.4 Fitting $H\alpha$ with consideration of $[NII]\lambda 6548$ and $[NII]\lambda 6584$

$H\alpha$ emission is blended with $[NII]\lambda 6548$ and $[NII]\lambda 6584$. We fit the $H\alpha$ emission line again by considering $[NII]\lambda 6548$ and $[NII]\lambda 6584$ lines. We show the FWHM distributions in Fig. 4.6 and mean value of the distributions in Table 4.2. We find both of the Seyfert 1.9 samples show slightly skewed $FWHM_{H\alpha, broad}$ distributions. The $FWHM_{H\alpha, broad}$ distribution of the 641 Seyfert 1.9 sample distributes slightly broader than previous result (Fig. 3.5). The $FWHM_{H\alpha, broad}$ distribution of the 656 Seyfert 1.9 sample does not show a high FWHM tail as the previous result (Fig. 3.5). We find the mean values of $FWHM_{H\alpha, narrow}$ are similar to previous results (Table 3.2). However, the mean values of $FWHM_{H\alpha, broad}$ in two Seyfert 1.9 samples are higher than previous results (Table 3.2) and it relates to overestimate broad $H\alpha$ component when the fitting process does not consider $[NII]$ emission line. We notice that the mean value of $FWHM_{H\alpha, broad}$ in 641 Seyfert 1.9 sample is higher than that in 656 Seyfert 1.9 sample in this fitting. This suggests that the fitting $H\alpha$ emission line without considering $[NII]$ may underestimate the $FWHM_{H\alpha, broad}$. We

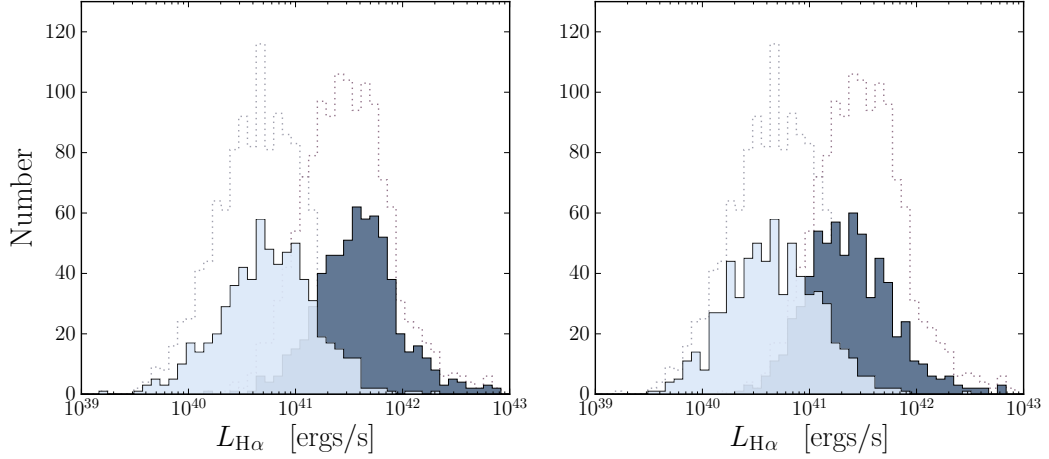


Figure 4.7. The distributions of $H\alpha$ luminosity of our Seyfert 1.9 samples. Left: 641 Seyfert 1.9 sources. Right: 656 Seyfert 1.9 sources. Dot-line represents the total 1297 Seyfert 1.9 sources. Light-blue represents the narrow component. Dark-blue represents the broad component.

estimate the t-test to check if the mean values of the $\text{FWHM}_{H\alpha, \text{broad}}$ are different in two Seyfert 1.9 samples. The t-statistic is 1.82 with a p-value = 0.07. This result suggests that we do not have enough confidence to reject the null hypothesis that the two populations have the same mean values.

We estimate the $H\alpha$ luminosity and show the distributions of $H\alpha$ luminosity in Fig. 4.7 and mean value of the distributions in Table 4.2. We find the $L_{H\alpha, \text{broad}}$ peak of the 641 Seyfert 1.9 sample steers high luminous while that of the 656 Seyfert 1.9 sample steers low luminous. The mean value of $L_{H\alpha, \text{broad}}$ in 641 Seyfert 1.9 sample is higher than that of 656 Seyfert 1.9 sample. We estimate the t-test for the two Seyfert 1.9 samples and the t-statistic is 4.74 with a p-value = 2.3×10^{-6} . The significant p-value indicates that we can reject the null hypothesis that the two populations have the same mean values. This result suggests that the 641 Seyfert 1.9 sample has an averagely luminous broad $H\alpha$ component than the 656 Seyfert 1.9 sample. The mean value in the 656 Seyfert 1.9 sample is not different from the previous result while the mean value in the 641 Seyfert 1.9 sample is slightly higher than the previous result. The distributions of broad and narrow $H\alpha$ luminosity do not show a significant difference from the previous result (Fig. 3.7). We also estimate total $H\alpha$ luminosity of our Seyfert 1.9 sources and show the results in Fig. 4.8 and Table 4.2. The total $H\alpha$ luminosity distribution of the 641 Seyfert sample is similar to the previous result. We find the total $H\alpha$ luminosity distribution of the 656 Seyfert 1.9 sample skews slightly to the low luminous while that of the 641 Seyfert 1.9 sample skews slightly to the high luminous.

We use the flux of $[\text{NII}]\lambda 6584$ derived in this fitting to plot the BPT diagram and show the result in Fig 4.9. We find most sources in the $[\text{NII}]/H\alpha$ diagram are shifting to the Seyfert region while a few sources are inside the composite region in two Seyfert 1.9 samples. In the 641 Seyfert 1.9 sample, there are 128 sources outside the AGN region. In the 656 Seyfert 1.9 sample, there are 84 sources outside the AGN region in the $[\text{NII}]/H\alpha$ diagram. This result indicates that fitting $[\text{NII}]$

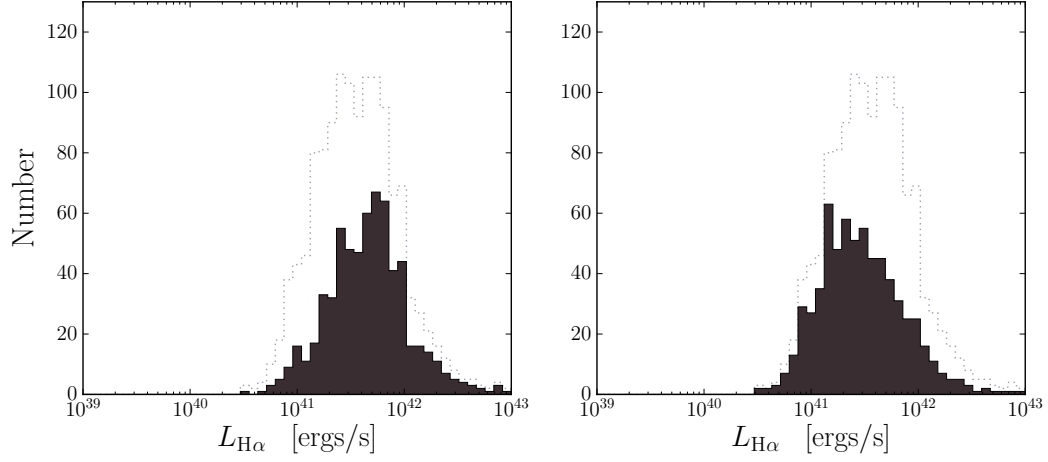


Figure 4.8. The distributions of total $H\alpha$ luminosity of our Seyfert 1.9 samples. Left: 641 Seyfert 1.9 sources. Right: 656 Seyfert 1.9 sources. Dot-line represents the total 1297 Seyfert 1.9 sources.

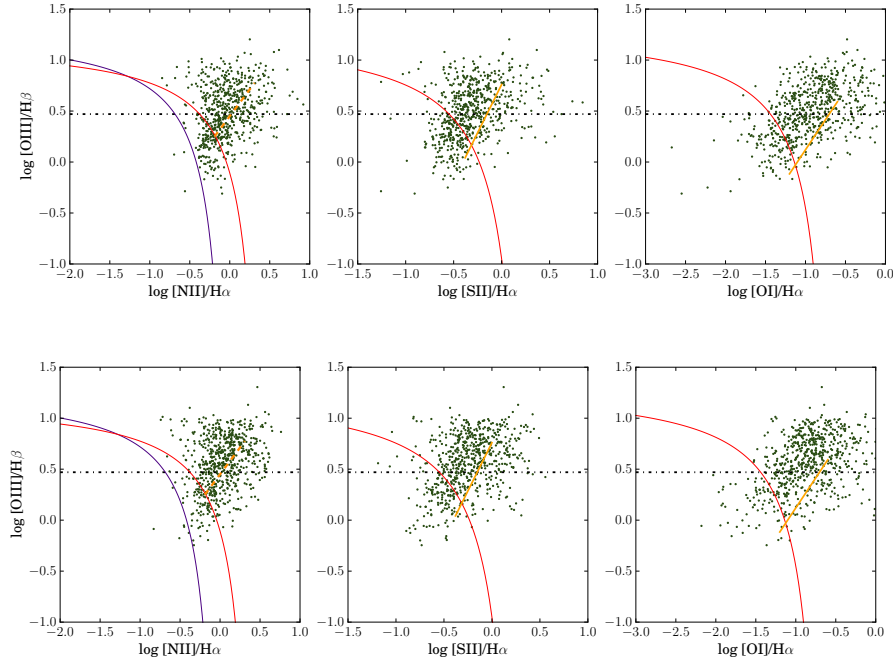
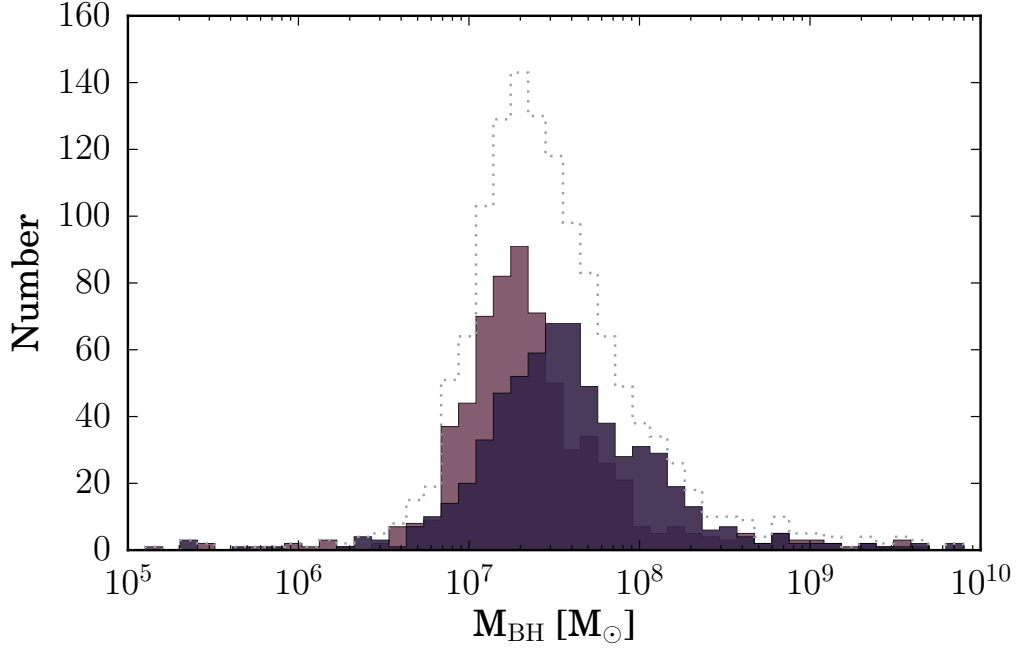


Figure 4.9. BPT diagram for our Seyfert 1.9 sample. Top: 641 Seyfert 1.9 sample. Bottom: 656 Seyfert 1.9 sample. The red line represents the definition of the starburst limit [15]. The dashed-orange line in $[\text{NII}]/H\alpha$ diagram represents the Seyfert-LINER [30]. The orange line represents the Seyfert-LINER line [16]. The purple solid line represents the AGN definition [14]. The black dashed line represents the ratio of $[\text{OIII}]/H\beta = 3$.

Table 4.2. Mean values of the Seyfert 1.9 that are fitted with consideration of [NII] lines

	1297 Sy19	641 Sy19	656 Sy19	
FWHM _{Hα} ,broad	4603.21 (3312.69)	4772.46 (3076.48)	4437.82 (3520.39)	[km/s]
FWHM _{Hα} ,narrow	339.88 (127.93)	335.03 (116.38)	344.63 (138.13)	[km/s]
L _{Hα} ,broad	5.47×10^{41} (1.20×10^{42})	7.05×10^{41} (1.56×10^{42})	3.92×10^{41} (6.34×10^{41})	[ergs/s]
L _{Hα} ,narrow	8.52×10^{40} (1.44×10^{41})	9.33×10^{40} (1.40×10^{41})	7.74×10^{40} (1.47×10^{41})	[ergs/s]
L _{Hα}	6.32×10^{41} (1.28×10^{42})	7.99×10^{41} (1.64×10^{42})	4.70×10^{41} (7.27×10^{41})	[ergs/s]

The values inside parenthesis are standard deviations of the samples.

**Figure 4.10.** The distributions of SMBH mass of our Seyfert 1.9 samples. Dark-purple represents the 641 Seyfert 1.9 sample. Twilight-lavender represents the 656 Seyfert 1.9 sample

emission lines without consideration of H α might cause an overestimate of [NII]. We find the BPT distributions of Seyfert 1.9 galaxies are similar to Seyfert 2 galaxies and this suggests that the narrow-line region of the Seyfert 1.9 galaxies have similar ionization levels with Seyfert 2 galaxies [21].

We estimate the SMBH mass of our Seyfert 1.9 sources and show the result in Fig. 4.10. We find the SMBH mass distributions distribute from 10^5 to 10^{10} M_{\odot} and the distribution range is broader than the previous result (Fig. 4.5). We find the 641 Seyfert 1.9 sample shows a small component around 10^8 M_{\odot} . The mean value of the SMBH mass distribution for the 641 Seyfert 1.9 sample is 1.42×10^8 (7.11×10^8) M_{\odot} and for the 656 Seyfert 1.9 sample is 1.11×10^8 (6.65×10^8) M_{\odot} . Both of the Seyfert 1.9 samples have a higher mean value of SMBH mass than the previous result and it might relate to the broader distributions in Fig. 4.10. The t-statistic for the tow mean values is 0.82 with a p-value = 0.41. The very high p-value indicates that we

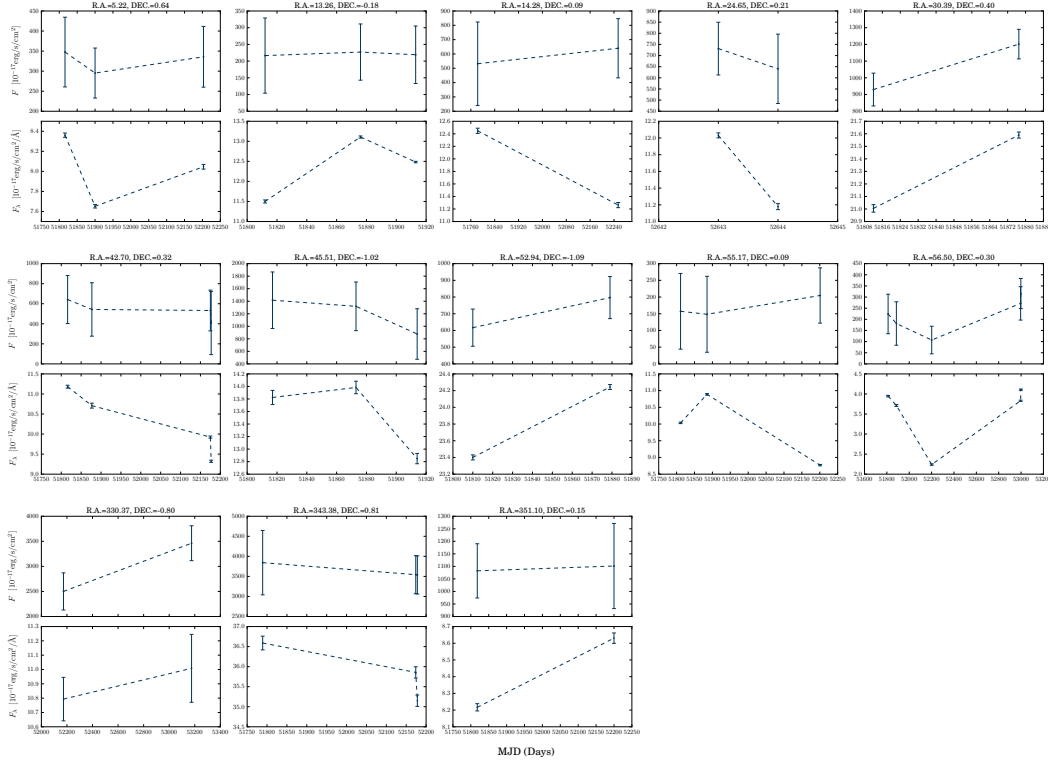


Figure 4.11. Light curves of the 13 Seyfert 1.9 sources.

can not reject the null hypothesis that the two populations have the same mean values. We also estimate the percentage of the SMBH mass larger than $10^8 M_{\odot}$. The percentage for the 641 Seyfert 1.9 sample is 19.2% (123/641) and for the 656 Seyfert 1.9 sample is 8.4% (55/656). We find no matter we consider [NII] in our fitting process or not, the SMBH mass distributions of our Seyfert 1.9 galaxies are dominated in $10^7 M_{\odot}$.

I also refit the spectra of our 13 Seyfert sources with multiple observations and the fitting results are in Appendix A. We show light curves in Fig. 4.11 and find three sources show different behavior in their light curves than the previous result. The first one is a source with R.A.=13.26 and DEC.=−0.18. From MJD=51876 to MJD=51913, the light curve of broad H α component and continuum level decrease. Therefore, the previous result that is an increasingly broad H α is caused by overestimating the broad H α component. The second one is a source with R.A.=14.28 and DEC.=0.09. The broad H α component increases with decreasing continuum level in 471 days. The third one is a source with R.A.= 55.17 and DEC.=0.09. This source still shows an increasing broad H α component with a decreasing continuum level from MJD=51885 to MJD=52201 in this fitting result. The variation time is 316 days (MJD=51811 to MJD=51885) and it shows a decreasing broad H α component with an increasing continuum level. This result is different behavior from the previous result. However, the large uncertainty in the broad H α component is hard to conclude if the variation is significant. I also show the broad H α profiles of 13 Seyfert 1.9 in Fig. 4.12 and find the profiles only show variations in their amplitude.

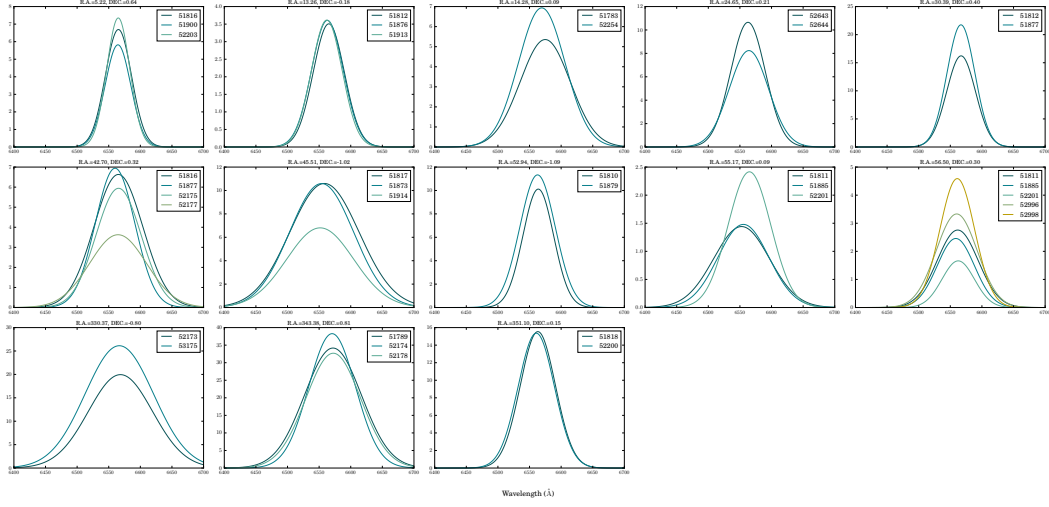


Figure 4.12. Broad H α profiles of 13 Seyfert 1.9 sources.

4.5 Host galaxy morphology

FracDev is an indicator of host galaxy morphology. This parameter describes the bulge contribution in galaxies and ranges from 0 to 1. A bulge-dominated galaxy has *FracDev* close to 1. We show the *FracDev* of our Seyfert 1.9 sample in Fig. 4.13. We find that both 641 Seyfert 1.9 and 656 Seyfert 1.9 samples have similar distributions which are dominant by *FracDev*=1. These results indicate that Seyfert 1.9 galaxies are more likely residing in bulge-dominated galaxies. We estimate the K-S test for these two Seyfert 1.9 samples and obtain statistic $D = 0.0967$ with a p-value = 0.0042, indicating that the two distributions have a low probability to be drawn from the same population. Besides, we compare our Seyfert 1.9 to Seyfert 1 and Seyfert 2 galaxies with the same redshift range [6]. We show the results of *FracDev* distributions in Fig. 4.13. We find that the distribution of Seyfert 1.9 is similar to the distribution of Seyfert 1 galaxies and the two distributions are dominated by *FracDev* = 1. We estimate the K-S test for Seyfert 1.9 and Seyfert 1 and obtain statistic $D = 0.0602$ with a p-value = 0.0014. The low p-value suggests that Seyfert 1.9 galaxies and Seyfert 1 galaxies have a low probability to be drawn from the same population. This result indicates that although Seyfert 1.9 galaxies are dominated by *FracDev*=1, its host galaxy morphology distribution is still different from Seyfert 1 galaxies. On the other hand, the *FracDev* distribution of Seyfert 1.9 galaxies is different from that of Seyfert 2 galaxies. We estimate a K-S test for Seyfert 1.9 and Seyfert 2 and obtain statistic $D = 0.2545$ with a p-value = 8.56965×10^{-54} . This suggests that the different types of the Seyferts might be related to its host galaxy morphology.

The [OIII] luminosity is isotropic and related to the power of AGNs. We limit our Seyfert to the same [OIII] luminosity to see if they have different morphology distributions. Our Seyfert 1.9 are pick out from a Seyfert 2 sample with $\log L_{\text{[OIII]}} [\text{ergs/s}] > 40.125$ [7], we compare the host galaxy morphology of the Seyfert 2 sample and our Seyfert 1.9 sample. We show the result of host galaxy morphology distribution in Fig. 4.14. We find that although Seyfert 1.9 and Seyfert 2 have

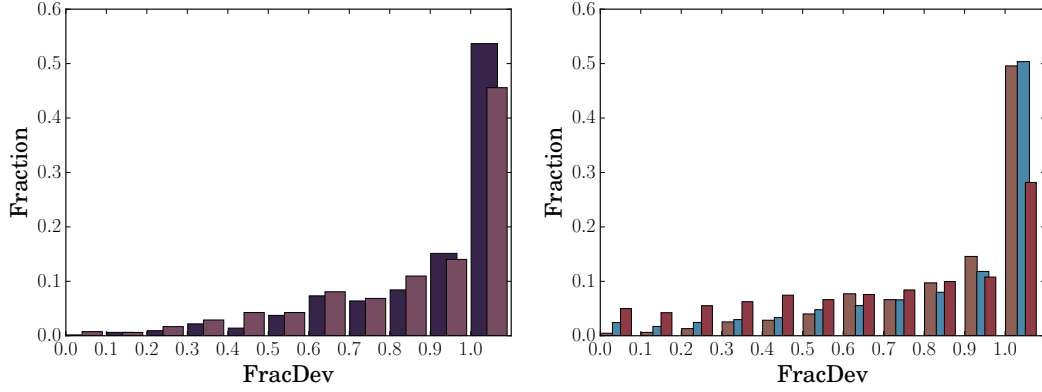


Figure 4.13. *FracDev* distributions. Left: Seyfert 1.9 sample. Dark-purple represents the 641 Seyfert 1.9 sample. Twilight-lavender represents the 656 Seyfert 1.9 sample. Right: Seyfert 1.9, Seyfert 1, and Seyfert 2. Brown represents the total 1297 Seyfert 1.9 sample. Steel-blue represents Seyfert 1 sample. Solid-pink represents Seyfert 2 sample.

similar [OIII] luminosity range, the distributions of their host galaxy mythologies are different. Seyfert 1.9 galaxies dominate $FracDev = 1$ whereas Seyfert 2 galaxies are dominated by $FracDev < 1$. This result indicates that the host galaxy of Seyfert 1.9 is more elliptical galaxies and Seyfert 2 is more disk galaxies. We estimate the K-S test for the Seyfert 1.9 galaxies and Seyfert 2 galaxies and find the $D = 0.18907$ with a $p\text{-value} = 3.0283 \times 10^{-37}$. This result indicates that Seyfert 1.9 and Seyfert 2 have an extremely low probability to draw from the same population. We also notice that the Seyfert 1.9 have relatively a high [NII]/H α value than Seyfert 2 galaxies. These results suggest that the difference between Seyfert 1.9 and Seyfert 2 is related to their host galaxy.

4.6 compare to selection method

In this work, we provide 641 Seyfert 1.9 galaxies and additional 656 Seyfert 1.9 galaxies which are picked out by our model with redshift range from 0 to 0.2. We compare the amount of our Seyfert 1.9 sample to Véron 13th catalog [42]. The amount of Seyfert 1.9 with the same redshift range in the Véron catalog is 151. Our amount of the Seyfert 1.9 galaxies is 1297 and is almost 10 times greater than the Véron catalog.

Due to the complicated shape of the emission line of the intermediate Seyferts, it is not easy to pick these sources out from the observation data. A traditional method is visual inspection. However, the method is not efficient when dealing with a huge amount of observation data. There are several different ways to select the intermediate Seyfert from huge observation data in the literature. One example is to use the excess flux near the wing of H α emission in blue wavelength [44]. An alternative and similar example is to use the excess flux near narrow H α component after continuum subtract [36]. Another way is to fit two components to H α emission line to see if the source exists the second component [41]. These methods shown above need to fit enormous candidate spectra at first and the fitting process usually

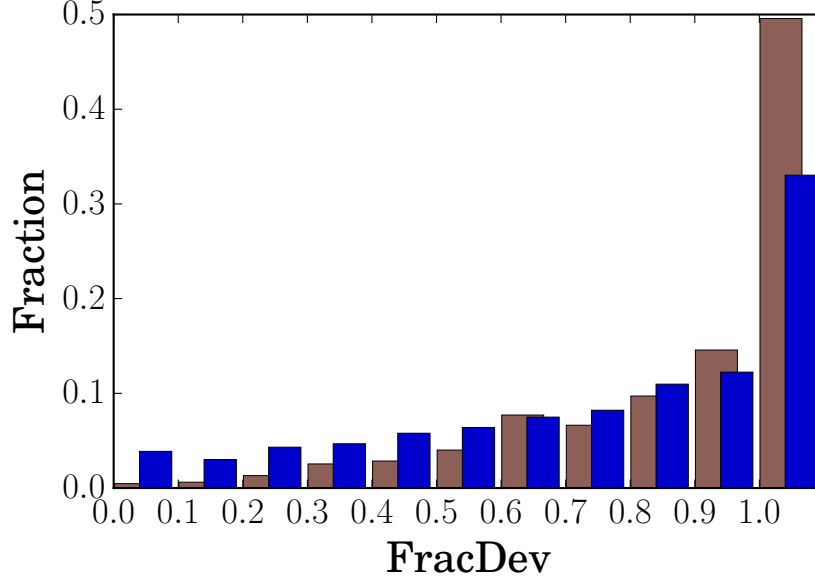


Figure 4.14. *FracDev* distributions. Brown represents our 1297 Seyfert 1.9 galaxies. Blue represents the Seyfert 2 galaxies with $\log L_{[\text{OIII}]} [\text{ergs/s}] > 40.125$.

takes a long time. Besides the difficult and lengthy picking way, the relatively weak broad $\text{H}\alpha$ of the Seyfert 1.9 galaxies usually cause contamination in the Seyfert 2 sample. The way to check if the Seyfert 2 sample exists Seyfert 1.9 galaxies is to stack the spectra of Seyfert 2 galaxies and to find if the stack result shows a weak wing. However, this method is still not efficient. Our model can pick out the Seyfert 1.9 galaxies without fitting the number of candidate spectra and can divide the Seyfert 1.9 sources from the Seyfert 2 sample in a short time. We provide a different way to solve the selection problem of the Seyfert 1.9 galaxies efficiently. For the training time, our model spends two hours for training and the test process just needs a few minutes. Our selection is quite fast and saves a lot of time because we do not fit the candidate spectrum that does not have two components in the $\text{H}\alpha$ emission line. After obtaining a sample of Seyfert 1.9 galaxies, we can focus on fitting two components of the $\text{H}\alpha$ emission line in the spectra of the Seyfert 1.9 galaxies.

Although early classification for intermediate Seyfert galaxies in the literature using flux ratio as an indicator is convenient, the flux and line width represent different physic parameters and do not have necessary dependence. We show the flux of broad $\text{H}\alpha$ component and FWHM of broad $\text{H}\alpha$ in Fig. 4.15. The correlation coefficient between flux of broad $\text{H}\alpha$ component and FWHM is 0.15 with a p-value= 3.32×10^{-8} . This indicates that flux of the broad $\text{H}\alpha$ component of our Seyfert 1.9 shows weak dependence on FWHM. This result indicates that using a flux or FWHM for picking out intermediate Seyfert is not a complete way and might lose some sources. Our CNN model has ability to pick out the intermediate Seyfert galaxies in an independent way. Besides, we notice that a weak broad $\text{H}\alpha$ component can be regarded as a fake feature when its flux is below $10^{-16} \text{ ergs/s/cm}^2$ [20]. We find the minimum flux of the broad $\text{H}\alpha$ component is $10^{-15} \text{ ergs/s/cm}^2$ and is large

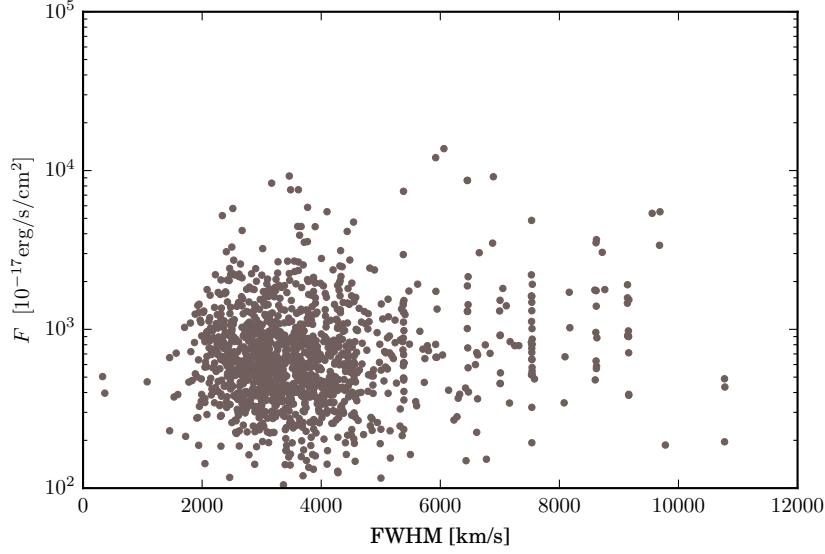


Figure 4.15. Flux and FWHM of the broad $H\alpha$ of our Seyfert 1.9.

than the flux limit.

Our Seyfert 1.9 galaxies were picked out from a sample of the Seyfert 2 spectra with $\log L_{[\text{OIII}]}$ [ergs/s] > 40.125 . This shows that a sample of the Seyfert 2 spectra does contaminate by Seyfert 1.9 galaxies. The Seyfert 2 are selected by line flux and BPT region, the weak broad $H\alpha$ makes some Seyfert 1.9 to be selected into this sample. Picking out Seyfert 1.9 spectra by visual inspection is time-consuming. By training our CNN model, it can pick out the Seyfert 1.9 galaxies from the Seyfert 2 sample. The training process takes few hours and testing the whole Seyfert 2 sample only takes few minutes. This indicates that our model can solve the contamination from Seyfert 1.9 galaxies in the Seyfert 2 spectra in a quick way.

The select method/picking out in the thesis provides a good and efficient way to collect intermediate Seyfert, especially Seyfert 1.9 galaxies. This method also reduces the contamination of Seyfert 1.9 in the Seyfert 2 spectra.

Intermediate Seyfert galaxies are hard to be picked out from observation data because of their complicated line shape. However, these sources are important to the unification model. By collecting a large sample of these sources, we can learn their properties and gain our knowledge to the low luminous end. In this thesis, we especially collect Seyfert 1.9 galaxies and provide the properties of the low luminous AGNs and low BH mass. A large sample can provide constrains for models. A catalog of broad-line AGNd indicates their sources with $L_{H\alpha, \text{broad}} < 10^{41}$ [ergs/s] is 8% [20]. In our sample, we have 10% sources with $L_{H\alpha, \text{broad}} < 10^{41}$ [ergs/s]. We hope that these Seyfert 1.9 sources identified in this thesis will provide people to investigate its properties in different wavelength in the future.

Chapter 5

Summary

We build a 1D CNN model to collect Seyfert 1.9 sources from Seyfert 2 sample. The model trained by intermediate Seyfert or Seyfert 1.9 galaxies has more than 80% precision for classifying correctly Seyfert 1.9 galaxies. Besides, we use the new Seyfert 1.9 sample as the training sample to improve the performance of our model and obtain 91% precision for classifying correctly Seyfert 1.9 galaxies.

In this work, we have the original 641 Seyfert 1.9 sources that are picked out by visual inspection and 656 Seyfert 1.9 sources that are picked out by our model. In total, we identify 1297 Seyfert 1.9 galaxies and 157 sources (Seyfert 1.2, 1.5, and 1.8) with a broad $H\beta$ component. We find the two Seyfert 1.9 samples have similar mean value in FWHM distribution but different mean value in broad $H\alpha$ luminosity distribution; the Seyfert 1.9 sample that is picked out by our model has lower luminosity than the Seyfert 1.9 sample that is picked out by visual inspection. This suggests that our model picks out the relatively weak Seyfert 1.9 sources that are usually missed by visual inspection. The SMBH mass distributions of the two Seyfert 1.9 samples are dominated by $10^7 M_\odot$ and we find the two Seyfert 1.9 samples have similar mean value in the SMBH mass distribution. The BPT diagrams of our Seyfert 1.9 samples have similar distributions with Seyfert 2 galaxies. We check the distribution of host galaxy morphology of the Seyfert 1.9 samples and find our Seyfert 1.9 galaxies are dominated by elliptical galaxies and the distributions of host galaxy morphology of the Seyfert 1.9 are more similar to Seyfert 1 galaxies instead of Seyfert 2 galaxies. This suggests that the difference between different types of Seyfert galaxies might be related to the host galaxy morphology.

We investigate our Seyfert 1.9 sample with Stripe 82 survey for possible spectral variation and only find 13 Seyfert 1.9 sources with multiple spectra. We find both of the continuum and broad $H\alpha$ component show variation.

In this thesis, we present a new method for picking out the intermediate Seyfert galaxies, especially Seyfert 1.9 galaxies, in an efficient way. This method is good at dealing with huge observation data without taking a long time on fitting candidate spectra. Besides, it can collect Seyfert 1.9 sample from the Seyfert 2 spectra and remove the contamination of the Seyfert 1.9 spectra in the Seyfert 2 sample at the same time.

Appendix A

Fitting results of our 13 sources

Our 13 Seyfert 1.9 sources have total 36 spectra and we show the fitting results of the 36 spectra in the following.

A.1 Do not consider $[\text{NII}]\lambda 6548,6584$ in fitting process

A.2 Consider $[\text{NII}]\lambda 6548,6584$ in fitting process

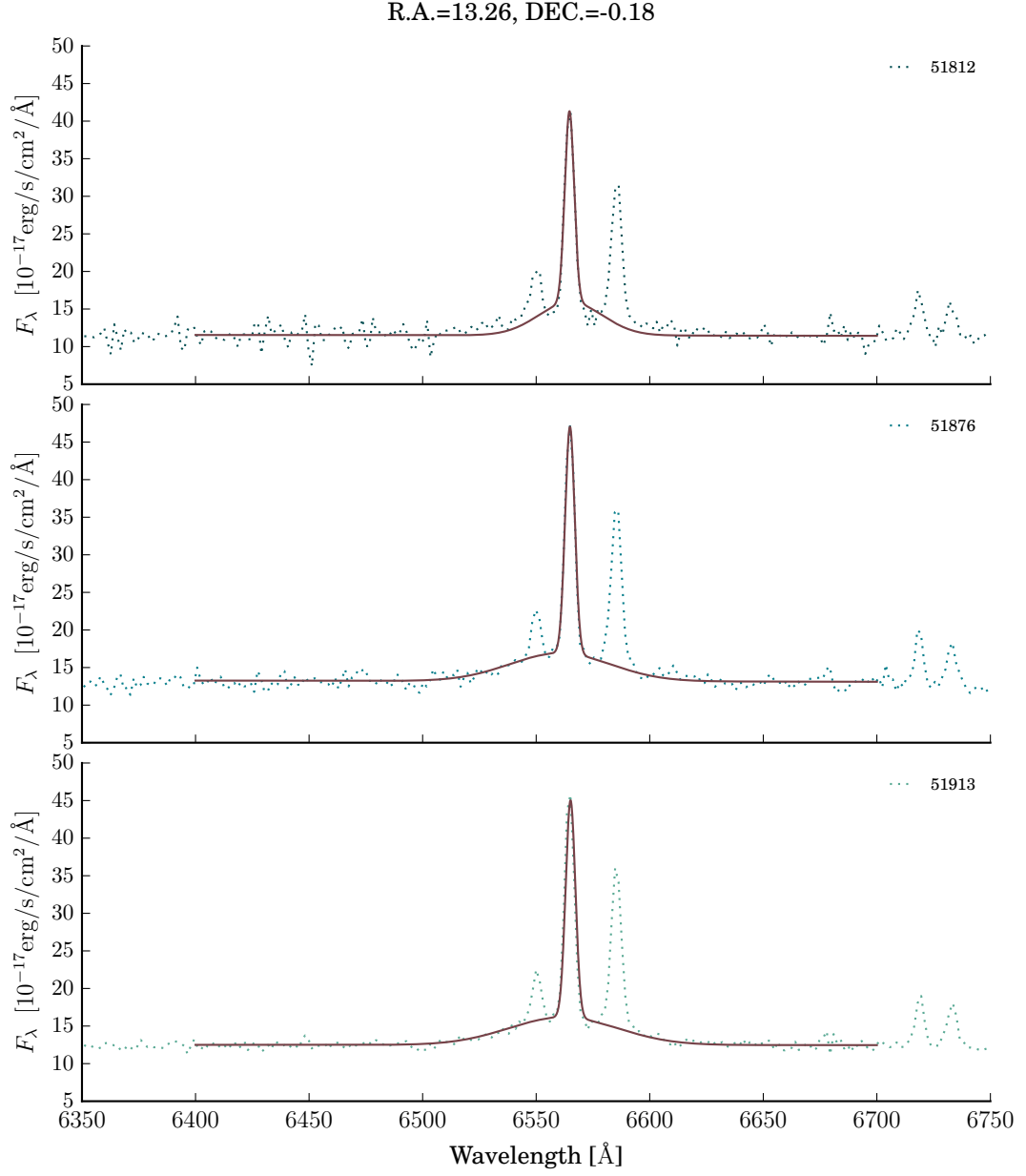


Figure A.1. Fitting results of two Gaussian components. Each panel represents different observation day. The dotted line represents observation spectrum. The solid line represents fitting results.

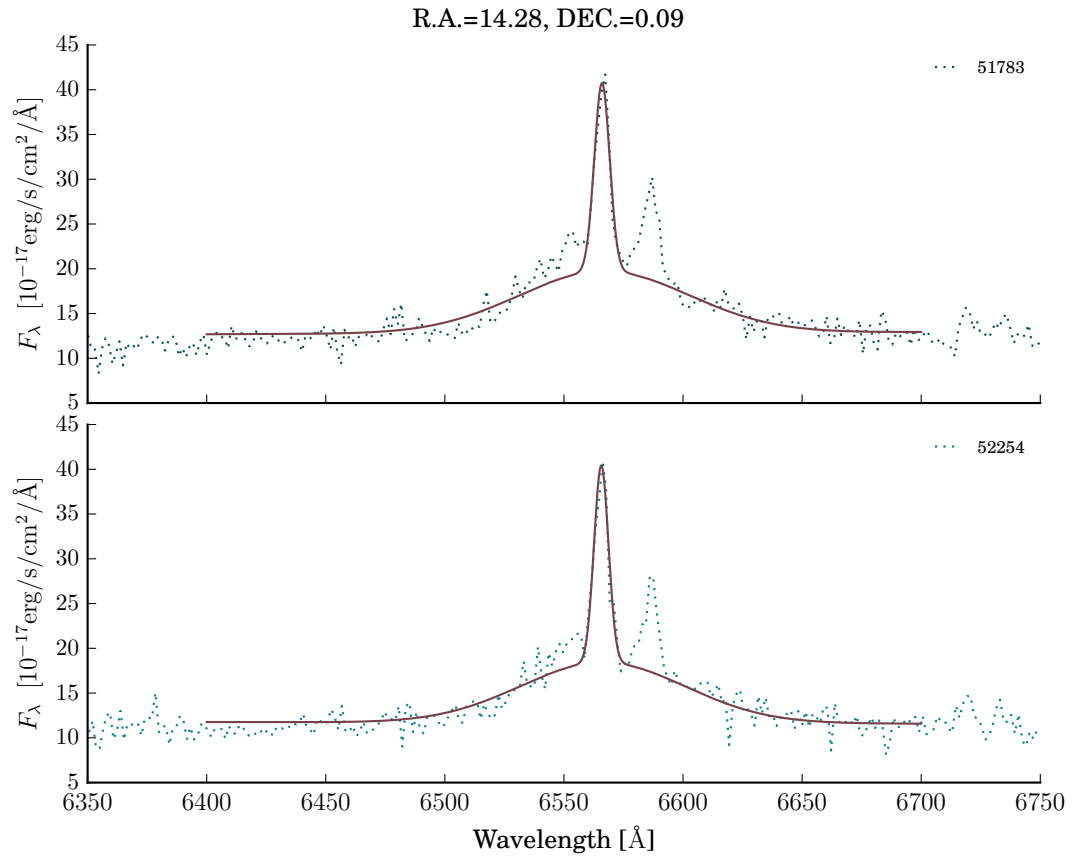


Figure A.2. Fitting results of two Gaussian components. Each panel represents different observation day. The dotted line represents observation spectrum. The solid line represents fitting results.

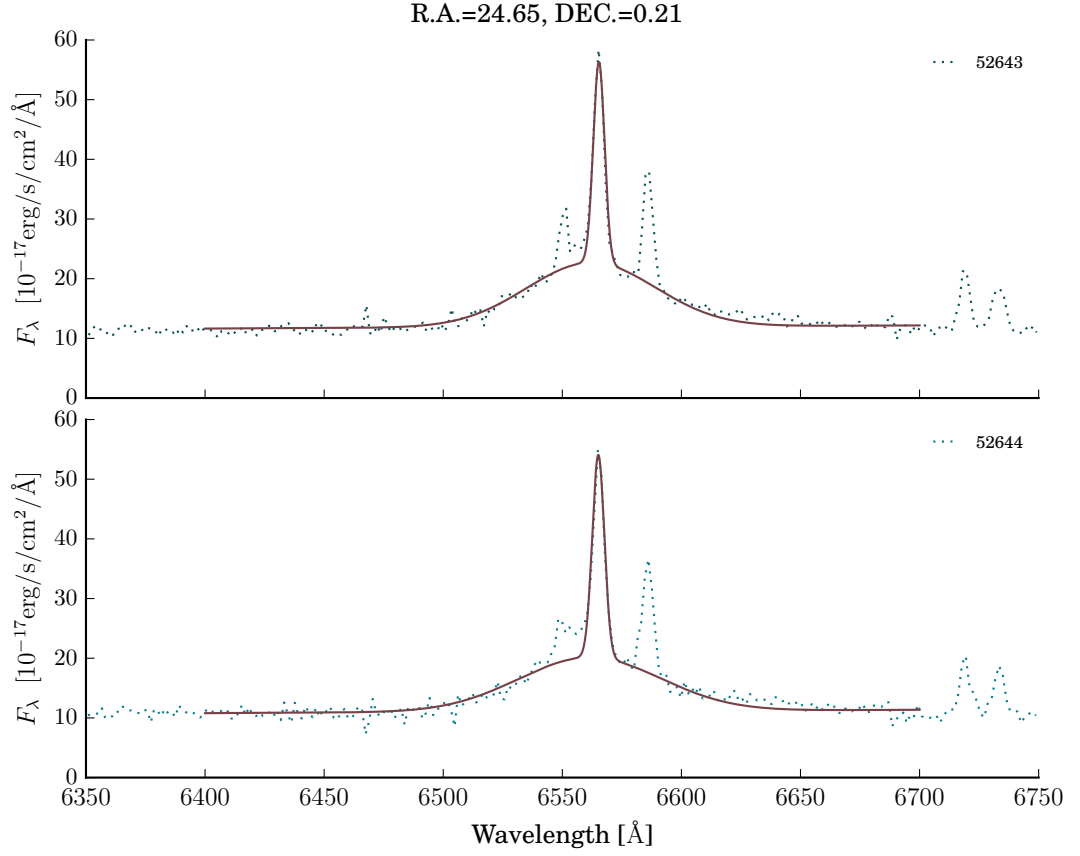


Figure A.3. Fitting results of two Gaussian components. Each panel represents different observation day. The dotted line represents observation spectrum. The solid line represents fitting results.

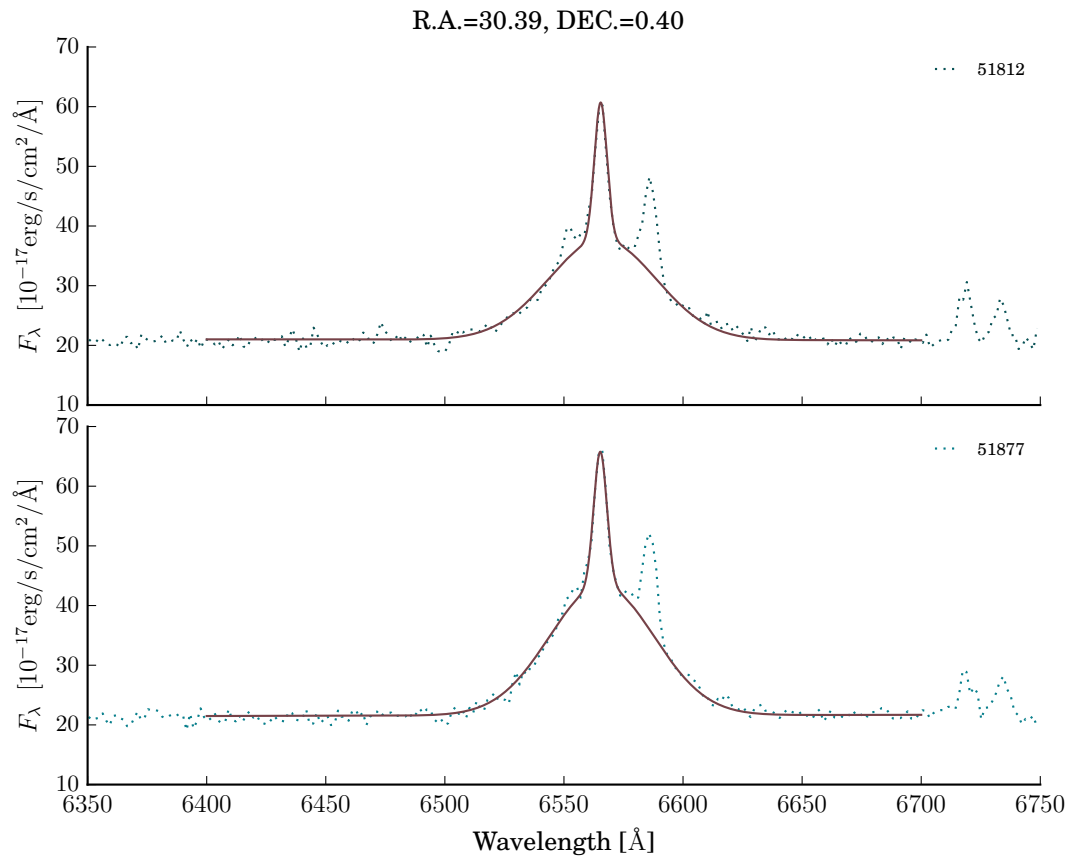


Figure A.4. Fitting results of two Gaussian components. Each panel represents different observation day. The dotted line represents observation spectrum. The solid line represents fitting results.

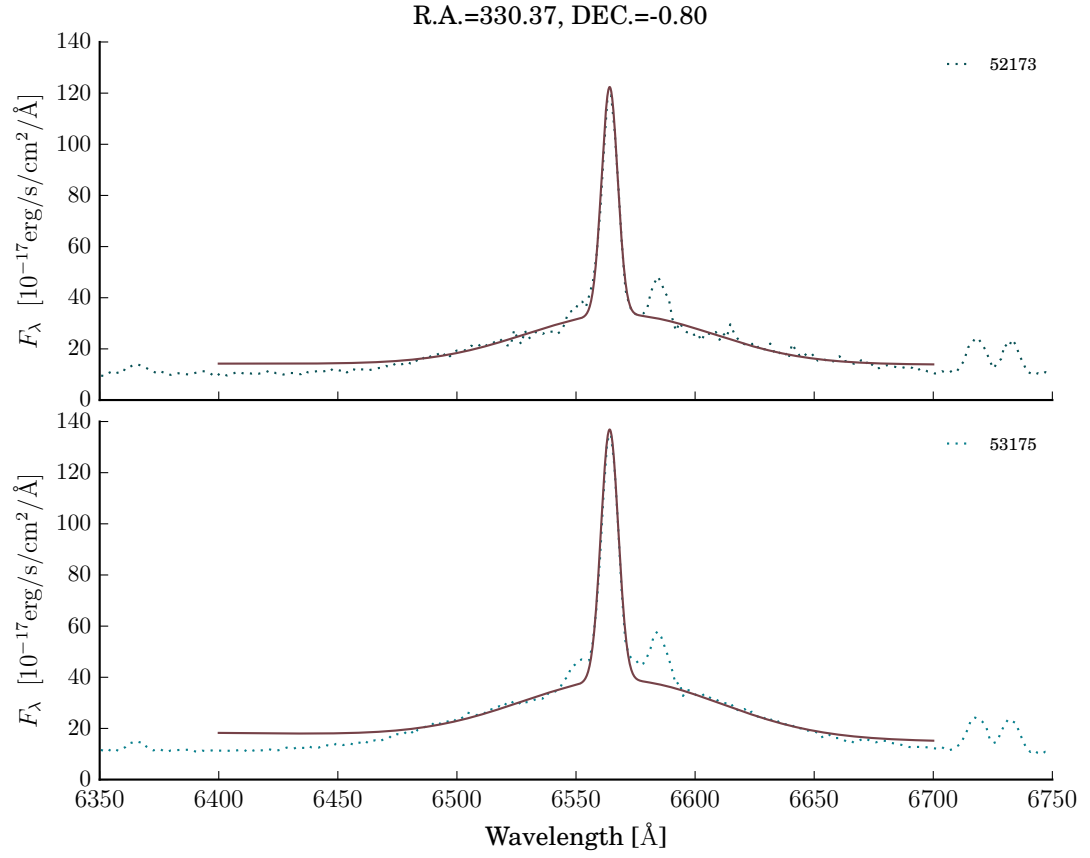


Figure A.5. Fitting results of two Gaussian components. Each panel represents different observation day. The dotted line represents observation spectrum. The solid line represents fitting results.

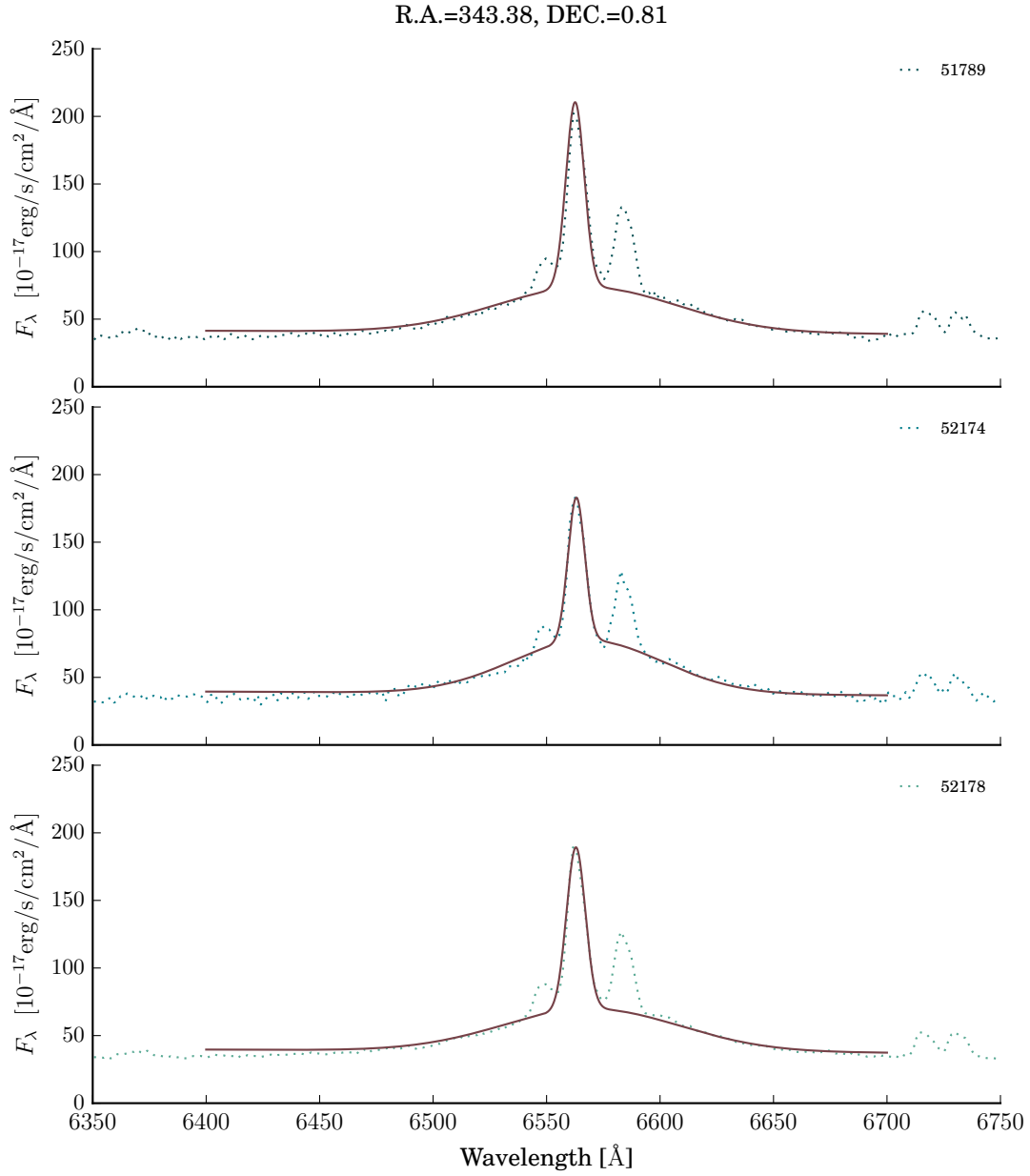


Figure A.6. Fitting results of two Gaussian components. Each panel represents different observation day. The dotted line represents observation spectrum. The solid line represents fitting results.

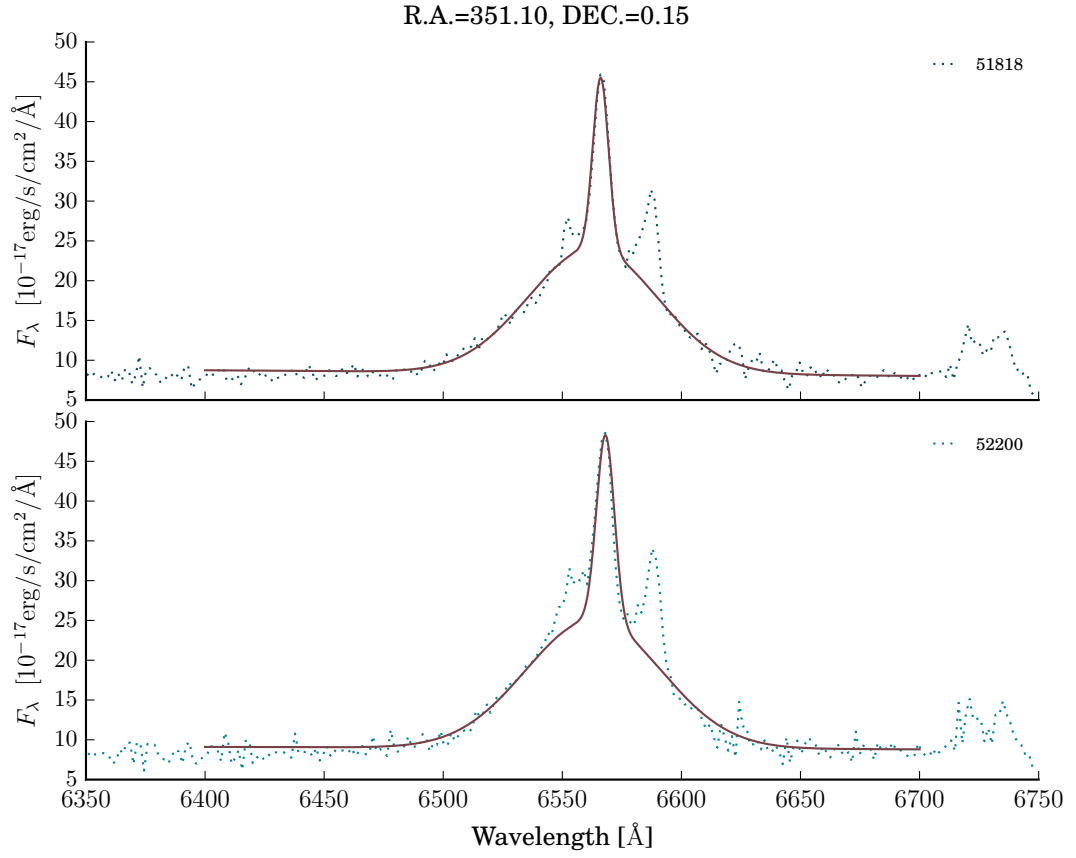


Figure A.7. Fitting results of two Gaussian components. Each panel represents different observation day. The dotted line represents observation spectrum. The solid line represents fitting results.

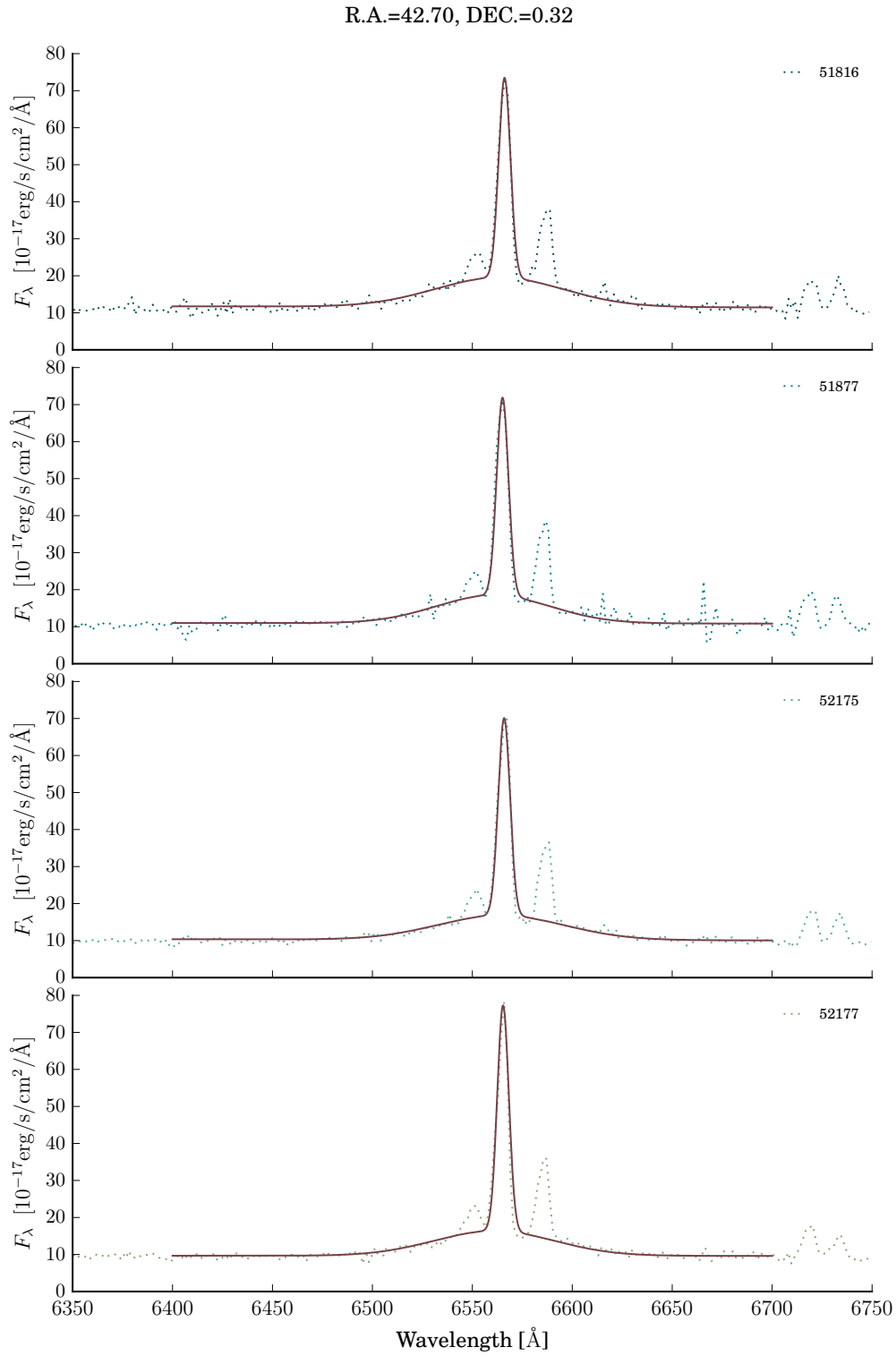


Figure A.8. Fitting results of two Gaussian components. Each panel represents different observation day. The dotted line represents observation spectrum. The solid line represents fitting results.

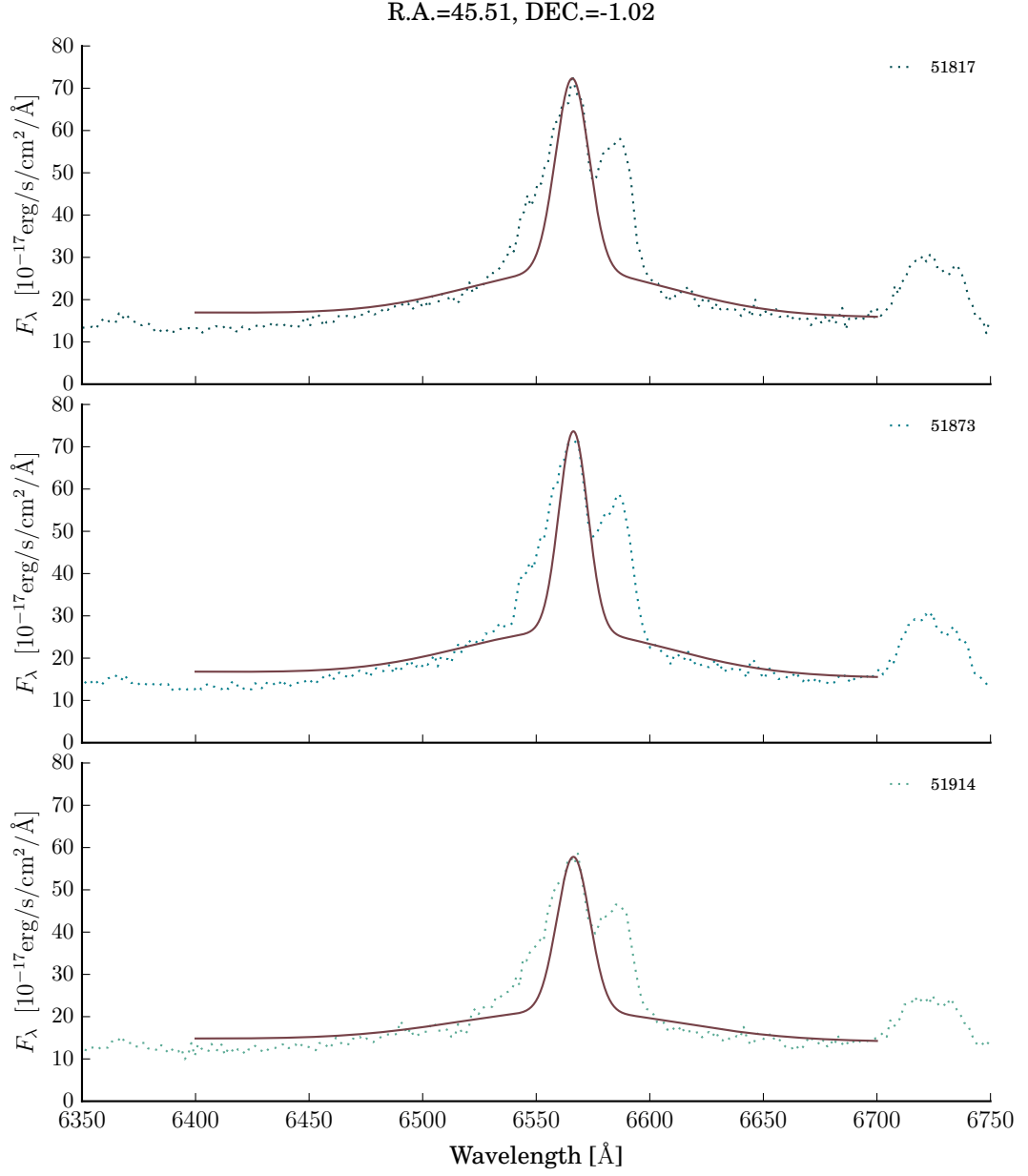


Figure A.9. Fitting results of two Gaussian components. Each panel represents different observation day. The dotted line represents observation spectrum. The solid line represents fitting results.

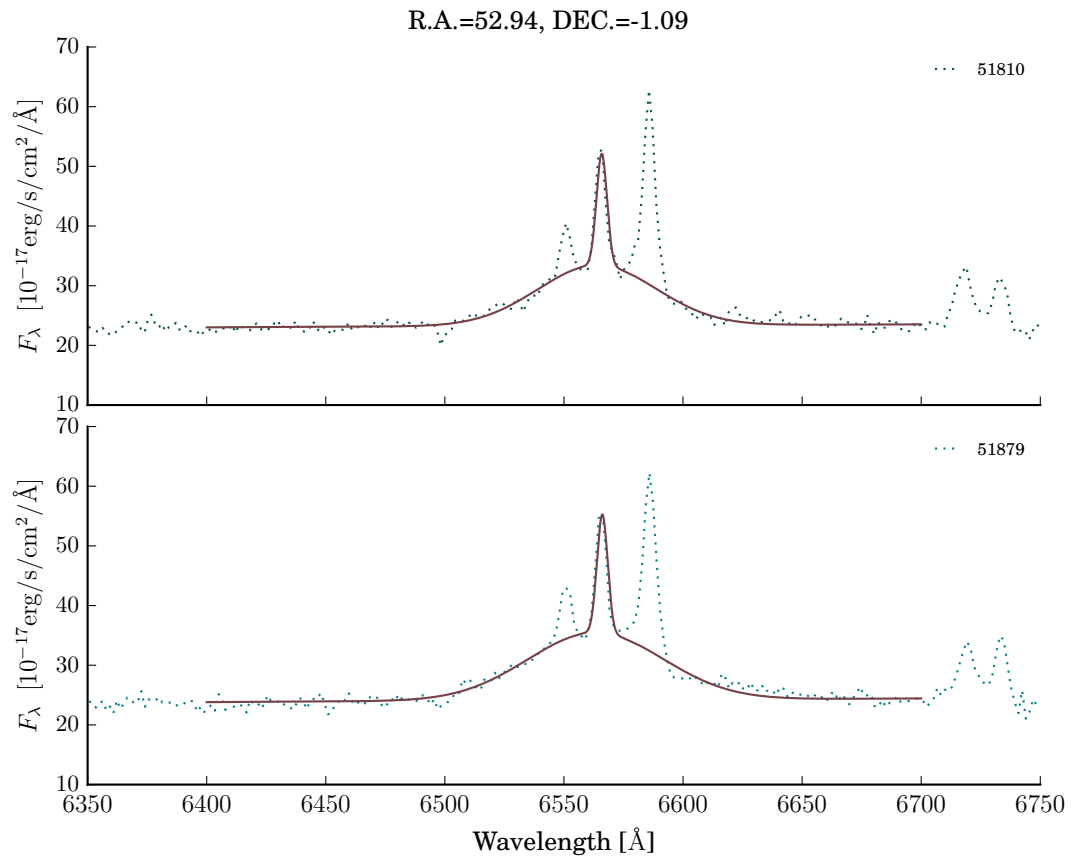


Figure A.10. Fitting results of two Gaussian components. Each panel represents different observation day. The dotted line represents observation spectrum. The solid line represents fitting results.

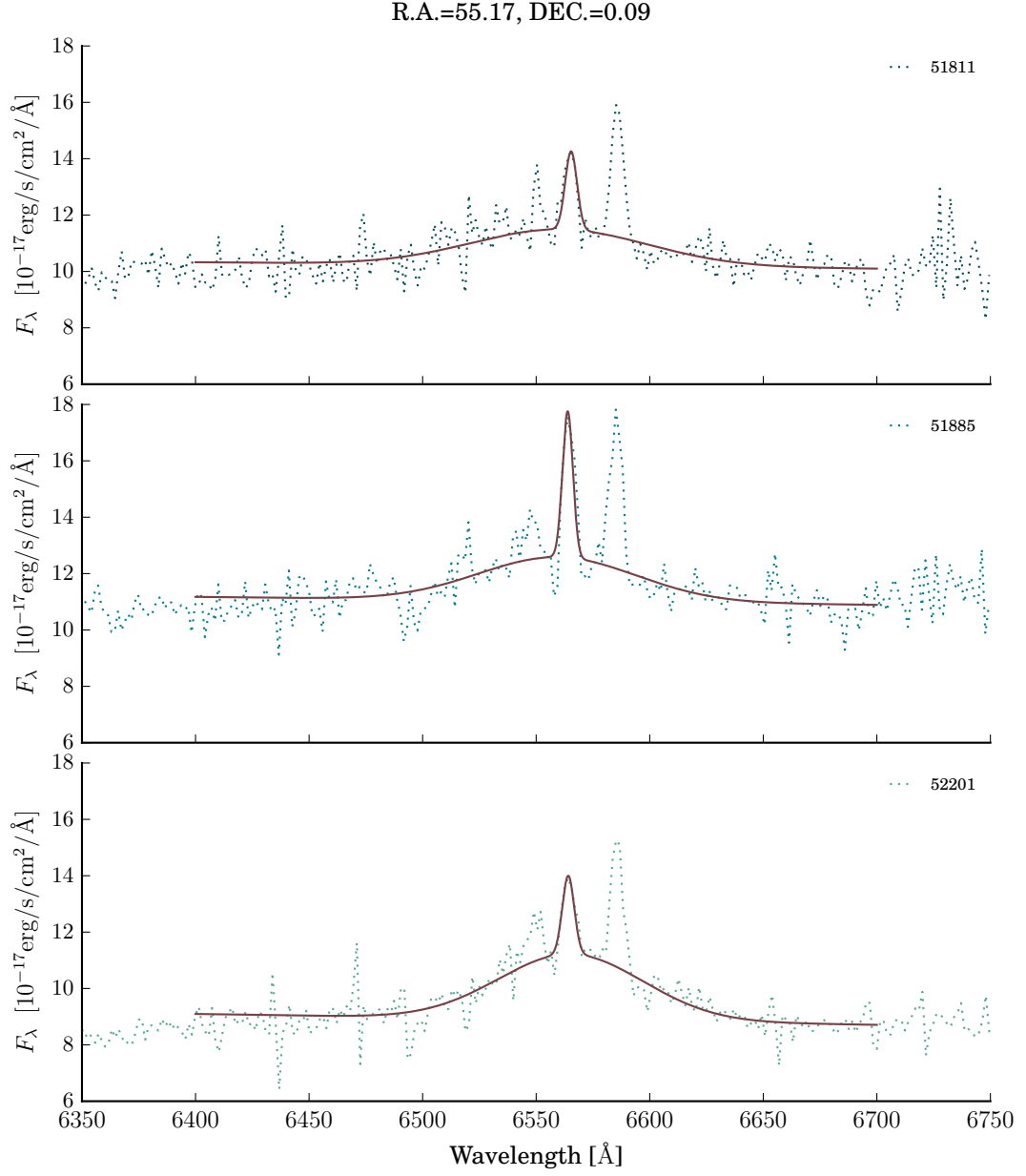


Figure A.11. Fitting results of two Gaussian components. Each panel represents different observation day. The dotted line represents observation spectrum. The solid line represents fitting results.

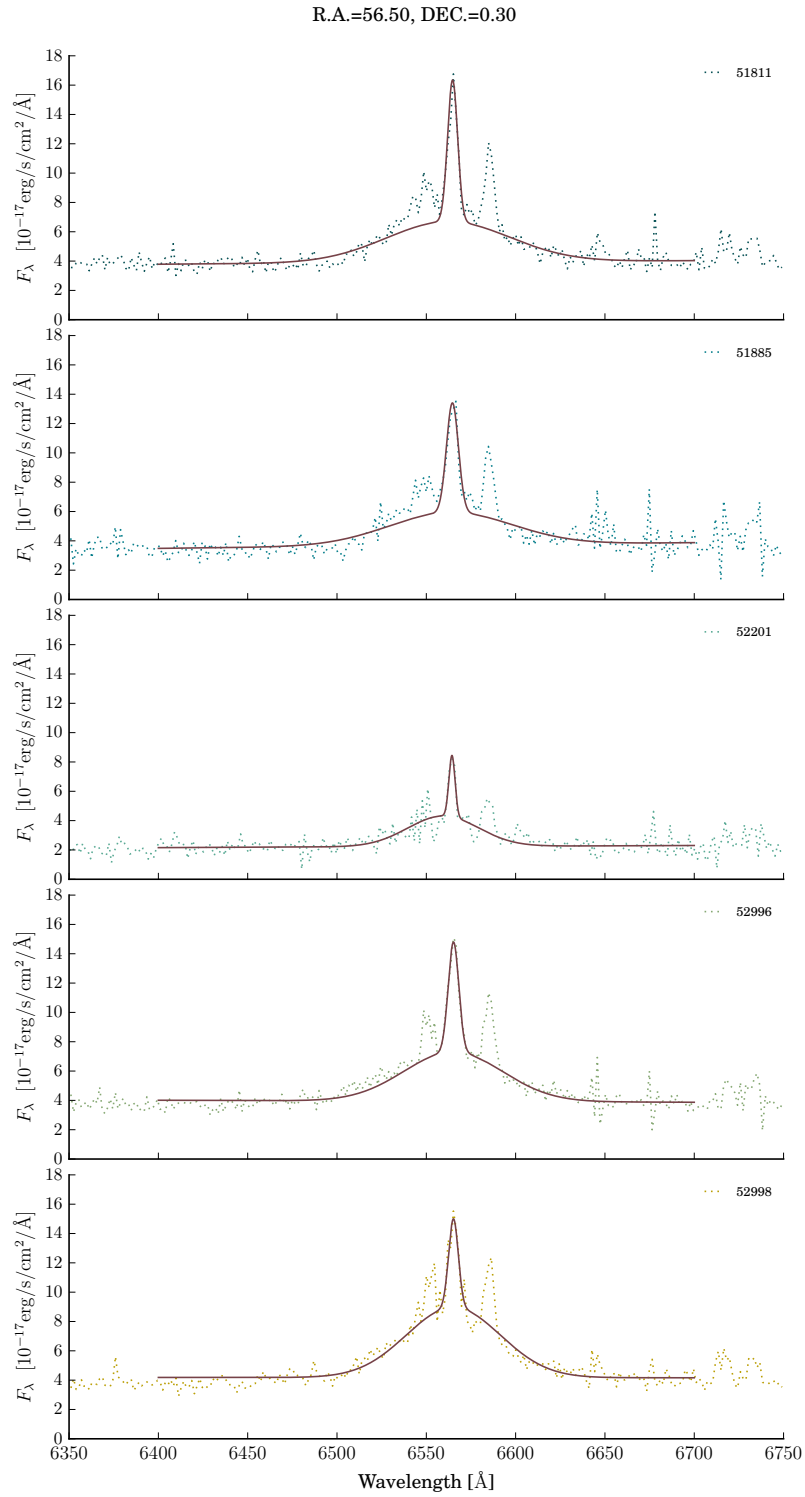


Figure A.12. Fitting results of two Gaussian components. Each panel represents different observation day. The dotted line represents observation spectrum. The solid line represents fitting results.

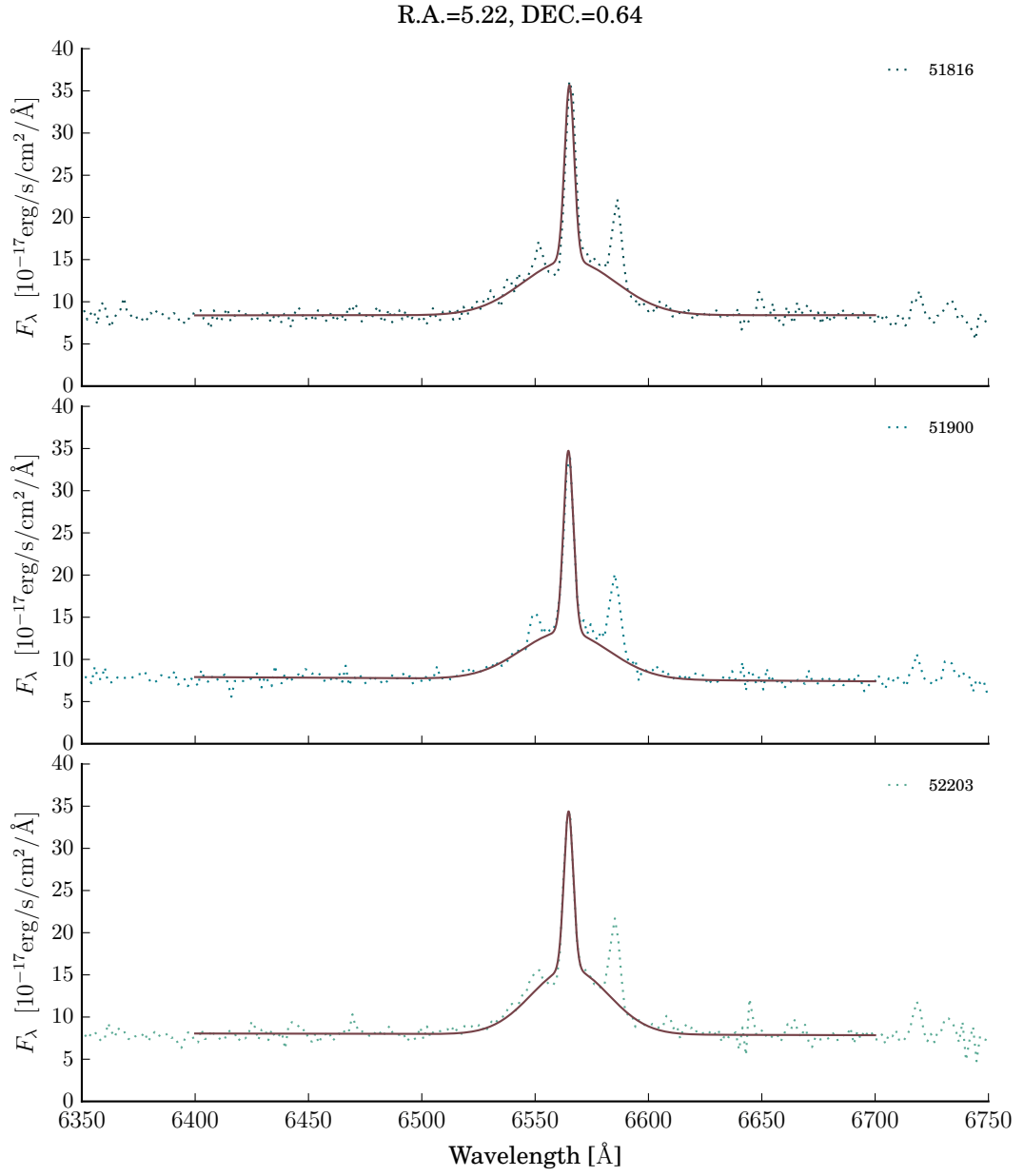


Figure A.13. Fitting results of two Gaussian components. Each panel represents different observation day. The dotted line represents observation spectrum. The solid line represents fitting results.

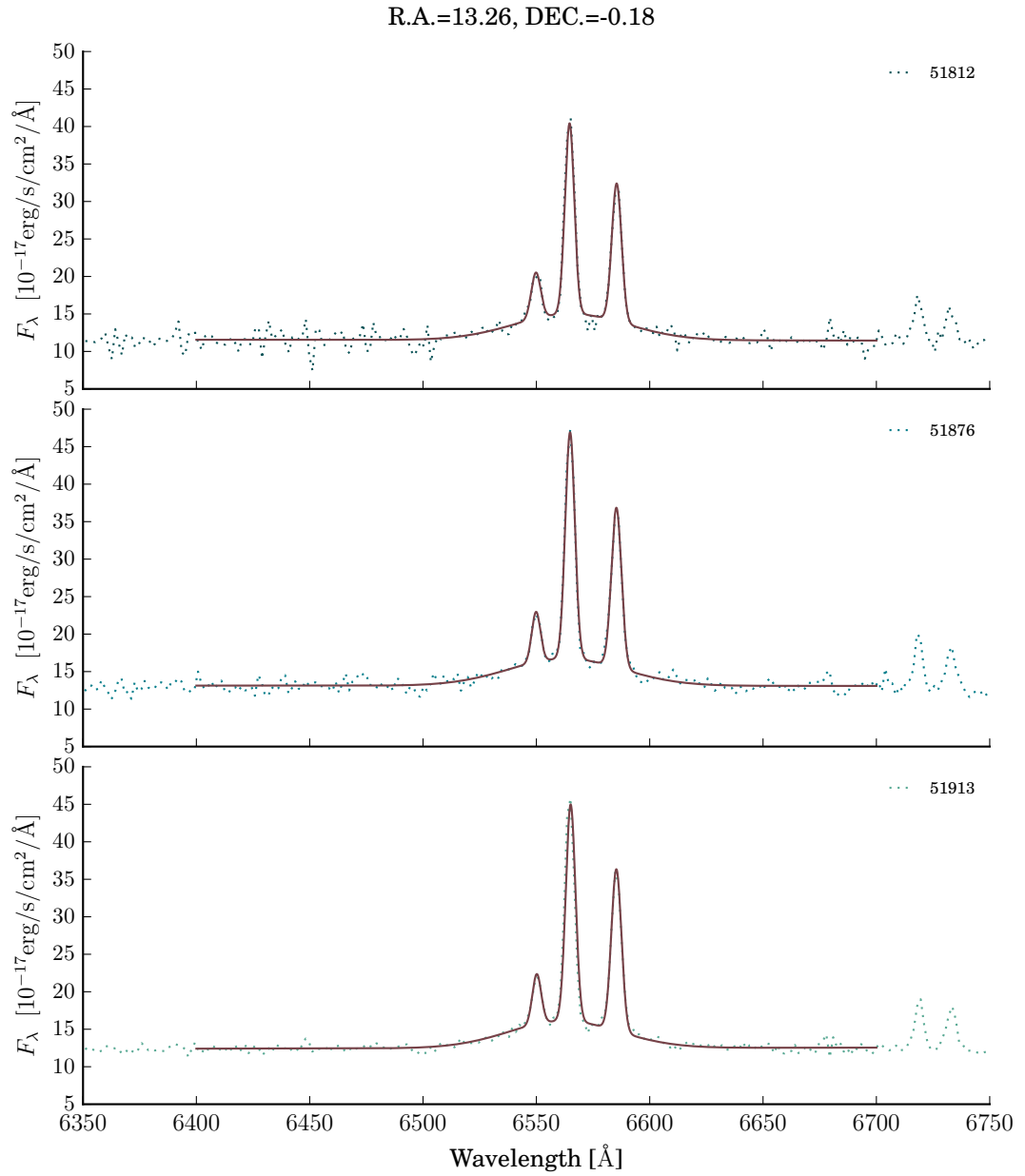


Figure A.14. Fitting results of four Gaussian components. Each panel represents different observation day. The dotted line represents observation spectrum. The solid line represents fitting results.

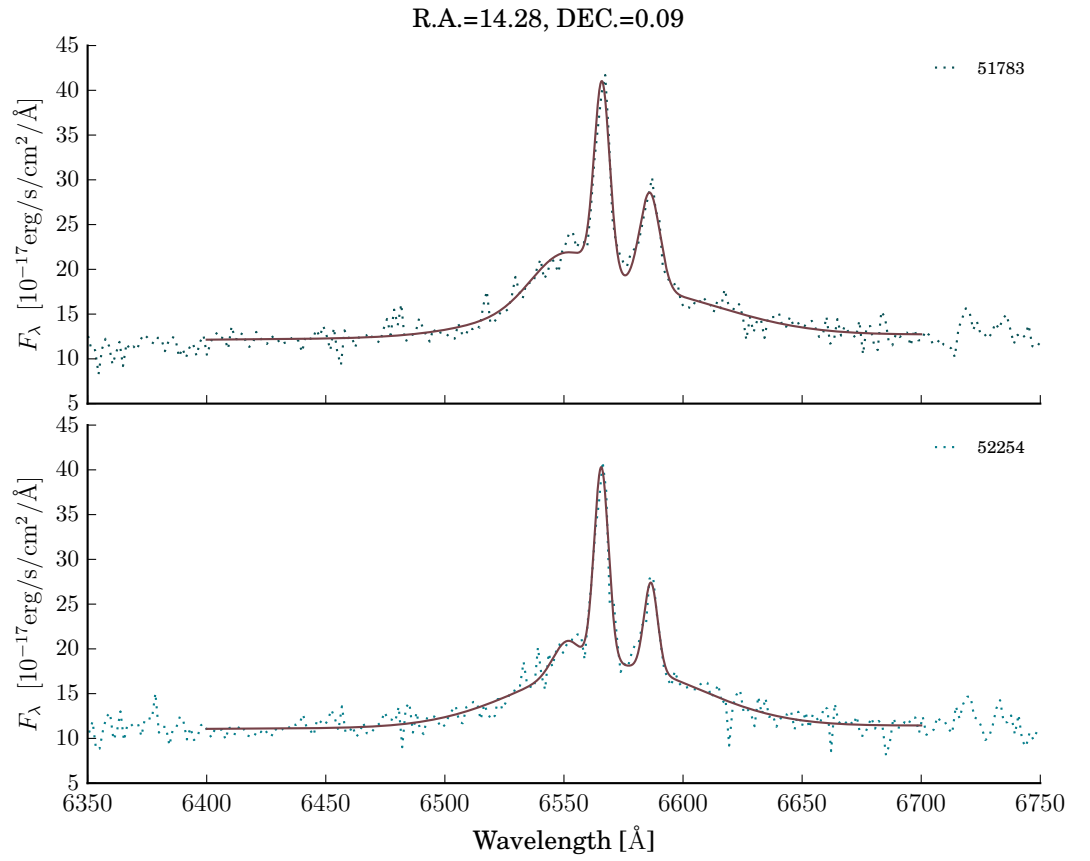


Figure A.15. Fitting results of four Gaussian components. Each panel represents different observation day. The dotted line represents observation spectrum. The solid line represents fitting results.

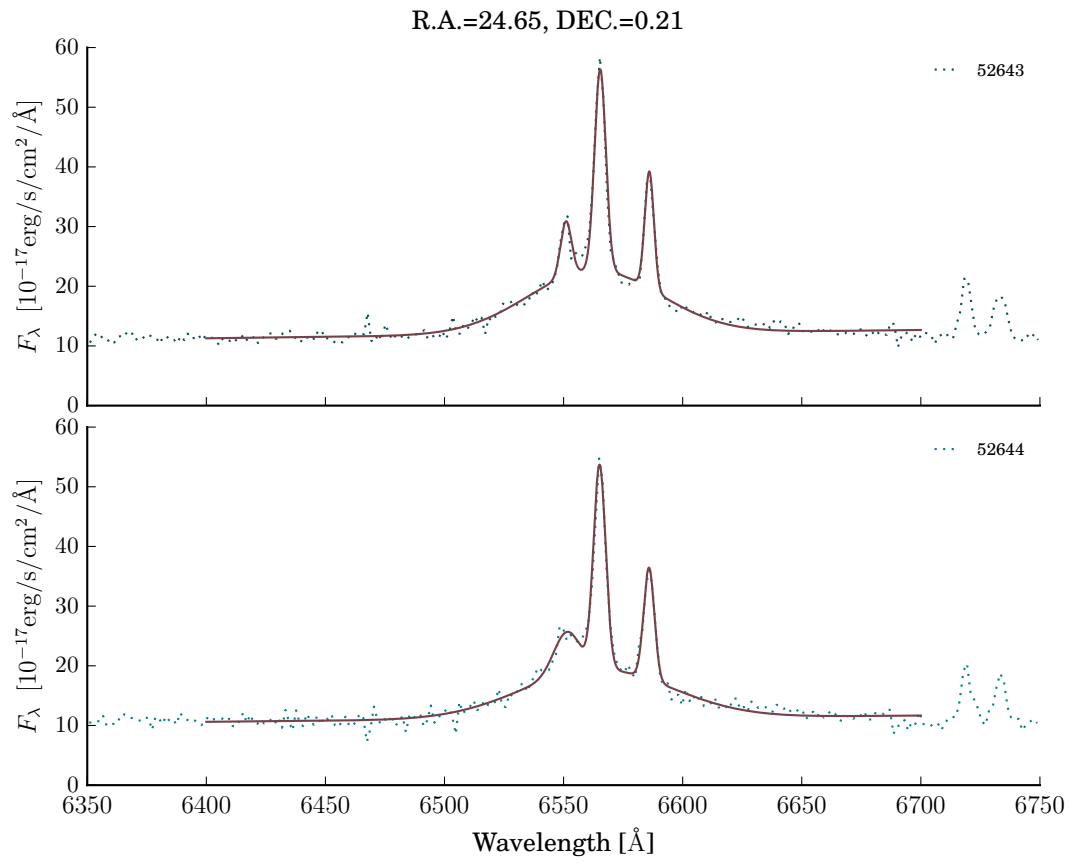


Figure A.16. Fitting results of four Gaussian components. Each panel represents different observation day. The dotted line represents observation spectrum. The solid line represents fitting results.

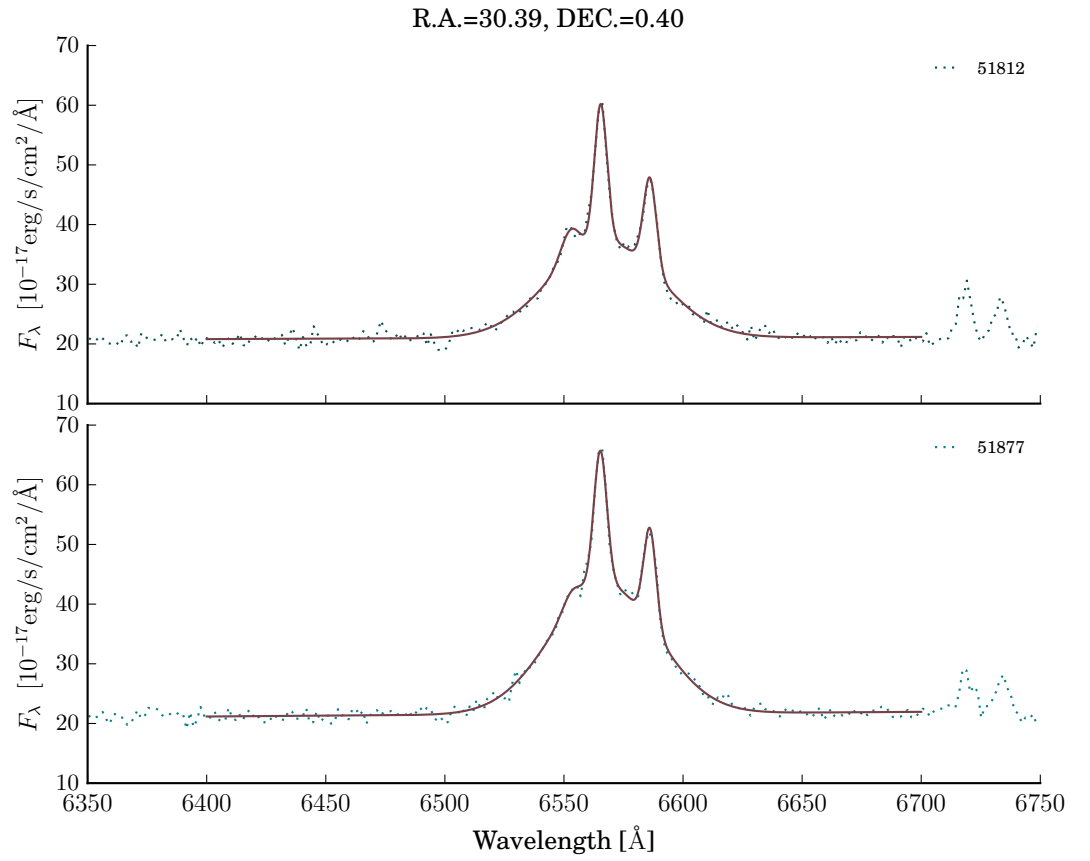


Figure A.17. Fitting results of four Gaussian components. Each panel represents different observation day. The dotted line represents observation spectrum. The solid line represents fitting results.

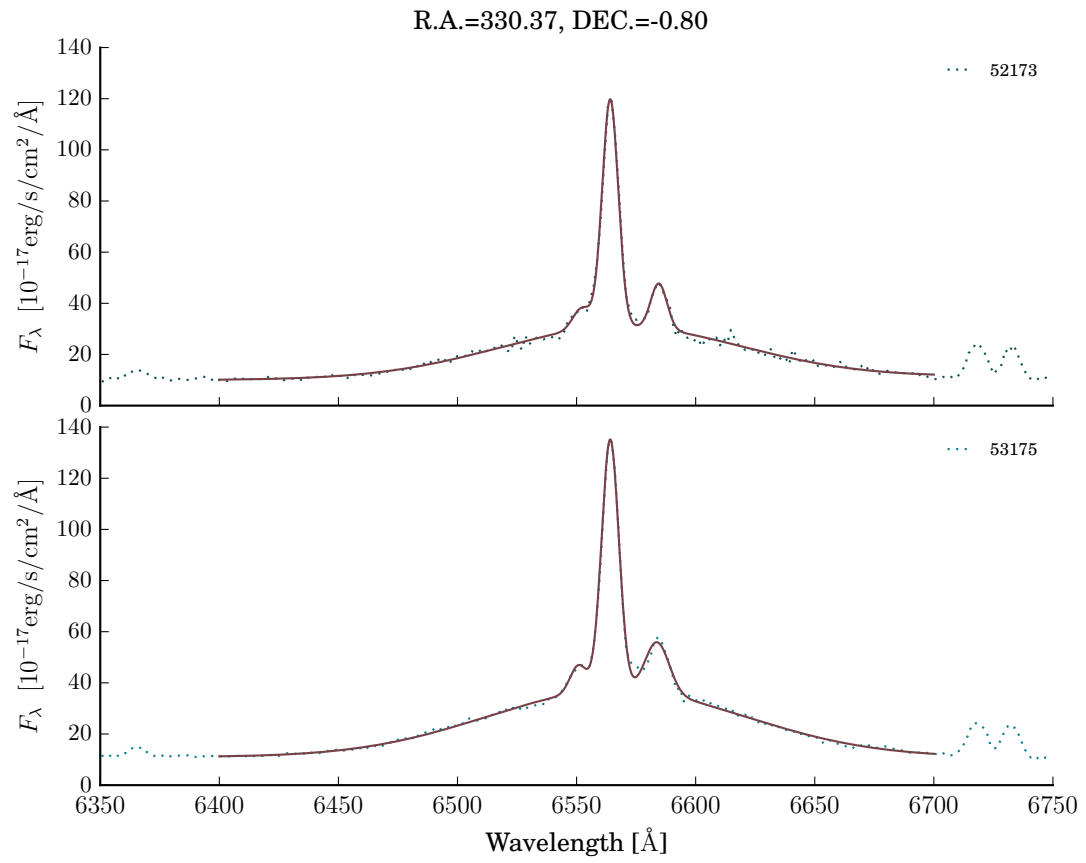


Figure A.18. Fitting results of four Gaussian components. Each panel represents different observation day. The dotted line represents observation spectrum. The solid line represents fitting results.

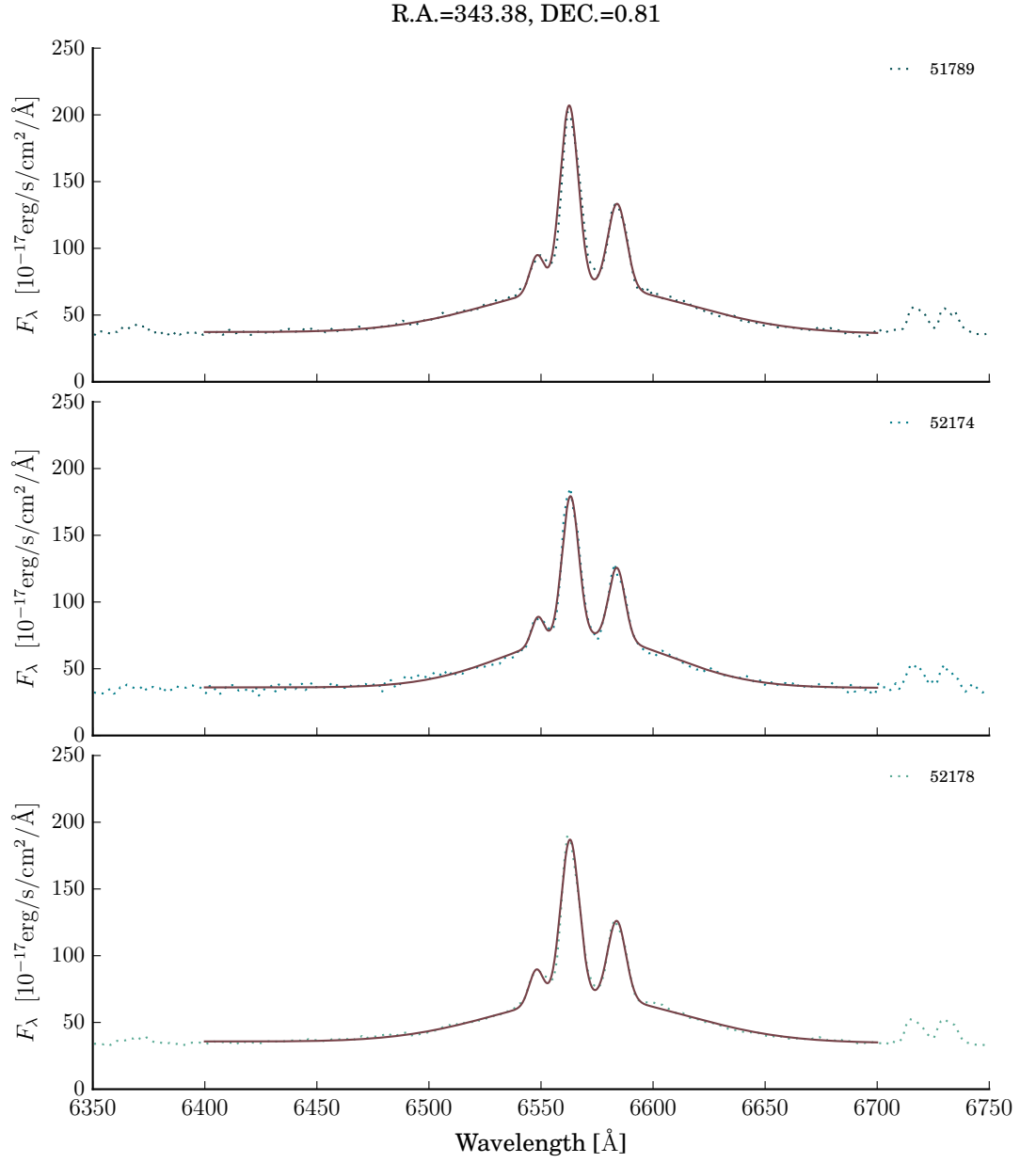


Figure A.19. Fitting results of four Gaussian components. Each panel represents different observation day. The dotted line represents observation spectrum. The solid line represents fitting results.

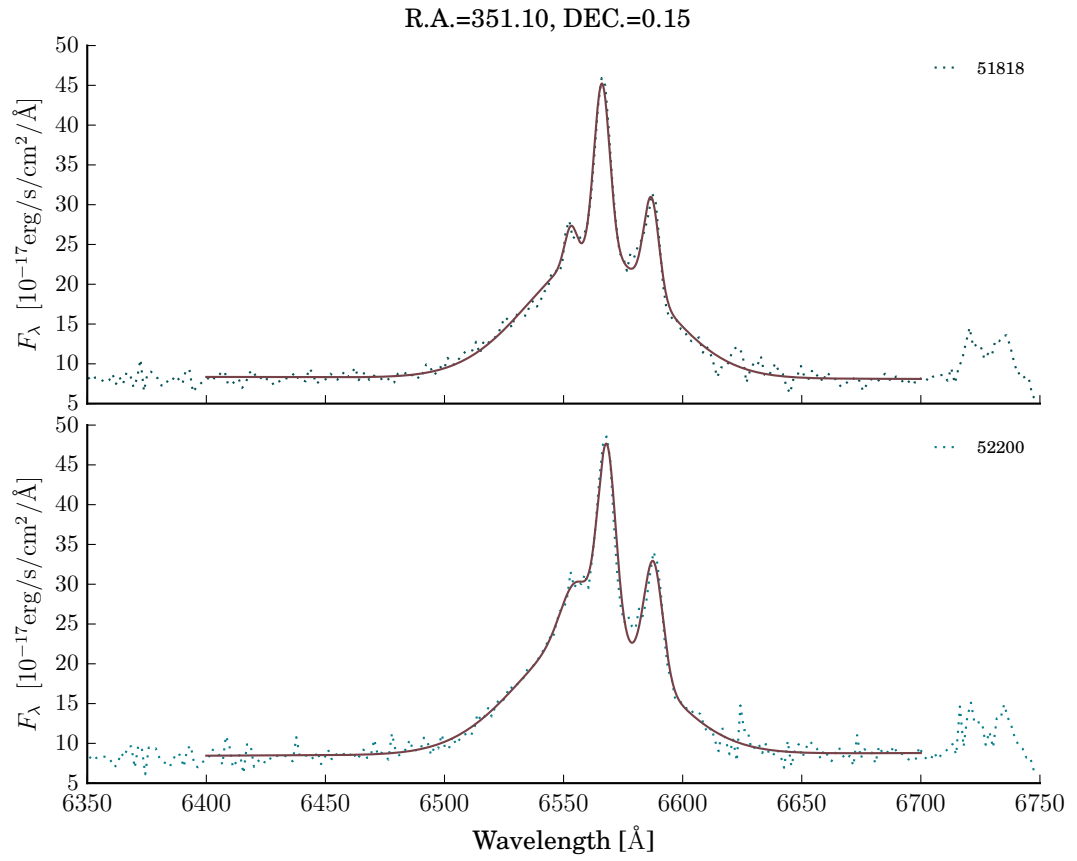


Figure A.20. Fitting results of four Gaussian components. Each panel represents different observation day. The dotted line represents observation spectrum. The solid line represents fitting results.

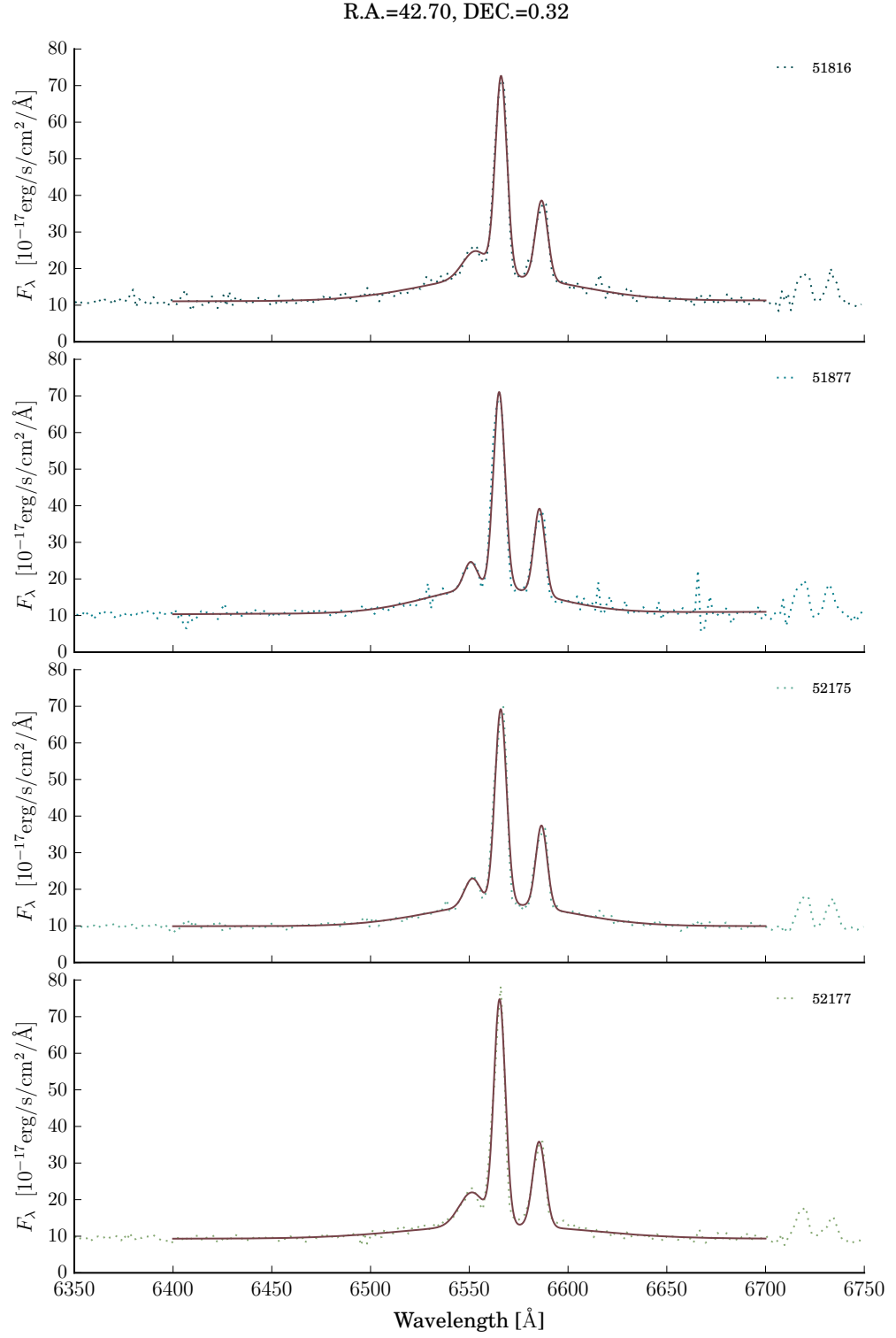


Figure A.21. Fitting results of four Gaussian components. Each panel represents different observation day. The dotted line represents observation spectrum. The solid line represents fitting results.

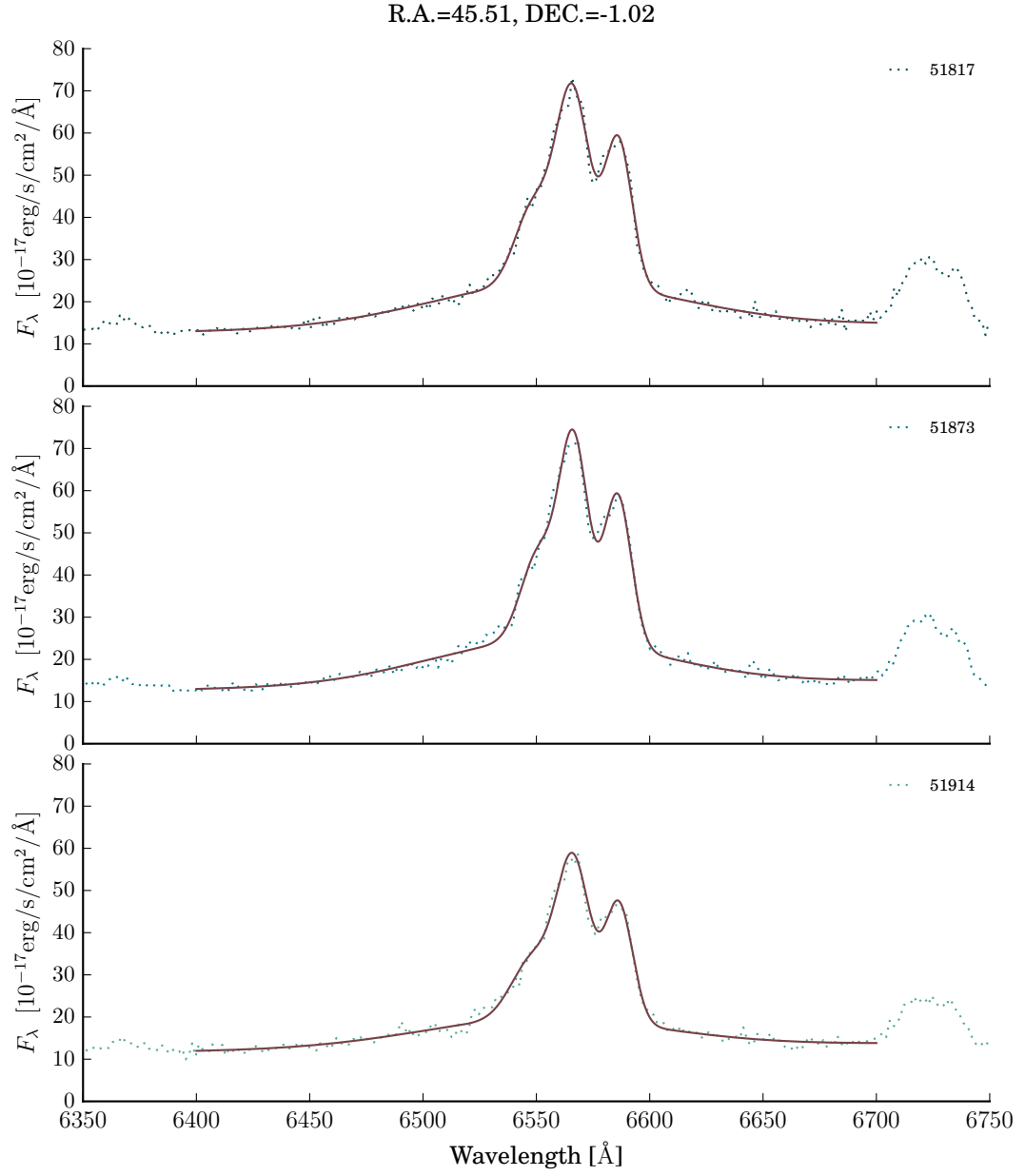


Figure A.22. Fitting results of four Gaussian components. Each panel represents different observation day. The dotted line represents observation spectrum. The solid line represents fitting results.

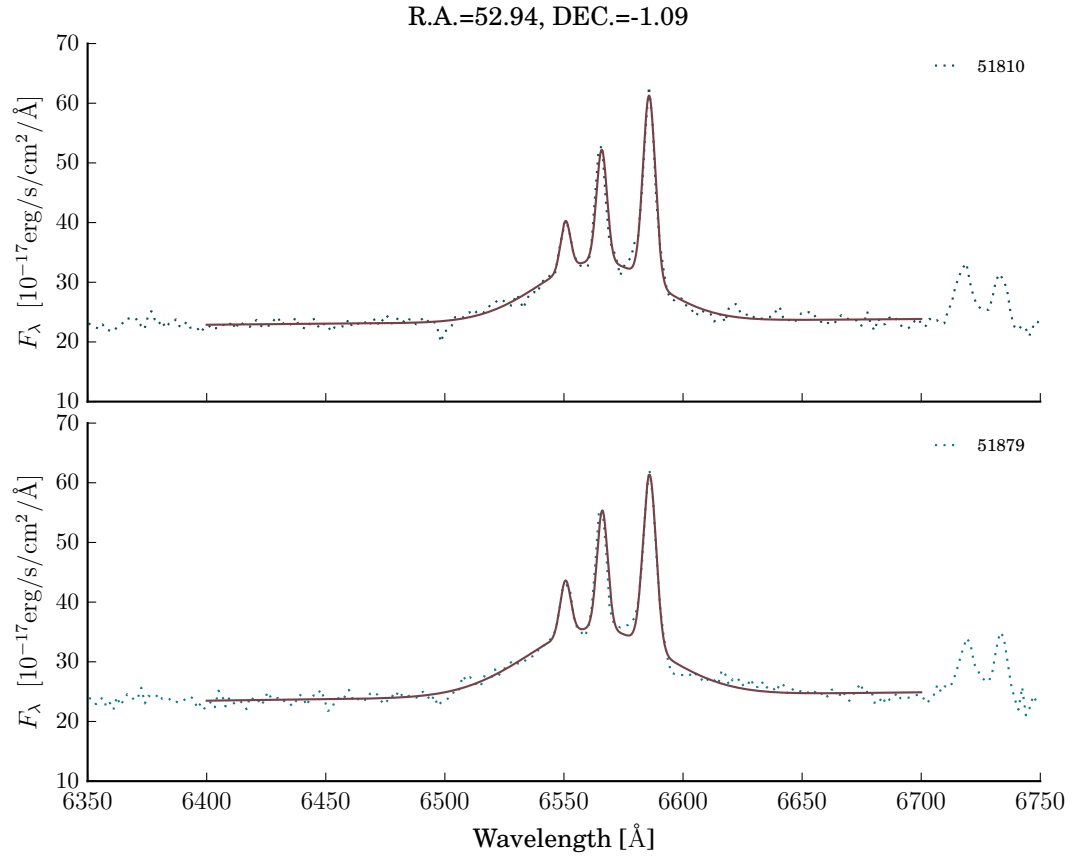


Figure A.23. Fitting results of four Gaussian components. Each panel represents different observation day. The dotted line represents observation spectrum. The solid line represents fitting results.

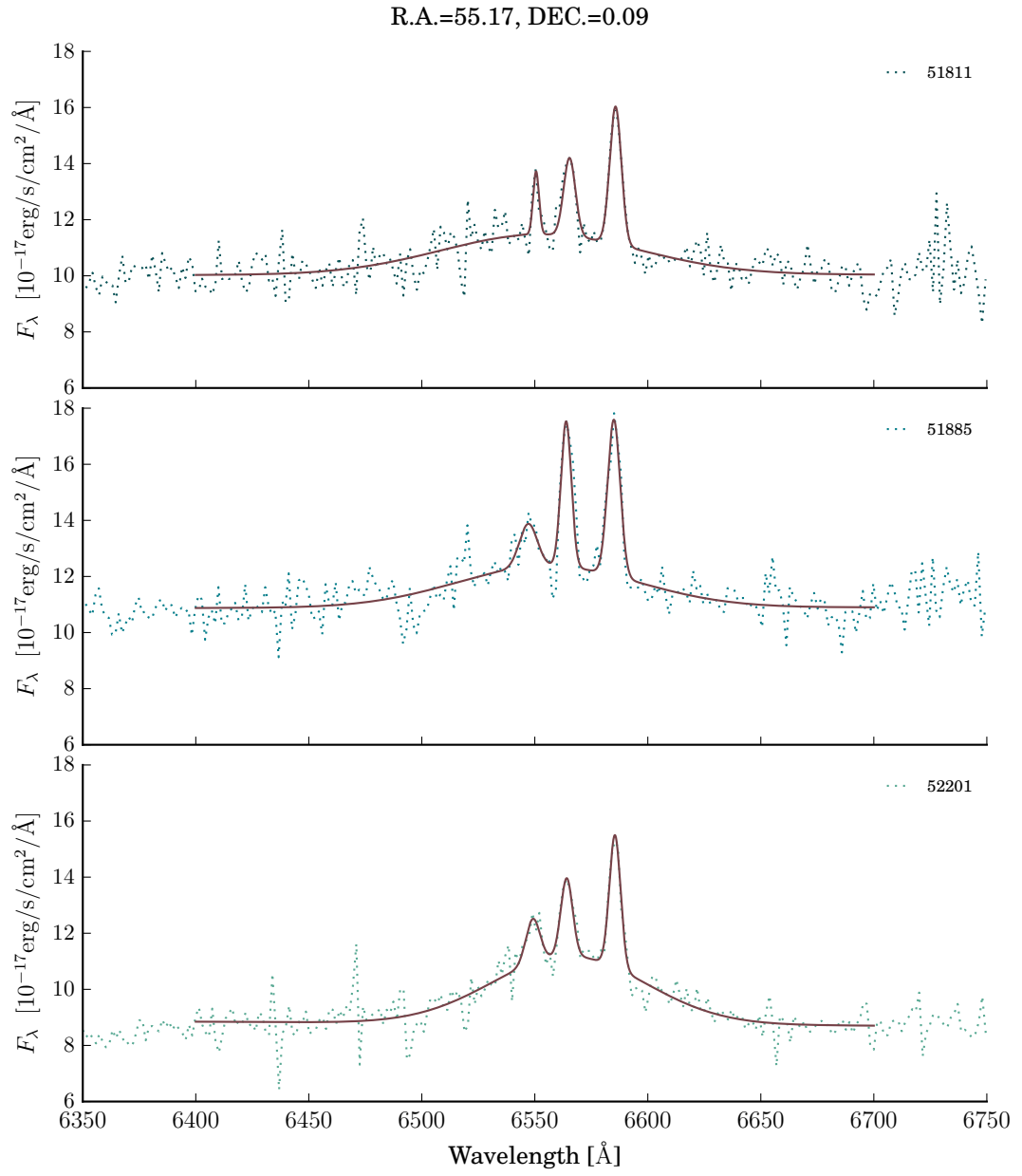


Figure A.24. Fitting results of four Gaussian components. Each panel represents different observation day. The dotted line represents observation spectrum. The solid line represents fitting results.

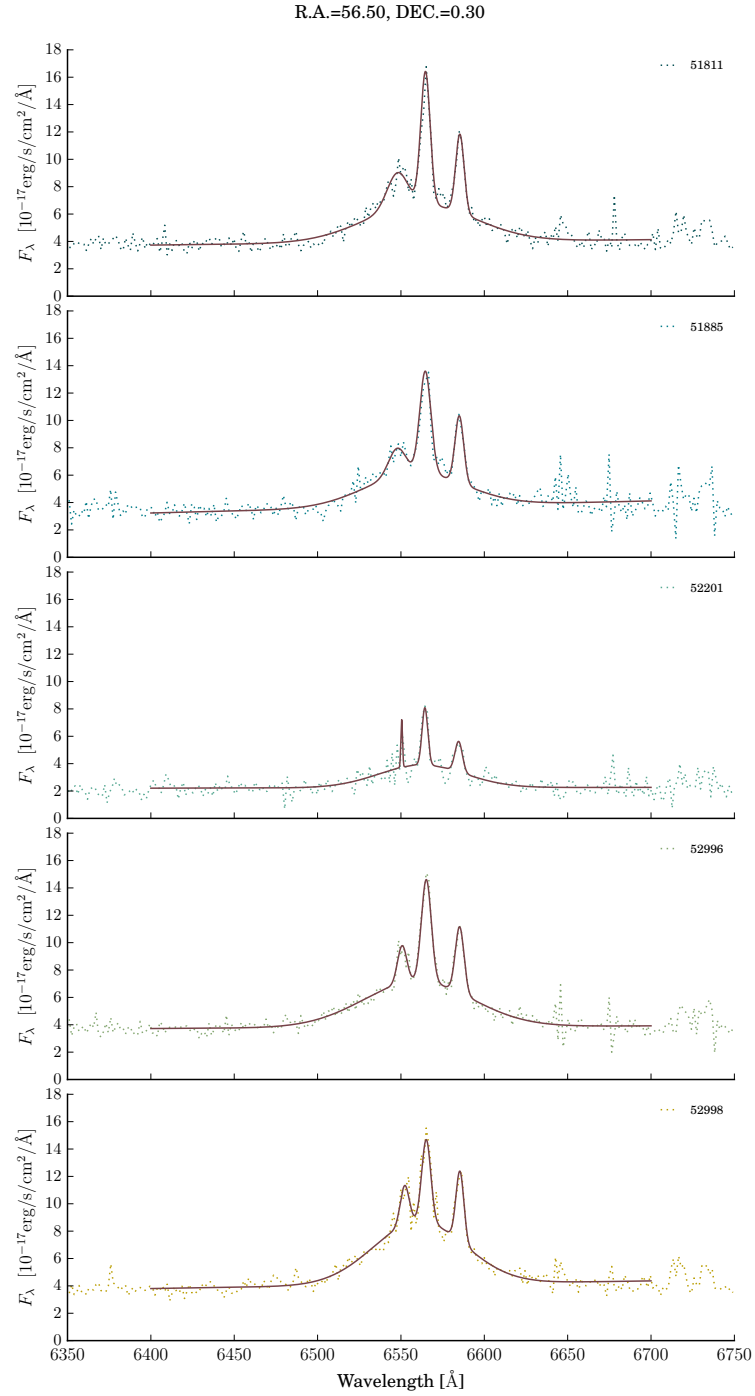


Figure A.25. Fitting results of four Gaussian components. Each panel represents different observation day. The dotted line represents observation spectrum. The solid line represents fitting results.

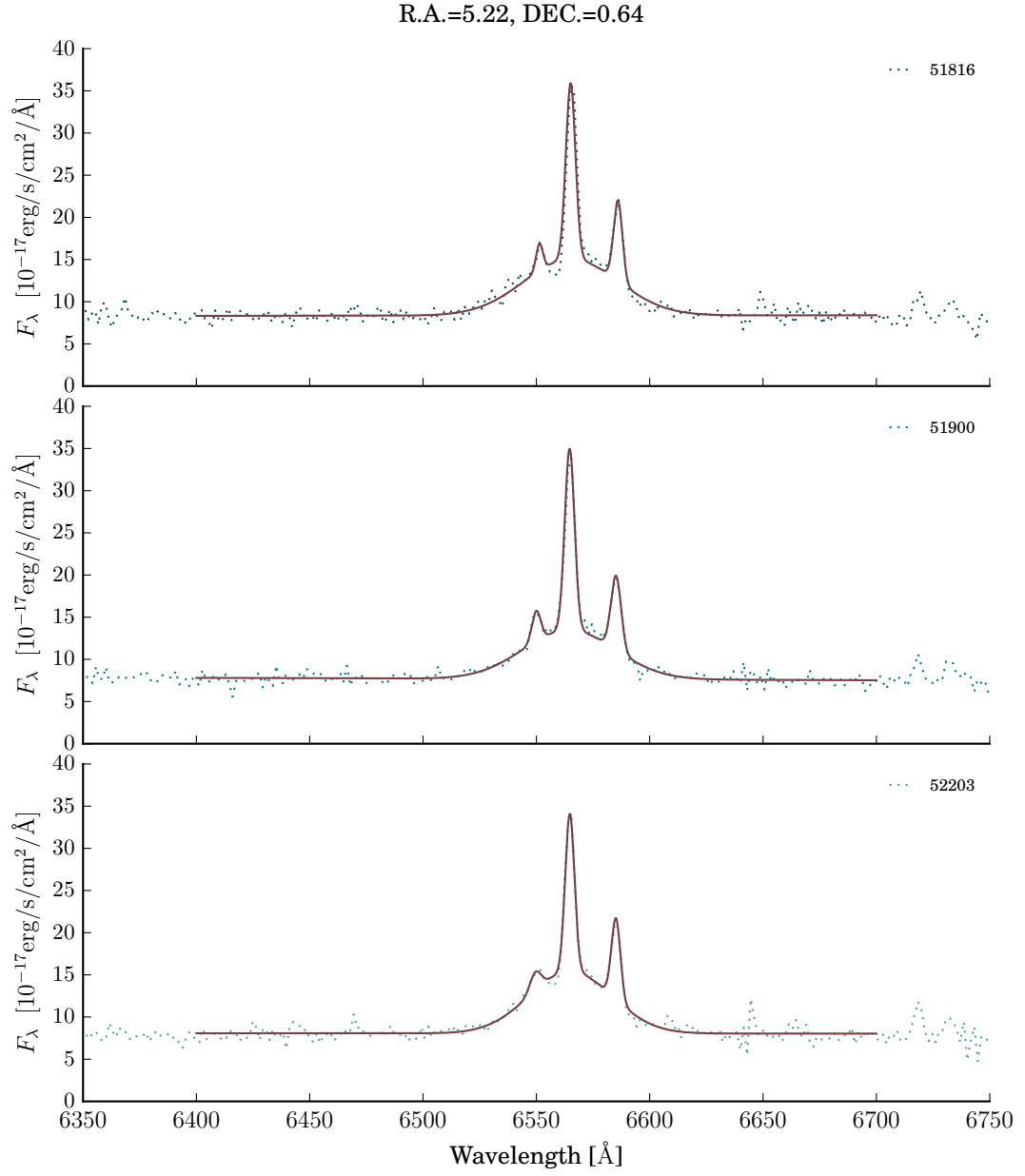


Figure A.26. Fitting results of four Gaussian components. Each panel represents different observation day. The dotted line represents observation spectrum. The solid line represents fitting results.

Appendix B

1D CNN model code (Pytorch)

```
import torch

class Net(torch.nn.Module):

    def __init__(self):
        super(Net, self).__init__()
        self.C1 = torch.nn.Conv1d(1, 10, 200)
        self.S2 = torch.nn.AvgPool1d(2, stride=2)
        self.C3 = torch.nn.Conv1d(10, 30, 100)
        self.S4 = torch.nn.AvgPool1d(5, stride=2)
        self.C5 = torch.nn.Conv1d(30, 36, 56)
        self.F6 = torch.nn.Linear(3384, 94)
        self.Out = torch.nn.Linear(94, 2)

    def forward(self, x):
        x = F.relu(self.C1(x))
        x = self.S2(x)
        x = F.relu(self.C3(x))
        x = self.S4(x)
        x = F.relu(self.C5(x))
        x = x.view(x.size(0), -1)
        x = F.relu(self.F6(x))
        x = self.Out(x)
        return x
```


Bibliography

- [1] AHN, C. P., ET AL. The Tenth Data Release of the Sloan Digital Sky Survey: First Spectroscopic Data from the SDSS-III Apache Point Observatory Galactic Evolution Experiment. *Astrophysical Journal Supplement*, **211** (2014), 17. [arXiv:1307.7735](#), [doi:10.1088/0067-0049/211/2/17](#).
- [2] ANNIS, J., ET AL. The Sloan Digital Sky Survey Coadd: 275 deg² of Deep Sloan Digital Sky Survey Imaging on Stripe 82. *The Astrophysical Journal*, **794** (2014), 120. [arXiv:1111.6619](#), [doi:10.1088/0004-637X/794/2/120](#).
- [3] ANTONUCCI, R. Unified models for active galactic nuclei and quasars. *Annual Review of Astronomy and Astrophysics*, **31** (1993), 473. [doi:10.1146/annurev.aa.31.090193.002353](#).
- [4] ANTONUCCI, R. R. J. AND MILLER, J. S. Spectropolarimetry and the nature of NGC 1068. *Astrophysical Journal*, **297** (1985), 621. [doi:10.1086/163559](#).
- [5] BALDWIN, J. A., PHILLIPS, M. M., AND TERLEVICH, R. Classification parameters for the emission-line spectra of extragalactic objects. *Publications of the Astronomical Society of the Pacific*, **93** (1981), 5. [doi:10.1086/130766](#).
- [6] CHEN, Y.-C. AND HWANG, C.-Y. Morphology of Seyfert galaxies. *Astrophysics and Space Science*, **362** (2017), 230. [arXiv:1707.08806](#), [doi:10.1007/s10509-017-3210-0](#).
- [7] CHEN, Y.-C. AND HWANG, C.-Y. Emission line luminosity distributions of Seyfert 2 galaxies. *Monthly Notices of the Royal Astronomical Society*, **485** (2019), 3402. [arXiv:1903.03581](#), [doi:10.1093/mnras/stz677](#).
- [8] FANAROFF, B. L. AND RILEY, J. M. The morphology of extragalactic radio sources of high and low luminosity. *Monthly Notices of the Royal Astronomical Society*, **167** (1974), 31P. [doi:10.1093/mnras/167.1.31P](#).
- [9] GIOMMI, P., PADOVANI, P., AND POLENTA, G. A simplified view of blazars: the γ -ray case. *Monthly Notices of the Royal Astronomical Society*, **431** (2013), 1914. [arXiv:1302.4331](#), [doi:10.1093/mnras/stt305](#).
- [10] GOODRICH, R. W. PA beta Measurements and Reddening in Seyfert 1.8 and 1.9 Galaxies. *Astrophysical Journal*, **355** (1990), 88. [doi:10.1086/168743](#).

- [11] GREENE, J. E. AND HO, L. C. A New Sample of Low-Mass Black Holes in Active Galaxies. *The Astrophysical Journal*, **670** (2007), 92. [arXiv:0707.2617](#), [doi:10.1086/522082](#).
- [12] HO, L. C. AND PENG, C. Y. Nuclear Luminosities and Radio Loudness of Seyfert Nuclei. *The Astrophysical Journal*, **555** (2001), 650. [arXiv:astro-ph/0102502](#), [doi:10.1086/321524](#).
- [13] IZOTOV, Y. I., THUAN, T. X., AND GUSEVA, N. G. Broad-Line Emission in Low-Metallicity Blue Compact Dwarf Galaxies: Evidence for Stellar Wind, Supernova, and Possible AGN Activity. *The Astrophysical Journal*, **671** (2007), 1297. [arXiv:0709.3643](#), [doi:10.1086/522923](#).
- [14] KAUFFMANN, G., ET AL. The host galaxies of active galactic nuclei. *Monthly Notices of the Royal Astronomical Society*, **346** (2003), 1055. [arXiv:astro-ph/0304239](#), [doi:10.1111/j.1365-2966.2003.07154.x](#).
- [15] KEWLEY, L. J., DOPITA, M. A., SUTHERLAND, R. S., HEISLER, C. A., AND TREVENA, J. Theoretical Modeling of Starburst Galaxies. *The Astrophysical Journal*, **556** (2001), 121. [arXiv:astro-ph/0106324](#), [doi:10.1086/321545](#).
- [16] KEWLEY, L. J., GROVES, B., KAUFFMANN, G., AND HECKMAN, T. The host galaxies and classification of active galactic nuclei. *Monthly Notices of the Royal Astronomical Society*, **372** (2006), 961. [arXiv:astro-ph/0605681](#), [doi:10.1111/j.1365-2966.2006.10859.x](#).
- [17] KHACHIKIAN, E. Y. AND WEEDMAN, D. W. An atlas of Seyfert galaxies. *Astrophysical Journal*, **192** (1974), 581. [doi:10.1086/153093](#).
- [18] KRIZHEVSKY, A., SUTSKEVER, I., AND HINTON, G. E. Imagenet classification with deep convolutional neural networks. *Commun. ACM*, **60** (2017), 84–90. Available from: <https://doi.org/10.1145/3065386>, [doi:10.1145/3065386](#).
- [19] LEUNG, H. W. AND BOVY, J. Deep learning of multi-element abundances from high-resolution spectroscopic data. *Monthly Notices of the Royal Astronomical Society*, **483** (2019), 3255. [arXiv:1808.04428](#), [doi:10.1093/mnras/sty3217](#).
- [20] LIU, H.-Y., LIU, W.-J., DONG, X.-B., ZHOU, H., WANG, T., LU, H., AND YUAN, W. A Comprehensive and Uniform Sample of Broad-line Active Galactic Nuclei from the SDSS DR7. *The Astrophysical Journal Supplement Series*, **243** (2019), 21. [arXiv:1906.05597](#), [doi:10.3847/1538-4365/ab298b](#).
- [21] MALKAN, M. A., JENSEN, L. D., RODRIGUEZ, D. R., SPINOGLIO, L., AND RUSH, B. Emission Line Properties of Seyfert Galaxies in the 12 μ m Sample. *The Astrophysical Journal*, **846** (2017), 102. [arXiv:1708.08563](#), [doi:10.3847/1538-4357/aa8302](#).
- [22] NETZER, H. AND LAOR, A. Dust in the Narrow-Line Region of Active Galactic Nuclei. *Astrophysical Journal Letters*, **404** (1993), L51. [doi:10.1086/186741](#).
- [23] OSTERBROCK, D. E. Spectrophotometry of Seyfert 1 galaxies. *Astrophysical Journal*, **215** (1977), 733. [doi:10.1086/155407](#).

- [24] OSTERBROCK, D. E. Seyfert galaxies with weak broad H alpha emission lines. *Astrophysical Journal*, **249** (1981), 462. [doi:10.1086/159306](https://doi.org/10.1086/159306).
- [25] OSTERBROCK, D. E. AND DAHARI, O. Spectra of Seyfert galaxies and Seyfert galaxy candidates. *Astrophysical Journal*, **273** (1983), 478. [doi:10.1086/161385](https://doi.org/10.1086/161385).
- [26] OSTERBROCK, D. E. AND KOSKI, A. T. NGC 4151 and Markarian 6 - two intermediate-type Seyfert galaxies. *Monthly Notices of the Royal Astronomical Society*, **176** (1976), 61P. [doi:10.1093/mnras/176.1.61P](https://doi.org/10.1093/mnras/176.1.61P).
- [27] PETERSON, B. M. Reverberation Mapping of Active Galactic Nuclei. *Publications of the Astronomical Society of the Pacific*, **105** (1993), 247. [doi:10.1086/133140](https://doi.org/10.1086/133140).
- [28] REES, M. J. Black Hole Models for Active Galactic Nuclei. *Annual Review of Astronomy and Astrophysics*, **22** (1984), 471. [doi:10.1146/annurev.aa.22.090184.002351](https://doi.org/10.1146/annurev.aa.22.090184.002351).
- [29] ROWAN-ROBINSON, M. On the unity of activity in galaxies. *Astrophysical Journal*, **213** (1977), 635. [doi:10.1086/155195](https://doi.org/10.1086/155195).
- [30] SCHAWINSKI, K., THOMAS, D., SARZI, M., MARASTON, C., KAVIRAJ, S., JOO, S.-J., YI, S. K., AND SILK, J. Observational evidence for AGN feedback in early-type galaxies. *Monthly Notices of the Royal Astronomical Society*, **382** (2007), 1415. [arXiv:0709.3015](https://arxiv.org/abs/0709.3015), [doi:10.1111/j.1365-2966.2007.12487.x](https://doi.org/10.1111/j.1365-2966.2007.12487.x).
- [31] SCHMIDT, M. 3C 273 : A Star-Like Object with Large Red-Shift. *Nature*, **197** (1963), 1040. [doi:10.1038/1971040a0](https://doi.org/10.1038/1971040a0).
- [32] SEYFERT, C. K. Nuclear Emission in Spiral Nebulae. *Astrophysical Journal*, **97** (1943), 28. [doi:10.1086/144488](https://doi.org/10.1086/144488).
- [33] SHUDER, J. M. Emission-line-continuum correlations in active galactic nuclei. *Astrophysical Journal*, **244** (1981), 12. [doi:10.1086/158678](https://doi.org/10.1086/158678).
- [34] SHUDER, J. M. AND OSTERBROCK, D. E. Empirical results from a study of active galactic nuclei. *Astrophysical Journal*, **250** (1981), 55. [doi:10.1086/159347](https://doi.org/10.1086/159347).
- [35] STERN, J. AND LAOR, A. Type 1 AGN at low z - II. The relative strength of narrow lines and the nature of intermediate type AGN. *Monthly Notices of the Royal Astronomical Society*, **426** (2012), 2703. [arXiv:1207.5543](https://arxiv.org/abs/1207.5543), [doi:10.1111/j.1365-2966.2012.21772.x](https://doi.org/10.1111/j.1365-2966.2012.21772.x).
- [36] STERN, J. AND LAOR, A. Type 1 AGN at low z- I. Emission properties. *Monthly Notices of the Royal Astronomical Society*, **423** (2012), 600. [arXiv:1203.3158](https://arxiv.org/abs/1203.3158), [doi:10.1111/j.1365-2966.2012.20901.x](https://doi.org/10.1111/j.1365-2966.2012.20901.x).
- [37] TRIPPE, M. L., CRENSHAW, D. M., DEO, R., AND DIETRICH, M. Long-Term Variability in the Optical Spectrum of the Seyfert Galaxy NGC 2992. *Monthly*

- Notices of the Royal Astronomical Society*, **135** (2008), 2048. [arXiv:0804.1381](#), [doi:10.1088/0004-6256/135/6/2048](#).
- [38] TRIPPE, M. L., CRENSHAW, D. M., DEO, R. P., DIETRICH, M., KRAEMER, S. B., RAFTER, S. E., AND TURNER, T. J. A Multi-wavelength Study of the Nature of Type 1.8/1.9 Seyfert Galaxies. *The Astrophysical Journal*, **725** (2010), 1749. [arXiv:1010.2750](#), [doi:10.1088/0004-637X/725/2/1749](#).
- [39] ULRICH, M. H., ET AL. Detailed observations of NGC 4151 with the IUE - III. Variability of the strong emission lines from 1978 February to 1980 May. *Monthly Notices of the Royal Astronomical Society*, **206** (1984), 221. [doi:10.1093/mnras/206.1.221](#).
- [40] URRY, C. M. AND PADOVANI, P. Unified Schemes for Radio-Loud Active Galactic Nuclei. *Publications of the Astronomical Society of the Pacific*, **107** (1995), 803. [arXiv:astro-ph/9506063](#), [doi:10.1086/133630](#).
- [41] VAONA, L., CIROI, S., DI MILLE, F., CRACCO, V., LA MURA, G., AND RAFANELLI, P. Spectral properties of the narrow-line region in Seyfert galaxies selected from the SDSS-DR7. *Monthly Notices of the Royal Astronomical Society*, **427** (2012), 1266. [arXiv:1210.5201](#), [doi:10.1111/j.1365-2966.2012.22060.x](#).
- [42] VÉRON-CETTY, M. P. AND VÉRON, P. A catalogue of quasars and active nuclei: 13th edition. *Astronomy and Astrophysics*, **518** (2010), A10. [doi:10.1051/0004-6361/201014188](#).
- [43] VOGES, W. The rosat all-sky survey. *Advances in Space Research*, **13** (1993), 391. Available from: <http://www.sciencedirect.com/science/article/pii/0273117793901474>, [doi:https://doi.org/10.1016/0273-1177\(93\)90147-4](#).
- [44] WANG, J. AND WEI, J. Y. Understanding the AGN-Host Connection in Partially Obscured Active Galactic Nuclei. I. The Nature of AGN+H II Composites. *The Astrophysical Journal*, **679** (2008), 86. [arXiv:0802.0548](#), [doi:10.1086/587048](#).
- [45] WHITTLE, M. Virial and Jet-induced Velocities in Seyfert Galaxies. I. A Compilation of Narrow Line Region and Host Galaxy Properties. *Astrophysical Journal Supplement*, **79** (1992), 49. [doi:10.1086/191644](#).
- [46] WINKLER, H. Variability studies of Seyfert galaxies - II. Spectroscopy. *Monthly Notices of the Royal Astronomical Society*, **257** (1992), 677. [doi:10.1093/mnras/257.4.677](#).
- [47] YEE, H. K. C. Optical continuum and emission-line luminosity of active galactic nuclei and quasars. *Astrophysical Journal*, **241** (1980), 894. [doi:10.1086/158403](#).
- [48] YORK, D. G., ET AL. The Sloan Digital Sky Survey: Technical Summary. *Astronomical Journal*, **120** (2000), 1579. [arXiv:astro-ph/0006396](#), [doi:10.1086/301513](#).

- [49] YU, P.-C. AND HWANG, C.-Y. IONIZATION MECHANISMS OF HBLR AND NON-HBLR SEYFERT 2 GALAXIES. *The Astronomical Journal*, **142** (2011), 14. Available from: <https://doi.org/10.1088/0004-6256/142/1/14>, [doi:10.1088/0004-6256/142/1/14](https://doi.org/10.1088/0004-6256/142/1/14).
- [50] ZHANG, K., SCHLEGEL, D. J., ANDREWS, B. H., COMPARAT, J., SCHÄFER, C., VAZQUEZ MATA, J. A., KNEIB, J.-P., AND YAN, R. Machine-learning Classifiers for Intermediate Redshift Emission-line Galaxies. *Astrophysical Journal*, **883** (2019), 63. [arXiv:1908.07046](https://arxiv.org/abs/1908.07046), [doi:10.3847/1538-4357/ab397e](https://doi.org/10.3847/1538-4357/ab397e).