technische universität
dortmund

The Data Mining Guide to the Galaxy

Active Galactic Nuclei in a Multi-Wavelength Context

Dissertation

zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften (Dr. rer. nat.)

vorgelegt von
M. Sc. Sabrina Einecke
geboren in Werne

2017

Lehrstuhl für Experimentelle Physik V
Fakultät für Physik
Technische Universität Dortmund

Erstgutachter:	Prof. Dr. Dr. Wolfgang Rhode
Zweitgutachter:	Prof. Dr. Carsten Westphal
Abgabedatum:	20. Oktober 2017

Abstract

Within the scope of this thesis, various topics from the field of multi-wavelength astronomy have been pursued, focusing on Active Galactic Nuclei (AGN).
The Large Area Telescope (LAT) aboard the *Fermi* satellite carried out the most sensitive all-sky survey to date in the high-energy gamma-ray regime from 100 MeV to 300 GeV. The corresponding point source catalogs of different generations contain a large number of sources, associated with a source type and a counterpart in a different wavelength. However, there are also numerous sources to which no counterparts were assigned, and sources for which the source type is uncertain.

The affiliation of source types and counterparts is subject to many studies. A list of AGN candidates from one of these studies, based on the second *Fermi* catalog and on machine learning methods, has been examined in this thesis. For this purpose, an analysis of data from the X-ray telescope *Swift*-XRT has been carried out and the presence of X-ray sources in the uncertainty region of the AGN candidates has been evaluated. A fraction of 74% is compatible with the results of another study and suggests that not all AGNs emit X-rays. Moreover, the AGN candidates have been compared with counterpart candidates from other studies at different wavelengths. During the preparation of this thesis, the third *Fermi* catalog was published, which included, among others, associations of source types for some of the considered AGN candidates. This enabled a validation of the predictions, showing on the one hand the quality of machine learning methods in this field of application, and on the other hand confirming the estimated false association rate. Besides, it has been shown that the additional demand for the presence of a counterpart reduces the false association rate by more than half. This proves the importance of multi-wavelength data.

Based on these conclusions, a new method has been developed to find AGN candidates and their corresponding counterparts for unassociated *Fermi* sources. This method is based on multi-wavelength data and machine learning methods. A radio, infrared and X-ray catalog have been combined with the *Fermi* catalog, and a model has been created and validated from each of these combinations. From these models, two candidate lists have been extracted: A list of medium and one of high-confidence. The medium-confidence candidates of the different wavelengths have been joined so that a high-confidence candidate list was created, which additionally contained counterparts of different wavelengths. For the most promising candidate, a literature search has been carried out, which revealed a BL Lac object as source type. This type of object is a blazar – a subtype of AGNs. Thus, the prediction has been confirmed. In addition, the measurements of this candidate in the different wavelengths have been set in context with the blazar sequence, which suggested a BL Lac object as well, but also allowed the estimation of a redshift.

Due to the success of the newly developed method, it has been extended to enable the search for blazar candidates – or rather candidates for BL Lac objects and flat spectrum radio quasars – and their counterparts. The corresponding models are not only capable to associate the type and to link a counterpart with the unassociated sources, but also to determine the type of sources of uncertain blazar type. Once more, the resulting candidates exhibited high precisions. The combination of the models of all mentioned wavelengths produced several candidates of BL Lac objects, but none for flat-spectrum radio quasars.

For the BL Lac object with the highest predictive value, an analysis of observations with the MAGIC telescopes in the very-high-energy regime has been conducted. Due to the low flux of the source and a redshift of $z = 0.17$, the source has not been detected significantly. Nontheless, flux upper limits have been determined, suggesting intrinsic absorption of the source in addition to the absorption by the extragalactic background light. The corresponding spectral energy distribution shows a two-hump structure, which is typical for BL Lac objects. The low variability in the high-energy regime is typical for this type of object as well.

In the further course of this thesis, methods of machine learning within the analysis chain of the MAGIC experiment have been examined. Classifiers are used to separate gamma rays and hadrons, and regressors to reconstruct the energy of the incident primary. Until now, however, lookup tables have been used by default for the energy reconstruction of stereoscopic data. Therefore, a regression for the energy reconstruction has been set up, optimized, evaluated and compared with the standard method. The deviation from the true energy, as well as the resolution improved, especially in the high-energy regime. In addition, different methods to evaluate the performance have been assessed, and according to that, a more robust method has been proposed. The method for the separation of gamma rays and hadrons has been optimized and evaluated as well. Both the separation and the energy reconstruction have been verified by observations of the Crab Nebula and it has been shown that the significance increases with the optimized methods. Particularly in case of the energy reconstruction, an improvement in the high-energy regime has been achieved. In this regime, the spectral energy distribution is now more stable and can be reconstructed to higher energies.

Kurzfassung

Im Rahmen dieser Arbeit wurden verschiedene Themen aus dem Bereich der Multiwellenlängenastronomie verfolgt, wobei der Fokus auf Aktive Galaktische Kerne (AGN) gelegt wurde.
Das Large Area Telescope (LAT) an Bord des *Fermi*-Satelliten hat die bisher sensitivste Himmelsdurchmusterung im Bereich der hochenergetischen Gammastrahlung von 100 MeV bis 300 GeV vorgenommen. Die dazugehörigen Punktquellenkataloge verschiedener Generationen enthalten eine Vielzahl von Quellen, denen ein Quelltyp und ein Gegenstück in einer anderen Wellenlänge zugeordnet werden konnte. Allerdings sind auch zahlreiche Quellen beinhaltet, denen keine Gegenstücke zugeordnet werden konnten, und Quellen, bei denen der Quelltyp unsicher ist.

Die Zuordnung von Quelltypen und Gegenstücken ist Thema vieler Studien. Eine Liste von AGN-Kandidaten von einer dieser Studien, basierend auf dem zweiten *Fermi*-Katalog und Methoden des maschinellen Lernens, wurde in dieser Arbeit genauer untersucht. Dazu wurde eine Analyse von Daten des Röntgenteleskops *Swift*-XRT durchgeführt und das Vorhandensein von Röntgenquellen im Unsicherheitsbereich der AGN-Kandidaten evaluiert. Ein Anteil von 74% ist kompatibel mit den Ergebnissen aus einer anderen Studie und lässt vermuten, dass nicht alle AGNs Röntgenstrahlung emittieren. Außerdem wurden die AGN-Kandidaten mit Gegenstück-Kandidaten aus anderen Arbeiten verschiedener Wellenlängen verglichen. Während der Erstellung dieser Arbeit wurde der dritte *Fermi*-Katalog veröffentlicht, der unter anderem Assoziationen von Quelltypen für einige der betrachteten AGN-Kandidaten beinhaltete. Dies ermöglichte eine Validierung der Vorhersagen, wodurch zum einen die Güte von Methoden des maschinellen Lernens in diesem Anwendungsgebiet gezeigt wurde und zum anderen die abgeschätzte Falsch-Assoziationsrate bestätigt wurde. Zudem zeigte sich, dass durch die zusätzliche Forderung eines Vorhandenseins eines Gegenstücks die Falsch-Assoziationsrate um mehr als die Hälfte gesenkt werden kann.

Ausgehend von diesen Schlüssen wurde eine neue Methode entwickelt, um AGN-Kandidaten und Gegenstücke zu den *Fermi*-Quellen ohne zugeordneten Typ zu finden. Diese Methode basiert auf Multiwellenlängendaten und Methoden des maschinellen Lernens. Ein Radio-, ein Infrarot- und ein Röntgen-Katalog wurden jeweils mit dem *Fermi*-Katalog kombiniert und aus jeder dieser Kombinationen wurde ein Modell erstellt und validiert. Aus diesen Modellen wurden jeweils 2 Kandidatenlisten extrahiert: Eine Liste mit einer mittleren und eine mit einer hohen Genauigkeit. Die Kandidaten mit den mittleren Genauigkeiten der verschiedenen Wellenlängen wurden vereinigt, so dass eine Kandidatenliste von hoher Genauigkeit entstand, die zudem Gegenstücke verschiedener Wellenlängen enthält. Für den vielversprechendsten Kandidaten wurde eine Literaturrecherche durchgeführt, die als Quelltyp einen BL-Lac-Objekt hervorbrachte. Dieser Typ zählt zu den Blazaren, ein Untertyp von AGNs. Somit konnte die Vorhersage bestätigt werden. Zudem wurden für diesen Kandidaten die Messungen in den verschiedenen Wellenlängen in Zusammenhang mit der Blazar-Sequenz gesetzt, wodurch ebenfalls ein BL-Lac-Objekt nahegelegt wurde, aber auch eine Rotverschiebung abgeschätzt werden konnte.

Aufgrund des Gelingens der neu entwickelten Methode wurde diese erweitert, so dass Blazar-Kandidaten – genauer gesagt Kandidaten für BL-Lac-Objekte und Flachspektrum-Radioquasare – und deren Gegenstücke gesucht werden konnten. Die zugehörigen Modelle eignen sich nicht nur dazu, den Typ und die Gegenstücke von Quellen ohne zugeordneten Typ zu finden, sondern auch um den Typ von Quellen mit Gegenstück, aber ohne sicheren Quelltyp, zu bestimmen. Die resultierenden Kandidaten wiesen erneut eine hohe Genauigkeit auf. Aus der Vereinigung der Modelle aller aufgeführten Wellenlängen gingen mehrere Kandidaten von BL-Lac-Objekten hervor, aber keine für Flachspektrum-Radioquasare.

Von dem BL-Lac-Objekt mit dem höchsten Vorhersagewert wurde eine Analyse von Observationen mit den MAGIC-Teleskopen im sehr hochenergetischen Bereich durchgeführt. Aufgrund des geringen Flusses der Quelle und einer Rotverschiebung von $z = 0.17$ konnte die Quelle jedoch nicht signifikant detektiert werden. Trotzdessen konnten Fluss-Obergrenzen bestimmt werden, die vermuten lassen, dass neben der Absorption durch das extragalaktische Hintergrundlicht auch intrinsische Absorptionseffekte der Quelle eine Rolle spielen. Die zugehörige spektrale Energieverteilung zeigte eine für BL-Lac-Objekte typische Zwei-Höcker-Struktur. Ebenso die geringe Variabilität im hochenergetischen Bereich ist typisch für diese Art von Objekten.

Im weiteren Verlauf der Arbeit wurden Methoden des maschinellen Lernens innerhalb der Analysekette des MAGIC-Experiments begutachtet. Klassifizierer lassen sich unter anderem zur Separation von Gammastrahlung und Hadronen verwenden und Regressoren zur Rekonstruktion der Energie der einfallenden Strahlung. Bislang wurden standardmäßig allerdings Lookup-Tabellen für die Energierekonstruktion von stereoskopische Daten verwendet. Daher wurde im Rahmen dieser Arbeit ein Regressionsverfahren für die Energierekonstruktion aufgesetzt, optimiert, evaluiert und mit der standardmäßigen Methode verglichen. Sowohl die Abweichung vom wahren Wert, als auch die Auflösung konnten so verbessert werden, insbesondere im hochenergetischen Bereich. Zudem wurden verschiedene Vorgehensweisen zur Bewertung der Güte begutachtet, wodurch eine robustere Methode vorgeschlagen werden konnte, die unter anderem keine Annahme einer speziellen Verteilung erfordert. Die Methode zur Separation von Gammastrahlung und Hadronen wurde ebenfalls optimiert und evaluiert. Sowohl die Separation als auch die Energierekonstruktion wurden anhand von Observationen des Krebsnebels verifiziert und es wurde gezeigt, dass durch die Optimierung die Signifikanz gesteigert werden konnte. Insbesondere bei der Energierekonstruktion zeigte sich die Verbesserung im Hochenergiebereich. In diesem Bereich ist die spektrale Energieverteilung nun viel stabiler und kann zu höheren Energien rekonstruiert werden.

Contents

1 Introduction

Space is big. Really big.
You just won't believe how vastly, hugely, mind-bogglingly big it is.

— Douglas Adams, *The Hitchhiker's Guide to the Galaxy*

How big is big? Hundreds of thousands of years are already necessary for light to traverse our Galaxy, comprising hundreds of billions of stars. The Universe is even bigger with 2 trillion of similar galaxies accessible to our observations. The size of the observable universe has been determined by the Planck space mission to be 13.8 billion light years, deduced from the most precise map of the cosmic microwave background to date. This is already vastly big, but the Universe beyond the observable one is once again by a multitude bigger. Arthur Dent – the main character in *The Hitchhiker's Guide to the Galaxy* – must have felt rather lost in this mind-bogglingly big space.

Such a big space comprises naturally also a huge amount of information. Multiple instruments collected data of several messengers, produced by various astrophysical objects and physical processes, summarized in a vast amount of catalogs. A large fraction of these objects beyond our Galaxy is made up of Active Galactic Nuclei, which provide insights into the cosmological evolution and the extragalactic background light. To probe this extragalactic background light, and thus, the star formation history, a sample of Active Galactic Nuclei in the very-high-energy regime at different redshifts can be exploited. Unfortunately, the available Active Galactic Nucleus population in the very-high-energy regime is very limited. Especially, if the sample is subdivided in different redshifts and in subclasses of Active Galactic Nuclei, the resulting samples are not representative. Correspondingly, it is of tremendous importance to enlarge these samples, either by more sensitive instruments, or by new analysis techniques. The most frequent subclass of Active Galactic Nuclei detectable in the very-high-energy regime is the blazar class, which is further subdivided in flat-spectrum radio quasars and BL Lac objects. Questions still to be answered concern the acceleration and emission processes of these sources, including which processes are the same and which are different for BL Lac objects and flat-spectrum radio quasars. Active Galactic Nuclei are also considered as potential sources of ultra-high-energy cosmic rays and neutrinos. Due to the small amount of detected astrophysical neutrinos, it is as of yet not possible to discover point sources, instead, so-called stacking analyses are often performed. In these analyses, multiple source of a particular type are combined in one analysis. Once more, it is important to have representative populations of diverse source classes to enable these stacking analyses. For transient events, such as extremely-high-energy neutrinos, gravitational waves, or flaring Active Galactic Nuclei, alerts are typically issued to enable follow-up observations by other instruments. But due to possible localization inaccuracies of the issuing instrument, the follow-up observations are often complicated and inefficient. The knowledge of the class and the position of the sources in question helps to develop the pointing strategy for the observations, and to make the most of the available observation time. Generally speaking, a more comprehensive sample of sources, including information regarding their class affiliation and position, provides a vast amount of possibilities for further investigations and physical insights.

For the determination of the class affiliation and the position of a gamma-ray emitting object, the lack of an associated counterpart often constitutes a major obstacle. The knowledge of these counterparts provides several advantages, such as accurate positions, since the measurements thereof are mostly conducted with instruments featuring a more accurate source localization than gamma-ray instruments. Thus, further counterparts can be associated and dedicated observations can be scheduled, for instance optical spectroscopy to determine the redshift and the class affiliation. Moreover, these counterparts contain information concerning the spectral energy distribution of the source, supporting the affiliation of the class.

This thesis deals with the search for Active Galactic Nuclei and blazars, together with their corresponding counterparts, using multi-wavelength data and machine learning methods. Moreover, it discusses one of the found BL Lac objects in a multi-wavelength context, with emphasis on the analysis of observations in the very-high-energy regime.

For these purposes, the thesis is structured as follows:

Chapter 2 gives an overview of astroparticle physics, focusing on Active Galactic Nuclei, gamma-ray astronomy, and the instruments to be considered in the following chapters.

Chapter 3 explains typical machine learning algorithms to generate classification and regression models, and introduces methods to validate them, establishing basic prerequisites for all further work.

Chapter 4 describes a follow-up study of Active Galactic Nucleus candidates, which evaluates counterpart associations based on various wavelengths. Therefore, the analysis of X-ray measurements and the connection of multiple wavebands are illustrated.

Chapter 5 presents several source catalogs, and explains the construction of samples thereof, important for the search of Active Galactic Nuclei and blazars in the subsequent chapters.

Chapter 6 illustrates a search for Active Galactic Nucleus candidates and their counterparts in other wavelengths among the unassociated gamma-ray sources in the third *Fermi*-LAT source catalog using machine learning techniques.

Chapter 7 extends the previous chapter to a search for blazar candidates and their counterparts among the unassociated gamma-ray sources and the blazar candidates of uncertain type.

Chapter 8 presents the Active Galactic Nucleus 3FGL J2346.7+0705 in a multi-wavelength context. The analysis of observations with the MAGIC telescopes is described, and the results are discussed together with the results derived by other experiments in various wavelengths.

Chapter 9 revises the Random Forest methods, used within the analysis chain of the MAGIC experiment. The optimization of the classification of gamma-rays and hadrons, and the setting up of the regression of the energy of the primary incident are presented.

2 Astroparticle Physics at a Glance

All you really need to know for the moment
is that the Universe is a lot more complicated than you might think,
even if you start from a position of thinking it's pretty damn complicated in the first place.

— Douglas Adams, *The Hitchhiker's Guide to the Galaxy*

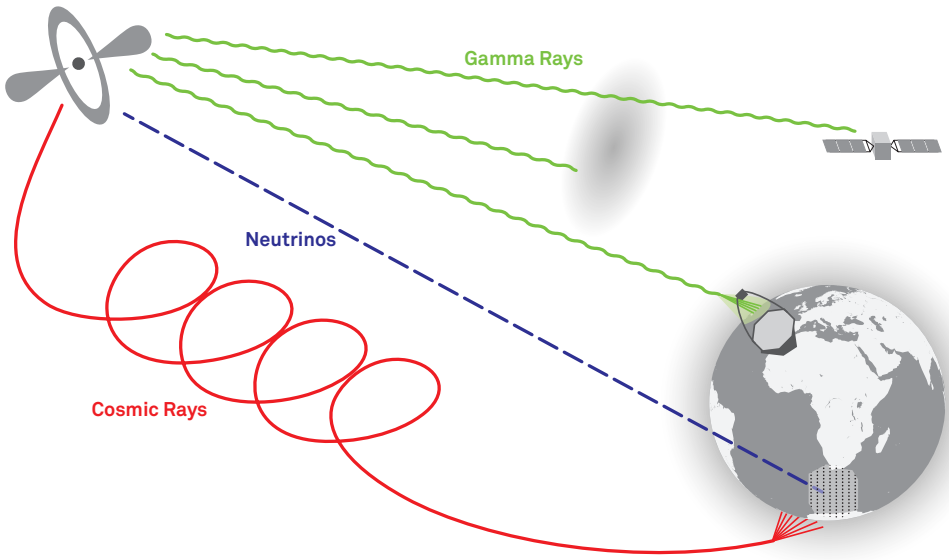

Figure 2.1: Astroparticle physics at a glance. Specific astronomical objects emit different messengers, such as neutrinos, cosmic and gamma rays, propagating through the universe. Depending on the type of messenger, they might interact with magnetic fields, interstellar clouds, the Earth's atmosphere or the Earth itself, and they can be detected with different instruments.

Astroparticle physics comprises multiple research fields, such as astronomy, particle physics, and cosmology, making it a very versatile research field. Figure 2.1 gives an illustrating overview of this field. To obtain information about astrophysical objects, such as Active Galactic Nuclei (AGN) or supernova remnants, three different types of messengers are available: Neutrinos, cosmic and gamma rays, featuring diverse – advantageous and disadvantageous – properties. Charged cosmic rays are deflected by intergalactic magnetic fields, complicating or even preventing to trace them back to their origin. When they strike the Earth's atmosphere, they interact with the atmosphere's molecules, inducing extensive air showers. The particles of these showers can be measured with ground-based detectors, such as the Pierre Auger Observatory (Aab et al., 2015) and the Telescope Array (Kawai et al., 2008). In contrast, gamma rays and neutrinos are uncharged, and correspondingly not deflected by magnetic fields. However, gamma rays can be absorbed, for instance by the extragalactic background light.

Gamma rays can be measured with satellite-based detectors, such as the *Fermi*-LAT (Atwood et al., 2009), or with ground-based detectors. Since they interact with the Earth's atmosphere and induce extensive air showers as well, the shower's particles can be measured with water Cherenkov detectors, such as HAWC (Abeysekara et al., 2013), or the Cherenkov light the particles emit can be detected with imaging air Cherenkov detectors, such as MAGIC (Aleksić et al., 2016a), VERITAS (Park et al., 2015), H.E.S.S. (Bonnefoy et al., 2017) or CTA (Acharya et al., 2017). Due to their small cross section, neutrinos are not absorbed when traveling the universe. While advantageous on the one hand, it is difficult to detect them on the other hand, and large detector volumes are necessary. Instruments like IceCube (Walter et al., 2007) and ANTARES (Ageron et al., 2011) utilize water and ice, respectively, as detector medium, to measure the Cherenkov light of secondary particles the neutrino induced.

2.1 Active Galactic Nuclei

Active Galactic Nuclei (AGN) are astrophysical objects, emitting radiation over the complete electromagnetic spectrum. A simplified schematic is presented in Figure 2.2. A supermassive black hole (SMBH) in the center of a galaxy is accreting matter from its surroundings, the accretion disc, and the main power source of an AGN. The mass of such an SMBH is in the range of $10^5 - 10^{10}$ $M_\odot$ (Beckmann and Shrader, 2012). Different processes in the accretion disc transform energy into thermal radiation, and the fast moving gas clouds above the disc are heated up – the broad line region. This region produces broad emission lines in the optical and ultraviolet range. A dust torus surrounds the accretion disc, obscuring often the inner part of the AGN depending on the viewing angle. Slower moving molecular clouds beyond the broad line region emit narrow emission and absorption lines. This region is referred to as the narrow line region. Depending on the accretion rate, two strongly collimated relativistic jets, perpendicular to the accretion disc, can be produced.
According to Padovani et al., 2017, the different wavebands provide different windows to AGN physics: The infrared band is mostly sensitive to obscuring material and dust, the optical/ultraviolet band is connected to the emission from the accretion disc, and the X-ray band traces the emission of a corona. In the radio and gamma-ray band strong non-thermal radiation is emitted, related to the jet.
AGNs are further subdivided into several classes of objects. A unified scheme (e. g. Urry and Padovani, 1995, Netzer, 2015) differentiates them according to their characteristics, such as the viewing angle, the presence of a jet, the presence of emission and absorption lines, and the intensity of the luminosity.

Seyfert galaxies Seyfert galaxies (Beckmann and Shrader, 2012) are radio-quiet, i. e. their radio emission is weak, and typically do not feature a jet. Their spectra show highly ionized emission lines. According to these emission lines, they are divided into two types: Seyfert I and II. In spectra of Seyfert I galaxies, the permitted lines are broader than the forbidden lines, while the permitted and forbidden lines in the spectra of Seyfert II galaxies are both narrow. Intermediate Seyfert galaxies are classified based on the width of the Hβ line.

Radio galaxies According to the classification proposed by Fanaroff and Riley (Fanaroff and Riley, 1974), radio galaxies can be subdivided into two types: Fanaroff-Riley of type I (FR-I) and of type II (FR-II). FR-I feature edge-darkened morphologies and are of low radio powers, while FR-II exhibit edge-brightened morphologies and high radio powers. Comparing the host galaxies, the hosts of FR-II are bluer, less massive and exhibit different lines in optical spectra. The differences are assumed to originate from different accretion rates, i. e. low accretion rates in case of FR-I and high accretion rates in case of FR-II (Saripalli, 2012).

Blazars It is commonly accepted that blazars are quasars, featuring a jet directed towards the observer and high variability. They are the most prominent class in the high-energy and very-high-energy gamma-ray regime (Acero et al., 2015, Wakely and Horan, 2008). Blazars can be subdivided into BL Lac objects (BLL) and flat-spectrum radio quasars (FSRQ), according to the width of their emission lines.

The main characteristics of the described AGNs are illustratively summarized in Figure 2.2.

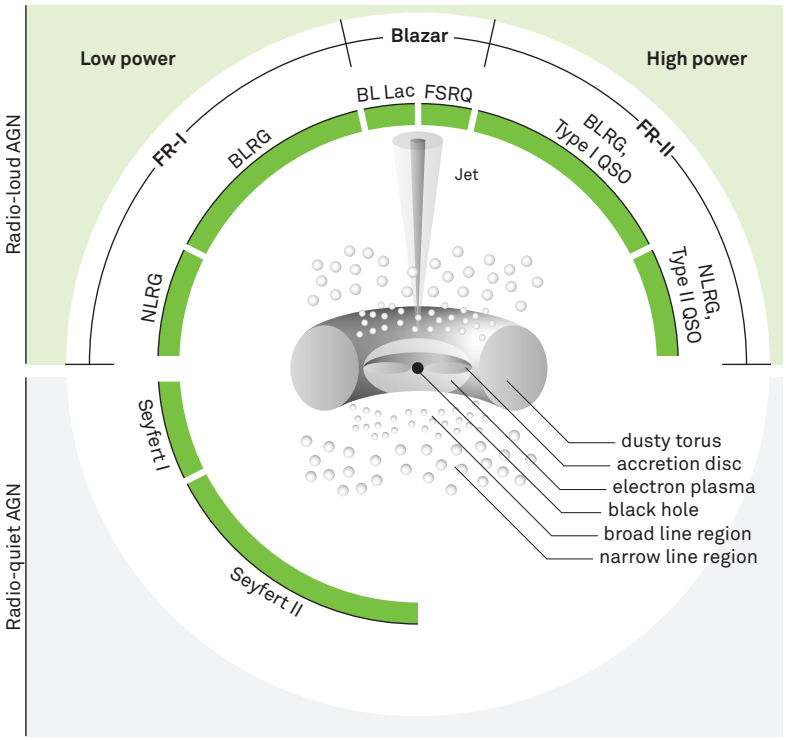

Figure 2.2: Simplified schematic of an Active Galactic Nucleus (AGN). The differentiation between different types of AGNs depends i. a. on the viewing angle, the presence of a jet, the presence of broad lines, and the luminosity according to the unified scheme. This figure is adapted from Beckmann and Shrader, 2012.

The main characteristics of blazars are the emission over the complete electromagnetic spectrum from radio up to very-high-energy gamma rays, large and fast variabilities, high luminosities, and a jet pointing to the observer. These characteristics are commonly explained by non-thermal emission from jet regions at relativistic velocities, inducing a Doppler boosting that boosts this emission, and exceeds thermal emission, such as from the accretion disc.(E. Massaro et al., 2009)

To further probe blazars, e. g. regarding their cosmological evolution and their phenomenology, the consideration of the average spectral energy distributions for different blazar and luminosity classes might be of interest. The so-called blazar sequence has been compiled by Fossati et al., 1998. The spectral energy distributions of several blazars have been divided according to their radio luminosity. Subsequently, they have been averaged, and phenomenological functions have been fitted. It turns out that with increasing luminosities the peak frequencies decrease, i. e. the blazars become redder, and that the ratio between the high- and low-energy hump increases. Moreover, the slope in the gamma-ray regime becomes softer, and the one in the radio regime harder with increasing luminosity. Ghisellini et al., 2017 compiled the sequence with more recent data, and binned the sources according to their gamma-ray luminosity, instead of their radio luminosity. In addition, they derived the sequence for BLLs and FSRQs separately. The corresponding sequences are depicted in Figure 2.3.

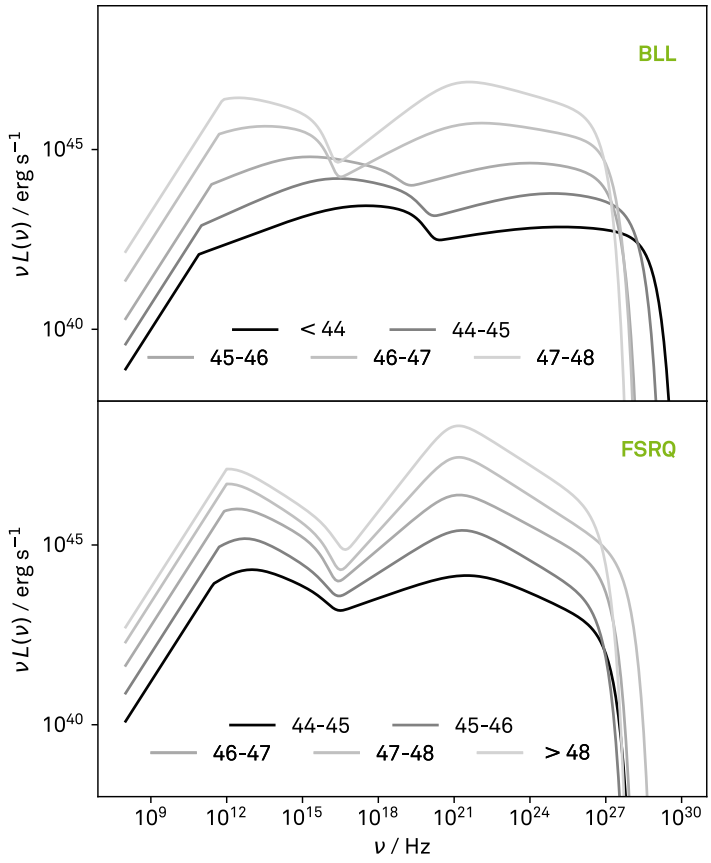

Figure 2.3: Simple phenomenological description of the average spectral energy distributions of BLLs (top) and FSRQs (bottom), dependent on their luminosities in the high-energy regime, referred to as the *blazar sequence*. The phenomenological functions originate from Ghisellini et al., 2017.

2.2 Gamma-Ray Astronomy

Gamma-ray astronomy (Degrange and Fontaine, 2015) is a rather young research field. It completed the upper part of the electromagnetic spectrum, and covers the MeV to TeV energy regime. While the high-energy gamma-ray astronomy started in the 1960s with satellite-based detectors, the access to the very-high-energy regime was only given in 1989 by ground-based detectors.

2.2.1 Detection of Gamma Rays

Depending on the targeted energy of the gamma rays, the measurement with satellite- or ground-based instruments, respectively, is advantageous. Satellite-based instruments are capable to measure the gamma rays directly (Knödlseder, 2016). However, their detector volume is limited and amounts to approximately 1 m^3. Due to the rapidly decreasing flux of typical gamma-ray emitting objects for increasing energy, the necessary observation time to detect very-high-energetic gamma rays in such a small volume would be too long and is not feasible. Thus, satellite-based instruments typically operate in the high-energy regime.

On Earth, it is not possible to measure gamma rays directly, since they do not reach the ground. An indirect method is to utilize the Earth's atmosphere as detector volume (Holder, 2012): Incident gamma rays interact with the atmosphere, creating electrons (e^-) and positrons (e^+) by pair production ($\gamma \rightarrow e^+ + e^-$), which in turn produce photons (γ) by Bremsstrahlung, which also produce electron-positron pairs. Like this, an extensive air shower (Grieder, 2010) emerges. Extensive air showers can also be produced by incident hadrons. In this case, the shower is formed differently, and can be subdivided in three components: The electromagnetic component is initiated by the decay of neutral pions (π^0) into gamma rays. The hadronic component consists of pions, kaons and lighter nuclei, propagating longitudinally along the momentum of the incident primary. The muonic component is produced by charged pions (π^+, π^-) that decay in muons (μ^+, μ^-) and neutrinos ($\nu_\mu, \bar{\nu}_\mu$). Figure 2.4 illustrates extensive air showers, induced by a primary gamma ray and a hadron.

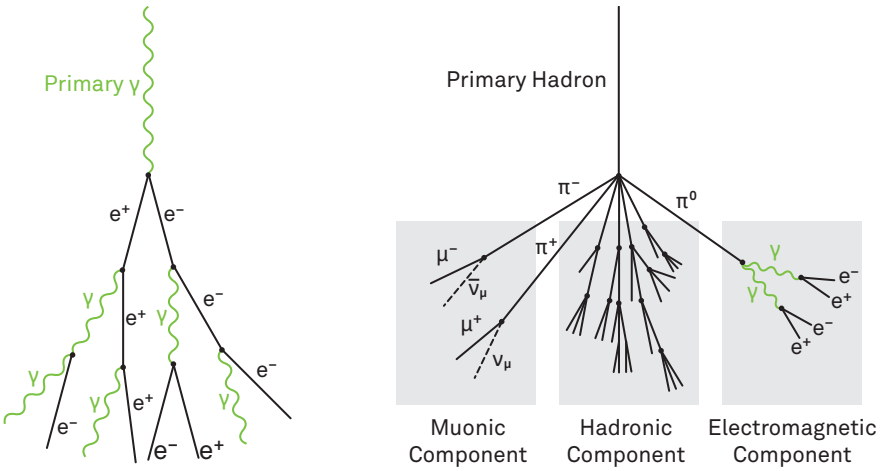

Figure 2.4: Schematic of an extensive air shower induced by a gamma ray (left) and a hadron (right), such as a proton. While the gamma-ray induced shower is of electromagnetic nature, the hadron induced one is composed of a muonic, a hadronic and an electromagnetic component.

The particles of the shower can be detected using the water Cherenkov method. Water tanks, equipped with multiple photon detectors, measure the Cherenkov light that a particle emits when traveling through the water faster than the speed of light in the medium (Cherenkov, 1937). With this method, the GeV to TeV energy regime is covered. The advantages are the wide field of view of 2 steradians, and the possibility to observe 24 hours a day. However, the angular resolution is limited.
An alternative is the use of the imaging air Cherenkov technique. Charged shower particles, traveling through the atmosphere faster than the speed of light in the medium, emit Cherenkov light. A cone of Cherenkov light is formed that produces an ellipse of several hundreds meter diameter on the ground. This Cherenkov light is reflected by a mirror of an Imaging Air Cherenkov Telescope (IACT), located within the ellipse, onto a camera, consisting of several hundreds of photon detectors. This method covers an energy regime from GeV to TeV as well, but in contrast to the water Cherenkov method, the field of view is only a few degrees, and the observation is to be conducted at night. However, the angular resolution is much higher.

2.2.2 Extragalactic Gamma-Ray Astronomy

Extragalactic gamma-ray astronomy probes gamma-ray emitting objects outside of our Galaxy. These objects are mostly AGNs and pulsars (Acero et al., 2015). However, extragalactic gamma-ray astronomy is limited by the Extragalactic Background Light (EBL), since very-high-energy photons γ_{VHE} are absorbed when they interact with low-energy photons γ_{EBL} from the EBL via pair production ($\gamma_{\mathrm{VHE}} + \gamma_{\mathrm{EBL}} \rightarrow e^+ + e^-$). The EBL photons, lying mostly in the ultraviolet to infrared regime, originate from emission from the formation and evolution of galaxies (Domínguez et al., 2011). The resulting attenuation of the intrinsic flux $\Phi_{\mathrm{intrinsic}}(E)$ of a source is modeled with an exponential attenuation factor as a function of the optical depth $\tau(E, z)$ dependent on both the energy E and the redshift z:

$$\Phi_{\mathrm{measured}}(E, z) = \Phi_{\mathrm{intrinsic}}(E) \cdot \exp\left(-\tau(E, z)\right) \ . \tag{2.1}$$

Domínguez et al., 2011, Franceschini, A. and Rodighiero, G., 2017, Kneiske, T. M. and Dole, H., 2010 and Mazin, D. and Raue, M., 2007 derived e. g. limits on the density of the EBL.
The understanding of the EBL is important to probe intrinsic characteristics of very-high-energy gamma-ray emitting objects. For this purpose, sources with different redshifts are needed. However, the number of sources with large redshifts are limited. Only 15 sources with $z > 0.3$, 5 sources with $z > 0.5$ and 2 source with $z > 0.9$ have been detected so far according to the TeVCat (`http://tevcat.uchicago.edu`).
To derive the intrinsic source spectrum from a measured very-high-energy spectrum, the spectrum has to be de-absorbed. Beyond that, the EBL can also be utilized to estimate the redshift of sources with unknown redshift.

2.3 The MAGIC Telescope System

Figure 2.5: Picture of the MAGIC telescopes at La Palma, Canary Islands (Ishio, 2016).

MAGIC is a stereoscopic system of two 17 m Imaging Air Cherenkov Telescopes, situated at 2200 m above sea level at the Roque de los Muchachos Observatory on the Canary Island of La Palma, and operates in an energy regime of about 50 GeV to 50 TeV. In Figure 2.5 a picture of the two telescopes is shown.
The physics program spans from astroparticle and fundamental physics to extragalactic and Galactic science. For instance, multi-wavelength, correlation and variability studies give insights into the emission processes of e. g. Active Galactic Nuclei, and the study of these sources with different redshifts constrains the model of the absorption by the extragalactic background light. In the field of astroparticle and fundamental physics, the search for dark matter, the understanding of the origin of cosmic rays and, besides others, conclusions from follow-up observations of astroparticle neutrinos, gamma-ray bursts and gravitational waves are of great interest.
In 2011 and 2012, MAGIC underwent a major hardware upgrade (Aleksić et al., 2016a) to unify the telescope system and to improve the overall performance, involving the replacement of the MAGIC-I camera, the receivers and the trigger systems, and the upgrade of the readout systems to Domino Ring Sampler version 4 chips. Since late 2014, the sampling speed of the readout system is 1.64 GSample/s. Both cameras are now equipped with 1039 photomultiplier tubes of 0.1°, resulting in a field of view of $\approx 3.5°$. The dish of the telescope is composed of multiple mirrors, adding up to a reflective surface of $\approx 235\,\mathrm{m}^2$. Signals, induced by photons and reflected by the mirror to the photomultiplier tubes, are transmitted through optical fibers to the trigger and readout systems. The standard trigger consists of different trigger levels: The level 0 trigger is fulfilled, if the signal in one pixel exceeds a specific discriminator threshold. If 3 neighboring pixels in a single telescope triggered the level 0, the level 1 is triggered, and if both telescopes issue a level 1 trigger in a given time range, also the level 3 is triggered, and the readout system stores the event.

Aleksić et al., 2016b evaluates the performance of MAGIC, subsequent to the upgrade, and based on Monte Carlo simulations and observations of the Crab Nebula, deriving an energy threshold at trigger level of ≈ 50 GeV for low zenith angle observations. The energy resolution, deduced from the relative deviations between the estimated and true energy, is ≈ 15 % at medium energies (few hundreds of GeV). The angular resolution, obtained from a two-dimensional Gaussian fit, is 0.07° at 250 GeV and improves with increasing energy. Another important performance parameter is the sensitivity, stating a minimum gamma-ray flux the telescopes can detect with a significance of 5 σ in 50 h of observations. The differential sensitivity, independent from the spectral properties of the source, and calculated using a simplified approach, is about 1.2 times the flux of the Crab Nebula at medium energies.

MAGIC Analysis and Reconstruction Software

The MAGIC Analysis and Reconstruction Software (*MARS*) is a software package of several executables and macros to obtain high-level data products such as energy spectra or light curves from raw data measured with the MAGIC telescopes. A short introduction of the executables used for a standard MAGIC analysis of observations performed in the year 2016 is given in the following. A detailed documentation of these executables is available at the internal websites[1] of the MAGIC Collaboration and serves as reference here.

merpp (**Mer**ging **P**rocess **P**rogram) converts the raw data files to standard *MARS* files in ROOT[2] format. In addition to the shower events recorded during data taking, information from subsystem reports are stored. This information is important for subsequent analysis steps. For instance the weather and trigger reports are used for the data quality check and the laser reports on the atmospheric transmission are necessary for an energy correction of observations under slightly cloudy conditions.

sorcerer (**S**imple, **O**utright **R**aw **C**alibration; **E**asy, **R**eliable **E**xtraction **R**outines) is mainly taking care of the calibration and the extraction of charge and time information.
The night sky background, comprising i. a. ambient lights or light from stars, induces continuously currents in every photomultiplier tube. Weather and moon conditions, influencing the night sky background, are varying only slowly, resulting in an approximately stable baseline in every readout sample. The baseline is calculated using pedestals (i. e. randomly triggered readout samples) by fitting a normal distribution to the distribution containing all signals of every readout slice of multiple pedestal readout samples.
For the extraction of charge and time information of the incident Cherenkov photons, the baseline is subtracted and the *sliding window method* is applied. The charge is determined by the largest sum of 5 readout slices (≈3 ns) in a readout range of 50 time slices. The same 5 readout slices are used to obtain the arrival time, by averaging the times weighted with the signal at that time. In order to convert the integrated readout counts into a number of photoelectrons, the *F-factor method* calculates conversion factors for each pixel from the averages and fluctuations of dedicated calibration light pulses. Timing differences between the pixels, caused e. g. by different optical fiber lengths, are also extracted from dedicated calibration pulses and are used to corrected the arrival times.

star (**St**andard **A**nalysis and **R**econstruction) is responsible for the cleaning of camera images and the parameterization thereof.

[1] `http://wiki.magic.pic.es`
[2] `https://root.cern.ch`

Before parameters characterizing the camera image can be derived, the camera image needs to be cleaned. Although the night sky background is taken into account during the charge and time extraction, fluctuations of the night sky or e. g. electronic noise remain in the image. Specifically within the edge region of the shower projected to the camera or in low-energy shower images, the number of Cherenkov photons is in the order of these fluctuations. On the one hand, it is important to keep as many Cherenkov photons as possible to keep as much information as possible and to keep low-energy events. On the other hand, fluctuations need to be suppressed to not distort the reconstruction and to ensure a robust parameter determination.
The *Sum Cleaning method* is performing a preselection of pixels by summing up the signals in groups of two, three and four neighboring pixels and keeping only those exceeding 1.8 phe, 1.3 phe and 1.0 phe per pixel, respectively, and by requiring an arrival time range of groups of 0.82 ns, 1.15 ns and 1.80 ns, respectively. Subsequently, a *Time-constrained Absolute Image Cleaning* is conducted on the preselected pixels. In this algorithm, pixels with signals above 6 phe are selected as core pixels and survive the cleaning if the difference between the mean arrival time of the core pixels and a specific core pixel is not larger than 4.5 ns. Neighboring pixels to these core pixels are selected as boundary pixels if they exceed 3.5 phe and if the time difference to the closest core pixel is smaller than 1.5 ns.
The first parameters to describe the camera image were introduced in Hillas, 1985, over the years the number of image parameters increased more and more. Some of the most common image parameters are e. g. `Size` (the total charge of all pixels), `Length` (second moment along the major axis of the shower ellipse) and `Width` (second moment along the minor axis of the shower ellipse). An illustration of some image parameters is shown in Figure 2.6. A full list of all image parameters available at this analysis level can be found in Appendix B.4.

superstar performs a stereoscopic reconstruction by combining the results of the single telescopes, allowing a 3-dimensional geometrical reconstruction of the shower.
The shower axis of the air shower is in line with the direction of the incident particle, and in case of gamma rays, it points to the corresponding gamma-ray source. Correspondingly, also the shower projected onto the camera image points to the direction of the incident particle, but only the direction vector and not the position vector is geometrically derived with a single telescope. But since the telescopes see the air shower from different viewing angles, the combination of both images provides the missing information. The point of impact of the shower axis on the ground is reconstructed by the intersection point of the elongated shower axes projected onto the two camera images, while the (geometrical) reconstructed source position in the camera is obtained by the intersection of the shower axes of both telescopes projected in one camera system. By means of these positions, further image parameters are calculated, as for instance the `Impact` (the distance between a telescope and the reconstructed core impact point) and `MaxHeight` (the estimated height above the telescope of the shower maximum), improving i. a. the suppression of hadrons. In turn, the `MaxHeight` is needed to determine the `CherenkovRadius` and `CherenkovDensity`, dependent on the light emitted by an 86 MeV electron at `MaxHeight`, improving the energy reconstruction. The reconstructed source position is illustrated in Figure 2.6. A full list of all image parameters available at the analysis level of *superstar* can be found in Appendix B.4.

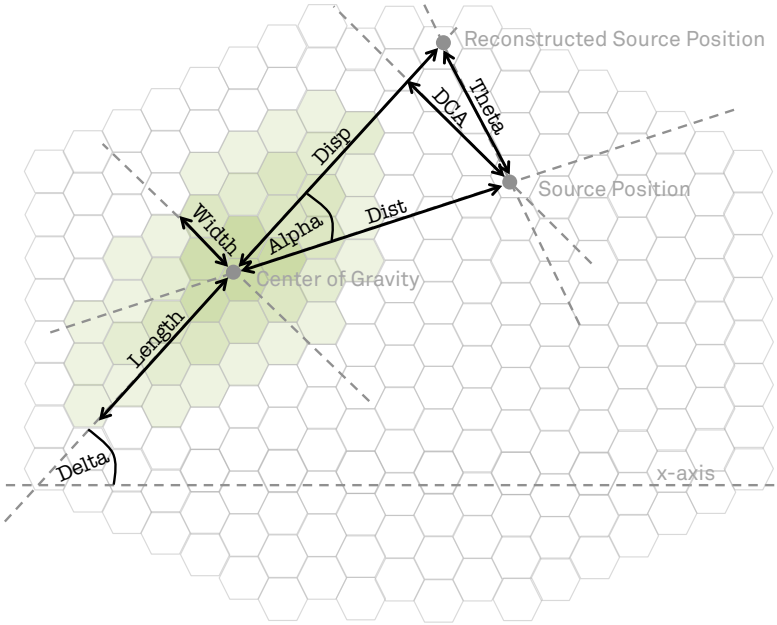

Figure 2.6: Sketch of some image parameters from different analysis levels, characterizing the shower in the camera image. `Length`, `Width`, `Delta`, the center of gravity and the source position are calculated with *star*, as well as the source dependent parameters `Alpha`, `Dist` and `DCA`. The reconstructed position is initially determined with *superstar*, and is refined with *melibea*, providing also the corresponding `Disp` and `Theta`.

coach (**C**ompressed **O**steria **A**lias **C**alculation of the **H**adronness parameter) generates several models to reconstruct the energy and the direction of an event, and to suppress hadronic background.
For stereoscopic observations the energy is reconstructed using the *Lookup Table method*. The lookup tables are two-dimensional tables in $\sqrt{\log10(\texttt{Size})}$ and `Impact`/`CherenkovRadius`, filled with `CherenkovDensity` $\cdot E_{\text{true}}$/`Size` with E_{true} from Monte Carlo simulations, and with additional corrections. These corrections concern the dependency of the reconstruction of `CherenkovDensity` on the zenith angle of the telescope and the geomagnetic field, and truncated shower images at the edge of the camera. For every bin of the table, the mean and the standard deviation is calculated and used in subsequent analysis steps to assign energies and energy uncertainties to the events.
In order to suppress hadronic background, a classification of gamma rays and hadrons is performed with a Random Forest. The Random Forest model assigns every event a so-called `Hadronness`, related to a probability of the event being a hadron. Details on the creation of Random Forest classification models in general and within *MARS* are given in chapter 3 and in section 9.1, respectively.
The direction reconstruction is performed for each telescope individually with a Random Forest regression. In this regression the continuous parameter `Disp` is estimated, representing the distance between the center of gravity and the reconstructed source position (cp. Figure 2.6). The Random Forest regression in general is described in detail in chapter 3.

melibea (**Me**rge and **L**ink **I**mage parameters **B**efore **E**nergy **A**nalysis) applies the models generated with *coach* to obtain estimations for the energy, direction and particle type. Based on them, *melibea* also calculates more sophisticated parameters.
While the determination of the `Hadronness` and its uncertainty is straight forward, the energy estimation is more complex. In a first step, the energy and its uncertainty are estimated for both telescopes individually. Subsequently, the energies of both telescopes are averaged and weighted with their uncertainties.
The reconstruction of the direction is even more complicated. In *superstar*, a geometrical reconstruction of the source position has already been conducted by using the intersection of the major shower axes of the two telescopes. Especially for small angles $\Delta\delta$ between the shower axes, the reconstruction is not very precise. A more refined method is the use of Random Forest regressions in combination with stereoscopic information. The application of the Random Forest models leads to `Disp` values for each telescope individually, but since they represent only the distance from the center of gravity, the reconstructed source position is ambiguous. To determine reconstructed source positions for each telescope separately, the position is chosen along the major shower axis in direction of the asymmetry. For a combined reconstructed source position, the weighted average of the closest pair of possible positions is selected (cf. Figure 2.7). The squared distance between the positions of the closest pair needs to be smaller than 0.05 deg^2, otherwise the stereo reconstruction is set as not valid. With this Stereo-`DISP` reconstructed source position, the major shower axes are refined by connecting the center of gravity and the reconstructed source position. Correspondingly, parameters such as `MaxHeight`, `Impact`, `CherenkovDensity` and `CherenkovRadius` are re-calculated. Moreover, additional validity checks are performed, as e. g. the check if the relation between the estimated energy and impact is reasonable. All parameters available at the analysis level of *melibea* are listed in Appendix B.4.

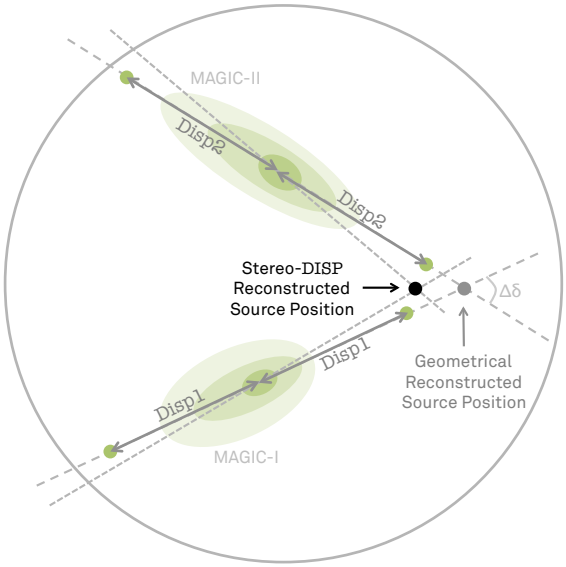

Figure 2.7: Sketch of the direction reconstruction approach. The Stereo-`DISP` reconstructed source position is determined by the weighted average of the positions of the closest pair of possible positions.

odie calculates significances for given sky coordinates, based on specified source (so-called on) and background (so-called off) regions.
Imaging Air Cherenkov Telescopes measure not only gamma rays from an assumed source, but also a certain amount of background at the same time, such as cosmic rays or gamma rays from the diffuse gamma-ray background (Fornasa and Sánchez-Conde, 2015). While the hadronic background can be suppressed to a large extent, it is not possible to distinguish diffuse gamma rays and gamma rays from the assumed source. In order to figure out how likely a gamma-ray source is located in the supposed region, a significance value is calculated. The most reliable (and in *MARS* implemented) significance value has been proposed in Li and Ma, 1983, and is based on a statistical hypothesis test. The null hypothesis says, that all measured events are due to background, and consequently, that no events stem from the supposed source, while the alternative hypothesis says, that the events originate from both, the background and a source. These hypotheses are embedded in a Likelihood-ratio test, where the significance value S is calculated as follows:

$$S = \sqrt{2\left(N_{\text{on}} \cdot \ln\left[\frac{1+\alpha}{\alpha} \cdot \frac{N_{\text{on}}}{N_{\text{on}}+N_{\text{off}}}\right] + N_{\text{off}} \cdot \ln\left[(1+\alpha)\frac{N_{\text{off}}}{N_{\text{on}}+N_{\text{off}}}\right]\right)}\,. \tag{2.2}$$

This significance is given in units of the standard deviation and indicates the probability of a rejection of the null hypothesis, related to the probability that only background events have been observed. For instance, a significance value of $S = 5\,\sigma$ implies a probability of 99.99997 %. The parameters N_{on} and N_{off} denote the number of events observed in the on and off region, respectively, while α specifies the ratio between the observation times of the on and off regions.
Standard stereo observations are performed in wobble mode, i. e. the telescopes are pointing at the source with a specific offset (0.4 deg in most cases). This allows on and off measurements at the same time under the same conditions, such as weather and light conditions. The specification of a source position is necessary for the calculation of the parameter θ (the distance between the reconstructed and the specified source position). The source position is different for the on and off regions, but at the same distance from the camera center. It is also possible to use multiple off regions at once, increasing the observation time and precision of the background measurement. A typical number of off regions is three, resulting in an α of 1/3. The on and off regions are defined by a cut in the θ^2 distributions, referring to a circular region around the specified positions of the off regions.
In addition to the θ^2 cut, cuts optimized for specific energy regimes are applied. The events have to fulfill `Hadronness` < 0.28 and `Size` > 60 phe for the low energy (LE) range and `Hadronness` < 0.16 and `Size` > 300 phe for the full range (FR). For the high energy (HE) range the events need to pass `Hadronness` < 0.1, `Size` > 400 phe and `Energy` > 1000 GeV.

2.4 Instruments for Multi-Wavelength Astronomy

The electromagnetic spectrum of AGNs ranges over multiple magnitudes in energy and flux. To measure the flux for different energies, multiple diverse instruments are necessary. In the course of this thesis, the data recorded by several instruments, capable to measure radio, infrared, X-ray or high-energy gamma rays, have been utilized. These instruments are described in the following.

2.4.1 *Fermi*-LAT

The *Fermi* Gamma-ray Space Telescope is a space observatory, equipped with two instruments: The Large Area Telescope (LAT), dedicated to monitor high-energy gamma-ray sources between 20 Mev to 300 GeV in an all-sky survey, and the Gamma-ray Burst Monitor (GBM), designed to detect transient events between approximately 8 keV and 40 MeV (Meegan et al., 2009). The *Fermi* satellite was launched in June 2008, orbiting now the Earth in about 96 minutes. To survey the whole sky, the LAT requires two orbits, i. e. every gamma-ray sources is monitored every three hours.
The *Fermi*-LAT – the primary instrument of the observatory – measures the energy, arrival time and direction of high-energy gamma-ray photons. The instrument is composed of different detectors with diverse functions (Atwood et al., 2009). Silicon strip detectors, interleaved with high-Z material, form the converter-tracker. Incident gamma rays, interacting with this material, convert to electron-positron pairs. The path of these charged particles through the detector is then tracked by the silicon strip detector. A calorimeter beneath the tracker measures the energy deposition of the electromagnetic cascade, the electron-positron pair is producing. It is made of multiple cesium iodide crystals. The anticoincidence detector is the outermost detector, surrounding the tracker, and rejects charged-particle background, such as the cosmic-ray background. The detector consists of plastic scintillators, read out by photomultiplier tubes, and is segmented to distinguish charged cosmic rays from charged particles induced by the electron-positron pair in the calorimeter.
More details about the technology and the performance of the LAT can be found in Atwood et al., 2009 and Ackermann et al., 2012b.

2.4.2 *Swift*-XRT

The *Swift* satellite (Gehrels et al., 2004) is a space observatory with three instruments on board, capable to measure photons in multiple wavelengths. The Burst Alert Telescope (BAT) is dedicated to detect gamma-ray bursts, and to provide fast alerts to other telescopes. It is designed with a large field of view, and a localization accuracy of a few arcminutes in time frames of a few seconds. The UltraViolet/Optical Telescope (UVOT) is capable to perform optical and ultraviolet photometry, and features localization accuracies in the sub-arcsecond regime. The X-Ray Telescope (XRT) conducts measurements in the energy regime of 0.2 – 10 keV. It is designed to perform spectroscopy of medium resolution and light curves of high resolution. Since almost all materials absorb X-rays, no conventional lens or mirror systems can be utilized to focus the incident radiation. However, if X-rays strike a material under a very large angle, total reflection occurs and the light is entirely reflected, i. e. they only graze the material. XRT's Wolter Type 1 telescope employs a mirror system, exploiting this effect of grazing incidence, to focus the X-rays to the *XMM/EPIC MOS* CCD camera.
More details about the technology and the performance of *Swift* and the XRT can be found in Burrows et al., 2005 and Gehrels et al., 2004.

2.4.3 VLA

The Very Large Array (VLA) is a radio astronomy observatory for wavelengths in the centimeter to decimeter regime, situated in New Mexico (Thompson et al., 1980). 27 individual telescopes of 25 m diameter can be combined to an interferometer. They are installed on rails in a Y-configuration, i. e. in three arms at an angle of 120° between them. The 9 telescopes per arm can be moved over a distance of 21 km. Like this, the array configuration can be varied between four common layouts, with the most distant telescopes between 590 m and 21 km from the intersection point of the arms, resulting in different performances and array properties. The interferometer is capable to measure at four different frequency bands, i. e. at 1.34 – 1.73 GHz, 4.5 – 5.0 GHz, 14.4 – 15.4 GHz, 22 – 24 GHz. In 2001, after 20 years of operation, major hardware and software changes have been initiated to keep up with state-of-the-art technologies, which have been completed in 2011.
More details about the technology and performance of the VLA can be found in Thompson et al., 1980 and about the expanded VLA in Perley et al., 2011.

2.4.4 WISE

The Wide-field Infrared Survey Explorer (WISE) is a space telescope, meant to perform an all-sky survey at four infrared wavelength, i. e. at the bands 3.4, 4.6, 12, 22 μm (E. L. Wright et al., 2010). An afocal infrared telescope with 40 cm diameter directs the radiation to four focal plane detector arrays. For the short-wavelength channels, an HgCdTe array is utilized, while the Si:As array is employed for the long-wavelength channels, each array consisting of 1024 × 1024 pixels. A cryostat, filled with solid hydrogen is cooling the telescope to less than 12 K. Every 11 seconds, the telescope is taking an image, covering a field of view of 47 arcminutes.
A more detailed description of the mission and its performance can be found in E. L. Wright et al., 2010.

3 The Data Mining Guide

The Guide is definitive.
Reality is frequently inaccurate.

— Douglas Adams, *The Hitchhiker's Guide to the Galaxy*

The term *Data Mining* refers to the extraction of knowledge from huge data samples and combines machine learning, statistics and methods to process a vast amount of data. In recent years, the amount of collected data increased heavily, requiring a fast processing and information extraction, which was often difficult with conventional methods. Thus, data mining techniques became more and more important. It has wide applications, ranging from science to economy and industry.

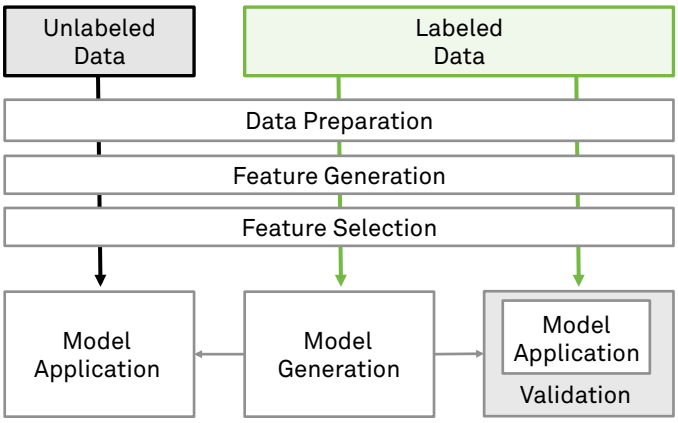

Figure 3.1: Flow chart of a typical data mining approach.

Figure 3.1 illustrates a typical data mining strategy. The flow chart – and the following sections as well – is adapted to strategies using supervised learning, i. e. using labeled data, and Random Forests, since they are relevant for the subsequent chapters. First, all data have to undergo a preprocessing, comprising a data preparation, a feature generation and a feature selection. The data preparation ensures that the subsequent algorithms are capable to cope with the data, i. e. invalid data are for instance removed, or categorical features are transformed. The feature generation creates multivariate relations between features that support the successive algorithms. The feature selection reduces the dimensionality of the feature space and improves thereby the description of the sample's density. In a next step, the labeled data are used to generate a model, such as a model to determine a class (classification) or to estimate a parameter (regression). This model is validated by applying the model to labeled data and comparing the label and the prediction. Based on this comparison, appropriate measures are calculated to estimate the performance of the model, such as the precision or accuracy of the model. Subsequently, the model is applied to the unlabeled data to estimate the information of interest.

3.1 Preprocessing

The preprocessing comprises the steps of data preparation, feature generation and feature selection, ensuring that subsequent algorithms can cope with the data and that the corresponding methods can get the best out of the data. This is a very important and complex step.

3.1.1 Data Preparation

The data preparation concerns various aspects: On the one hand the features themselves, and on the other hand the values of the features. These aspects and how to deal with them are listed below.

Range of values Some algorithms benefit or even require specific value ranges or values that are standardly distributed. Therefore, appropriate scalers can be applied. In case of algorithms based on decision trees, this is not necessary.

Missing values Often a feature contains missing values. This can occur, if for instance an information was simply not collected, or if a method returns invalid results, such as a failed fit. One possibility to remove these missing values is to simply remove the corresponding feature or sample. However, this also removes information, and is often not advisable. Instead, the missing value can be set to a specific value, such as the mean, median, or a value outside of the range of the remaining samples.

Categorical features The information is not always collected in continuous values, but in specific categories, such as a particle type, which are not sortable. Many algorithms cannot handle these kind of features, and they have to be transformed beforehand. One procedure is the *one-hot* encoding, creating a feature for every category with binary values.

Correlated features Highly correlated features unnecessarily inflate the feature space, increase the dimensionality and contain a large amount of redundant information. Before they can be removed, the correlation needs to be calculated. Typical correlation coefficients are the Pearson correlation coefficient, detecting linear dependencies between features, or the Spearman's rank correlation coefficient, detecting monotonic dependencies between features. A common threshold to define two features as highly correlated is 0.95, i. e. features with coefficients with absolute values above 0.95. To decide which of the two highly correlated feature is to be removed, a criterion judging the relevance of the features to the classification or regression task is considered.

3.1.2 Feature Generation

Frequently, the available information is not beneficially represented for the given problem by the features. Thus, it can be valuable to generate new features, containing admittedly the same information, but expressed differently in a more appropriate form. Like this, the feature space and dimensionality can subsequently be reduced by removing irrelevant features. This is specifically important for limited sample sizes, since the description of the sample's density is difficult for a large dimensionality. To generate these new features, the best procedure is to use expert knowledge to combine and transform the available information smartly. Alternatively, some kind of brute force can be applied, calculating a vast of features by combining and transforming the features virtually blindly.

3.1.3 Feature Selection

As described previously, it is important to avoid too large features spaces, since the description of the sample's density is made more difficult. For the reduction of the number of features, several algorithms (e. g. Hastie et al., 2009, Guyon and Elisseeff, 2003) are available. A selection of algorithms is presented in the following:

Brute force Every possible feature set is tested, and the most performant one, according to a chosen performance measure, is selected. Such an approach is able to find the optimal feature set. However, it is very computationally intensive, and practically not feasible.

Minimum Redundancy Maximum Relevance The *Minimum Redundancy Maximum Relevance* algorithm (Peng et al., 2005) selects iteratively features, exhibiting a large relevance regarding the target label, and a small redundancy regarding the other features in the set. Drawbacks of this algorithm are that neither multivariate characteristics in the feature set, nor the algorithm to generate the model in the subsequent step are taken into account.

Univariate selection In an univariate feature selection approach, a user-defined number k of best features is selected, according to e. g. specific rankings. For instance, the k features with the largest mutual information or with the best F-test values are selected. Another possibility is to select the k most important features. Methods like this are usually very fast, but do not consider any multivariate dependencies and is thus very unlikely to find an optimal feature set.

Forward selection The forward selection procedure is initialized with an empty feature set and then iteratively complemented until a certain abort criterion is fulfilled. In each step, every single feature is added at a time to the already selected feature set, and an estimator, such as a Random Forest, is trained and evaluated. The feature that improves the performance the most, is added to the feature set, and the procedure is repeated. In contrast to univariate methods, the importance as judged by an estimator considers multivariate relationships in the same way the estimator itself does.

Backward selection The backward selection procedure starts by training an estimator using all available features, and calculating the importance of every feature. The least important feature is removed from the feature set, and the estimator is once again trained and evaluated. This procedure is recursively repeated until only one feature remains in the feature set. For every feature set, a performance measure is determined, and the feature set corresponding to the highest performance is chosen. This approach involves a reasonable computational load, while considering the estimator and multivariate structures within the data sample.

3.2 Classification and Regression

Typical machine learning methods solve classification and regression tasks. A regression refers to the estimation of a continuous parameter, while a classification concerns the estimation of a class affiliation. Most commonly, the classification tasks consist of the discrimination between two classes, referred to as a 2-class problem. Nonetheless, machine learning methods are also capable to cope with so-called multi-class problems. In case of a multi-class problem, different approaches can be pursued, depending on the classifier to be used. Many classifiers are capable to handle multiple classes, and generate one model that provides a score for each class. However, some classifiers can only deal with binary classes.

In this case, the multi-class problem has to be transformed to a 2-class problem. In the *One-vs-Rest* approach, one model is generated for each class, while the *One-vs-One* approach generates a model for every pair of classes. Multi-class problems are not to be confused with multi-label problems. In multi-label problems, a sample has multiple labels, which are estimated at once.
Widely used supervised learning methods, i. e. methods generated with labeled data, are tree-based methods, such as the Random Forest. Figure 3.2 illustrates a binary decision tree. Starting at the root node, i. e. starting from all samples, a split in a particular feature is chosen, and the sample is split in two child nodes. Proceeding from these parent nodes, the samples are again divided, according to splits in different features. This is continued until an abort criterion is fulfilled, such as a minimum number of samples in a node. This process is also referred to as the growth of the tree. The final nodes are called leaves, and the number of layers of nodes is the depth of a tree. According to the properties of the samples in a leaf, a value, related to the assigned task, is calculated. For a classification task, this value indicates the class affiliation, and is determined by averaging the classes of the remaining samples. For a regression task, this value is the average over the continuous parameter to be estimated.

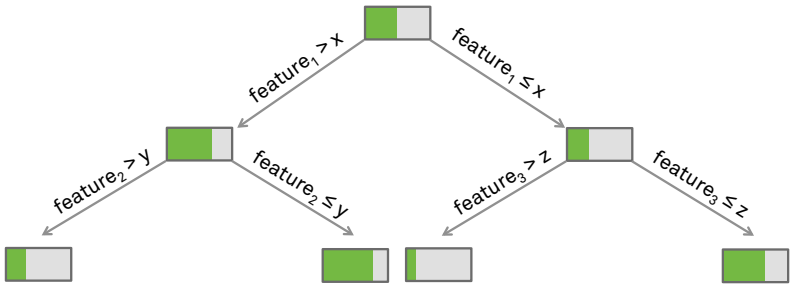

Figure 3.2: Illustration of a decision tree.

To determine a feature and a value to split a specific node, a procedure referred to as split selection is applied. The criterion for the best split is based on the maximum difference of a quality measure Q_{parent} of the parent node and the sum of the quality measures Q_{left} and Q_{right} of the left and right child nodes, scaled to the number N of samples at a node:

$$\text{criterion}_{\text{best split}} = Q_{\text{parent}} - \frac{N_{\text{left}}}{N} \cdot Q_{\text{left}} - \frac{N_{\text{right}}}{N} \cdot Q_{\text{right}} \,. \tag{3.1}$$

Different quality measures (Hastie et al., 2009) can be utilized, such as the Gini impurity

$$Q^{\text{Gini}}_{\{\text{node}\}} = \sum_{i=0}^{n_{\text{classes}}} p_i(1 - p_i) \tag{3.2}$$

or the cross-entropy

$$Q^{\text{Entropy}}_{\{\text{node}\}} = -\sum_{i=0}^{n_{\text{classes}}} p_i \log(p_i) \tag{3.3}$$

with $p_i = N_i/N$ in case of classifications, or the mean squared error for the regression of the continuous parameter x

$$Q^{\text{MSE}}_{\{\text{node}\}} = \frac{\sum x_i^2}{N} - \left(\frac{\sum x_i}{N}\right)^2 . \tag{3.4}$$

A Random Forest (Breimann, 2001) is an ensemble of decision trees that can be utilized for both classification and regression tasks (Breiman et al., 1984). By combining the predictions of the multiple decision trees, the robustness and generalization is improved compared to a single decision tree. This robustness is achieved by inducing randomness in different aspects during the generation of the ensemble. On the one hand, a subset of features is randomly chosen and considered for the split selection. On the other hand, each decision tree is generated using a different bootstrapped subsample. The predictions of the individual trees are averaged to receive the final prediction, or – depending on the implementation – a majority vote is conducted. The increase of robustness of a Random Forest compared to a single decision tree is proved by the decrease of the variance. However, the decrease of the variance is usually accompanied by an increase of the bias.
The most important hyperparameters that can be tuned to optimize a Random Forest are listed below:

Split selection criterion In case of classifications, mostly the Gini impurity or the cross-entropy are considered as quality measures for the split selection, while for regressions the mean squared error is commonly used as a criterion.

Number of trees Increasing the number of decision trees in the Random Forest, the variance is decreased up to a specific value. The number of trees should be large enough to reach this value.

Number of features considered at a node The larger the number of features that are considered at a node for the split selection, the larger the bias. The smaller the number, the smaller the variance. Aiming for a trade-off between a small bias and a small variance, the square root of the number of available features turned out to be beneficial for classification problems, and the maximum number of features for regressions.

Minimum number of samples for split By demanding a minimum number of samples in a node to further split the node, the growing of the tree is stopped.

Minimum number of samples in a leaf A split of a node is only allowed, if the resulting nodes contain a minimum number of samples. This demand aborts the growing as well.

Maximum depth of the tree The maximum number of layers of nodes stops the expansion of the tree as well.

3.3 Validation

The validation of a generated model, such as a classification or regression model, is essential in many regards. On the one hand, it is important for the selection of the final model. For instance, during the feature selection, the performance of various feature sets needs to be evaluated to decide which of the tested feature sets is best-suited. Moreover, when optimizing the hyperparameters of a model, a criterion is necessary to judge the model obtained with a specific set of hyperparameters. On the other hand, it is of interest how reliable the prediction of the model is. Various measures to evaluate the performance are available. In case of classification tasks, many of them are based on the following quantities:

True positives (tp) Number of samples which are correctly predicted as the positive class.

False positives (fp) Number of samples which are falsely predicted as the positive class.

True negatives (tn) Number of samples which are correctly predicted as the negative class.

False negatives (fn) Number of samples which are falsely predicted as the negative class.

Derived from these quantities, the following performance measures can be calculated:

$$\text{Precision, Purity} = \frac{\text{tp}}{\text{tp} + \text{fp}} \tag{3.5}$$

$$\text{True Positive Rate, Recall, Efficiency} = \frac{\text{tp}}{\text{tp} + \text{fn}} \tag{3.6}$$

$$\text{False Positive Rate} = \frac{\text{fp}}{\text{fp} + \text{tn}} \tag{3.7}$$

These performance measures are specifically defined for 2-class problems – a positive and a negative class. To utilize them in the context of multi-class problems as well, the *One-vs-Rest* strategy can be pursued: One class is considered the positive class, while the remaining ones are treated as the negative class. This is repeated for every class. Like this, the multi-class problem is decomposed into multiple 2-class problems.
A further performance measure is derived with the Receiver Operating Characteristic (ROC) curve, i. e. the true positive rate as a function of the false positive rate. Figure 3.3 illustrates the ROC curve for a successful classification, and for a classification model which is randomly guessing. The Area Under this Curve (AUC) quantifies the ROC curve. The advantage of the AUC score is that its calculation does not require the application of a cut in the output of the classifier.

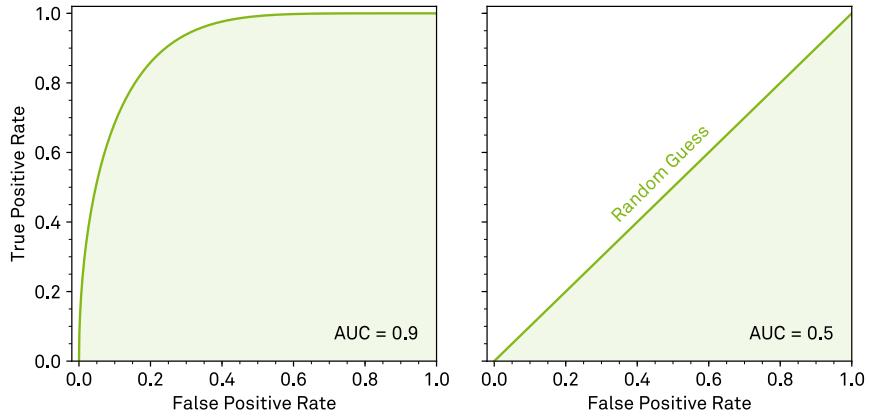

Figure 3.3: Examples of a Receiver Operating Characteristic (ROC) curve for a successful classification and a classification model which is randomly guessing. The corresponding Areas Under the Curves (AUCs) quantifies the performance, derived from the ROC curve.

In case of regression tasks, it is common to calculate the difference between the true and the predicted parameter. The resulting distribution, possibly for different ranges of the predicted parameter, can be quantified by calculating the bias in terms of the median, and the resolution in terms of an interquantile range.

The determination of these performance measures should not be carried out on the same data sample the model has been generated with. In principle, the model can memorize the data, and if the model is tested with the same data, the performance is greatly overestimated. This behavior is known as overfitting.

A beneficial method to properly evaluate the performance is cross validation, especially if the data sample is too small to provide a separate test sample. The method of cross validation is illustrated in Figure 3.4. In a k-fold cross validation, the labeled data sample is divided in k subsamples. $(k-1)$ samples are utilized to train the model, and the remaining one is used for testing. This is repeated k times, such that every subsample has been used once for testing. The k performance measures of each iteration are then averaged. The corresponding uncertainty in terms of the standard deviation indicates the robustness of the model. Values for k of 5 and 10 are recommended in the literature (Breiman and Spector, 1992) as a compromise between training samples large enough to derive a representative estimate of the performance, and training samples of sufficiently large dissimilarity. In case of $k = 2$, the two training samples are disjoint, but the samples are only half the size of the sample the final model is generated with, introducing a bias. For k equaling the sample size, referred to as *leave-one-out*, the training samples are almost the same, underestimating the variance, and thus, overestimating the robustness of the model.

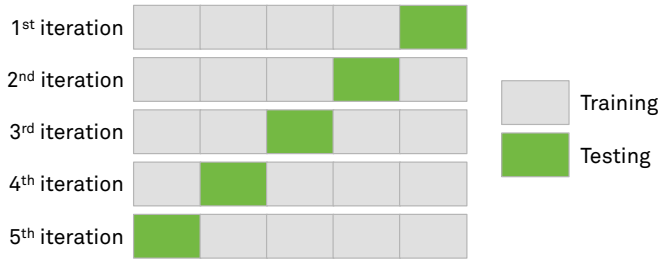

Figure 3.4: Illustration of a 5-fold cross validation.

A comprehensive overview regarding statistical learning is given in e. g. Hastie et al., 2009 and Efron and Hastie, 2016.

4 Follow-up Studies of High-Confidence AGN Candidates in the 2FGL Catalog

We demand rigidly defined areas of doubt and uncertainty!

— Douglas Adams, *The Hitchhiker's Guide to the Galaxy*

The second *Fermi*-LAT (2FGL) catalog (Nolan et al., 2012) covers 1873 gamma-ray sources, detected in the energy regime between 100 MeV and 100 GeV with the Large Area Telescope (LAT) on board the *Fermi* satellite during the first 24 months of operation (see also subsection 5.1.1 for more details regarding *Fermi*, the catalog creation, etc.). An association with source classes, such as pulsars, supernova remnants, BL Lac objects, flat-spectrum radio quasars or radio galaxies, has been performed using different approaches and is included in the catalog. Nonetheless, 575 sources remain unassociated, the so-called Unassociated Gamma-ray Sources (UGS).
The gamma-ray sky, as seen e. g. by *Fermi*-LAT, is a superposition of resolved sources, diffuse Galactic emission, assigned to cosmic-ray interactions with the interstellar medium, and residual diffuse emission, referred to as the isotropic diffuse gamma-ray background (Ackermann et al., 2015a). The origin of the isotropic gamma-ray background is commonly attributed to unresolved sources, such as faint blazars, star-forming galaxies, or radio galaxies, and diffuse processes, such as dark matter annihilation or decay, or interactions of ultra high-energy cosmic rays with the extragalactic background light. More sensitive strategies to detect and resolve individual sources, but also the association of the UGS will allow to infer more advanced conclusions regarding the isotropic gamma-ray background, and correspondingly regarding the extragalactic gamma-ray background, important to derive upper limits on diffuse processes. The UGS could also constitute, at least partially, new source populations, and their association will extend the number of objects for the conduction of population and evolution studies. In particular, the study of blazar populations is of interest, e. g. to probe the so-called blazar sequence. The blazar sequence contains spectral energy distributions, averaged over multiple blazars in bins of their radio luminosity (Fossati et al., 1998), or their gamma-ray luminosity (Ghisellini et al., 2017). Studying this sequence, it can i. a. be inferred that for increasing luminosities, the gamma-ray slope becomes softer, while the X-ray slope gets harder, and the peak frequencies decrease. For more conclusions and details regarding the blazar sequence, the reader is referred to section 2.1.
Most of the Active Galactic Nuclei (AGNs) in the 2FGL catalog are blazars, although generally blazars constitute only a small subclass of AGNs, since the most energetic gamma rays are expected to be produced in the jet, which is directed towards the observer in case of blazars. The exploration of blazars was initiated in the mid-1970s, when the term *blazar* was created to summarize rapidly variable extragalactic objects like BL Lacs and variable quasars (Marscher, 2016). The first blazar samples (E. Massaro et al., 2009) were based on large radio and X-ray surveys, and comprised only a small number of objects. Correspondingly, the necessity of an optical identification for the classification of these objects was viable. Over the years, the sample sizes increased, and sufficient observation time was practically not grantable to spectroscopically analyze all candidate sources with optical telescopes.

As a result, a preselection based on multi-wavelength data became more and more important to reduce the number of candidates for optical spectroscopy.
Extensive work regarding the unveiling of the nature of the UGS in the 2FGL catalog has been conducted by F. Massaro et al., considering different wavelengths, and following multiple approaches. In F. Massaro et al., 2012 the so-called *WISE* blazar strip has been introduced. This strip is a distinct region in the color-color space, measured by the Wide-field Infrared Survey Explorer (*WISE*), comprising predominantly blazars. This is related to the finding that the infrared spectrum of blazars is dominated by non-thermal emission in contrast to sources dominated by thermal emission, such as stars. In addition, blazars detected in gamma rays populate an even narrower region – the *WISE* gamma-ray blazar strip. D'Abrusco et al., 2013 refined this approach by using the three-dimensional color space, and by applying a principal component analysis to find the low-energy counterparts of the UGS. The obtained results have been correlated in F. Massaro et al., 2013c to further radio, infrared, optical and X-ray counterparts by large literature searches. They also showed that all their candidates with an available optical spectrum (despite one) exhibit blazar-like features in the optical spectrum, confirming the validity of their method. A similar method as for the *WISE* blazar strip has been adapted in F. Massaro et al., 2013b: The radio survey at 325 MHz with the Westerbork Synthesis Radio Telescope has been employed to study the properties of known blazars at low radio frequencies. Subsequently, sources with similar properties in other radio catalogs have been selected, and merged with infrared and optical counterparts from archival data. In this way, it has been shown that low radio surveys can be used to specify blazar candidates as well, and not only high radio frequencies, as it was the case until that time. Beyond, it revealed the flat radio spectrum of blazars at 325 MHz. Paggi et al., 2013 analyzed X-ray observations conducted with *Swift*, covering the positional uncertainty region of the UGS. It turned out that around 70 % of the UGS have at least a single X-ray counterpart candidate. Within the X-ray positional uncertainty region, further counterparts have been searched for in major radio, infrared, optical and ultraviolet surveys. Coincidences of blazar-like counterpart candidates with the before-mentioned approaches have been found as well. However, it has also been implied that possibly not all blazars feature X-ray emission, since often no X-ray candidates have been found within the UGS' uncertainty region, despite a sufficient observation time.
Beyond that, more statistical strategies have been pursued by e. g. Ackermann et al., 2012a and Mirabal et al., 2012. Ackermann et al., 2012a applied logistic regression and boosted decision trees from the field of data mining to UGS of the 1FGL catalog to classify between AGNs and pulsars. This strategy has been evaluated in various ways: The data mining models have been built on subsets of the associated gamma-ray sources, and have been applied to the remaining associated sources. That way, the results of the method and the association in the catalog could be compared, and performance measures were computed, proving the validity. Besides, the results have been compared with those of physical approaches, such as the studies by Massaro et al., and several matches have been found. Analogously, a classification between AGNs and pulsars has been conducted by Mirabal et al., 2012, but the 2FGL catalog and a Random Forest classifier were used instead. Again, the accuracy of the method has been demonstrated by a validation of the associated sample.
Likewise, Doert and Errando, 2014 classified the UGS of the 2FGL catalog with a Random Forest classification, and proposed high-confidence AGN candidates. Follow-up studies of these candidates have been conducted within the framework of this thesis. While a radio/gamma-ray connection has been established by the observation that all 2FGL blazars exhibit a radio counterpart, the X-ray/gamma-ray connection is still unclear, as outlined earlier. Thus, it is reasonable to begin the follow-up studies with X-ray observations.

As part of a program for Project-Related Personal Exchange of the German Academic Exchange Service, a cooperation with M. Doert and M. Errando has been created in 2013. During the exchange, an online analysis of *Swift*-XRT observations has been conducted in collaboration with A. Overkemping. Based on the results obtained, a proposal (lead by M. Errando) to observe 24 UGS with *Swift*-XRT in 2013/2014 has been submitted and accepted (Errando et al., 2013). These sources have been rated as high-confidence AGN candidates, and have not been previously observed with *Swift*. The corresponding offline analysis[1] of the *Swift*-XRT observations and its results are described in the following. Further follow-up studies comprise the correlation with other wavelengths, and with work presented in the literature mentioned above.

4.1 High-Confidence AGN Candidates in the 2FGL Catalog

A search for gamma-ray emitting AGNs in the sample of UGS of the 2FGL catalog using machine learning is presented in Doert and Errando, 2014. In the corresponding analysis, supervised multivariate classification methods have been applied, requiring a data sample with class labels. For this purpose, the associated gamma-ray sources available in the 2FGL catalog have been utilized. This sample contains 1092 AGNs, such as BL Lac objects or flat-spectrum radio quasars, and 205 non-AGNs, such as pulsars or supernova remnants. For every object in the 2FGL catalog, information regarding i. a. the fluxes in different energy ranges, the shape of the spectrum, and the flux variability is available. Based on these features, a feature generation and a feature selection have been performed. So-called hardness ratios HR_{ij} between two energy bands E_i, E_j and their respective fluxes F_i, F_j have been generated as follows:

$$\mathrm{HR}_{ij} = \frac{F_i E_i - F_j E_j}{F_i E_i + F_j E_j} \,. \tag{4.1}$$

Proceeding from these ratios, a `hardness slope` has been derived:

$$\texttt{hardness slope} = \mathrm{HR}_{23} - \mathrm{HR}_{45} \,. \tag{4.2}$$

The following features have been selected: `spectral index`, `hardness slope`, `normalized variability`, HR_{12}, HR_{23}, HR_{34}, and HR_{45} (for a more detailed description of the features the reader is referred to the corresponding paper). In general, the use of features correlated to the brightness of a source has been avoided, since the associated sample tends to be brighter compared to the UGS sample, and it is crucial that the classification model performs similarly for both samples. Only one feature (`normalized variability`) is dependent on the significance of the source.
The associated sample has been divided into two sets – one set to build and optimize the models, and a further set to evaluate the performance. The discrimination between AGN and non-AGN, as defined above, is called a two class problem. For this classification, a Random Forest classifier and a neural network have been applied. Both methods have been optimized separately, first regarding their hyperparameters, and secondly, concerning their thresholds of the classifier output to assign the AGN class to a source. For the final classification, a combination of both classifiers has been conducted. The performance has been evaluated in terms of the recall – the fraction of correctly classified AGNs – and the false-association rate – the ratio between sources falsely classified as AGN and all sources classified as AGN.

[1] The analysis has been conducted in 2013/2014 and does not consider more recent observations or literature.

In addition, the false-association rate has been re-weighted according to the distributions of the significance and the Galactic latitude of the UGS to take into account that the UGS sample tends to contain fainter sources and more objects in the Galactic plane as the associated sample.
Within the performance evaluation, the method combining the Random Forest classifier and the neural network obtained a recall of 70.1% and a false-association rate of 11.2%. The application of the combined model to the sample of UGS led to 231 AGN candidates of high confidence. Comparing these candidates with the similar approach by Mirabal et al., 2012, 156 of 159 AGN candidates outside the Galactic plane ($|b| \geq 10^{\circ}$) coincide. Further coincidences with the AGN candidates and corresponding counterparts of other frequencies have also been found in searches by F. Massaro et al., 2013a, Paggi et al., 2013, and Acero et al., 2013. The AGN candidates together with the found coincidences with other studies are listed in Table 4.3. Still, follow-up observations are necessary to finally unveil the nature of the UGS. A proficient strategy is to start with X-ray observations, exhibiting typically high angular resolutions and small numbers of potential counterpart candidates, and to proceed with dedicated radio and optical spectroscopy to determine the source and spectral class for the X-ray selected AGN candidates.

4.2 *Swift*-XRT Analysis

The X-Ray Telescope (XRT) aboard the *Swift* satellite is measuring X-ray photons in the energy regime of $0.2 - 10$ keV (see subsection 2.4.2 for more details). The following analysis is based on observations of the 2FGL uncertainty regions of the 231 AGN candidates presented previously, performed before September 2014 with the *Swift*-XRT. Due to its field of view of $23.6' \times 23.6'$ and a position accuracy of a few arcseconds, it is not required that the XRT points directly towards the 2FGL position. Nonetheless, the uncertainty region of the 2FGL position of about a few arcminutes was covered for most of the observations.
The corresponding data files[2] have been downloaded from the High Energy Astrophysics Science Archive Research Center (HEASARC, `https://heasarc.gsfc.nasa.gov/`). The analysis is conducted with *HEASoft* (version 6.16, `https://heasarc.nasa.gov/lheasoft/`), a software package containing all software distributed from the HEASARC, together with the calibration database *CALDB* (version from January 20th 2014 for the XRT analysis).
The common initial step in the analysis is the reduction of Level 1 data, which has been executed with the script *xrtpipeline* (version 0.13.1, Capalbi et al., 2005). For the production of Level 1 event files, the recorded data are reformatted and ancillary information is added. During the reduction, the data are prepared for further analysis, i. e. the data are calibrated, screened and filtered, and Level 2 event files are produced. The screening process produces a list of Good Time Intervals (GTI), which is used for the extraction of the events, and the filtering selects events according to their position, time or energy. For the selection of observations to be considered, the observation mode has been set to select only pointed observations and the data mode has been set to the photon counting mode. Only one screening criterion has been adapted compared to the standard *xrtpipeline* settings: The upper temperature limit of the CCDs has been set from $-47\,^{\circ}$C to $-50\,^{\circ}$C. Due to the temperature dependency of dark currents and hot pixels, the low-energy background is reduced when lowering the temperature, increasing the signal-to-noise ratio (D'Elia, V. et al., 2013), which is crucial for faint sources.

[2] This research has made use of data obtained through the High Energy Astrophysics Science Archive Research Center Online Service, provided by the NASA/Goddard Space Flight Center.

The command line interface *Xselect* (version 2.4c, Ingham et al., 2006) can be utilized for multiple missions to perform X-ray analyses, such as the extraction of images, spectra, or light curves for different event selection criteria. In the beginning of a session, the corresponding mission (here: *Swift*) has to be set, before reading and merging event files of Level 2 from individual observations. Based on these merged event lists, selection criteria concerning the energy regime have been applied, and images containing X-ray photons in the energy regime of $0.3 - 10$ keV have been extracted.
The multi-mission X-ray display and analysis program *XIMAGE* (version 4.5.1, Angelini and Padgett, 2010) is used to inspect the extracted images. Via the command `detect`, point sources are localized on an image with a sliding-cell algorithm: The mean background intensity is estimated by multiple small square cells distributed over the image. Each of these cells is checked for consistency with the statistical Poisson expectation, and is rejected in case of inconsistency. Cells with intensity values deviating more than 3σ from the mean of the background values are also rejected. Thereafter, the final background intensity is the average over all non-rejected cells. Subsequently, the signal-to-noise ratio is calculated for different positions and cell sizes. Cells maximizing the signal-to-noise ratio are selected, and if this ratio exceeds a specific threshold, a source is detected at the considered position. For the presented analysis, a threshold of 2σ has been chosen. In addition, several corrections are applied, e. g. regarding dead times, vignetting and the point spread function. The output of this command provides information regarding the detected point sources, such as the position, the count rate, the signal-to-noise ratio, and the size of the used cell, but also more general information about the exposure, or the time range of the observations.

4.3 *Swift*-XRT Results and Counterpart Coincidences

A total number of 151 out of 231 high-confidence AGN candidates have been observed with the *Swift*-XRT, in accordance with the following criteria: The minimum observation time amounts to more than 1000 s, and the complete 95% uncertainty region of the 2FGL sources is covered by the *Swift* field of view. For every source, an image section of the size $0.4^\circ \times 0.4^\circ$ has been cropped. The detected X-ray point sources are denoted by boxes on the images. Their sizes correlate to the size of the cell, used to determine i. a. the signal-to-noise ratio, the position and the count rate. The corresponding signal-to-noise ratio is stated next to the box. The 95% uncertainty region provided by the 2FGL catalog is illustrated to evaluate the number of XRT point sources within this region, and accordingly, the number and position of X-ray candidates to be considered as counterparts.
The obtained X-ray counterpart candidates have been compared to the candidates in F. Massaro et al., 2013c, which are based on a distinct region in the infrared color-color space – the WISE blazar strip – and an extensive literature search. All WISE point sources within the 2FGL uncertainty region, and likewise located within the WISE blazar strip, are considered as WISE counterpart candidates. For some AGN candidates, no WISE candidates have been found, while for others, multiple candidates have been found. According to the location in the WISE blazar strip, a blazar type (BZB = BL Lac object, BZQ = flat-spectrum radio quasar, UND = undetermined) has been associated. Within a region of a few arcseconds around the position of the obtained WISE candidates, a search in major radio, infrared, optical and X-ray surveys has been conducted to link further counterpart candidates, and to support the association with a blazar. The candidates from WISE and other wavelengths are designated in the images as well. The described images are presented in Figure 4.1, together with additional information, which is also listed in Table 4.3.

The numbers of detected *Swift*-XRT point sources per 95% uncertainty region of the 2FGL source are summarized in Table 4.1. While 55 AGN candidates feature a single X-ray point source, 57 AGN candidates exhibit multiple possible X-ray counterparts. Nevertheless, for 39 AGN candidates no X-ray counterpart have been detected. This could be related to an insufficient exposure to detect faint sources, although 35 out of these have been observed for more than 3000 s, supporting the assertion that possibly not all blazars feature X-ray emission (Paggi et al., 2013). The fraction of high-confidence AGN candidates with at least a single X-ray counterpart candidate of 74% is similar to the one for the UGS obtained in Paggi et al., 2013 of about 70%.

Table 4.1: Number of 2FGL sources dependent on the number of detected XRT point sources in the corresponding 95% uncertainty region. A total number of 151 out of 231 high-confidence AGN candidates have been observed by the *Swift*-XRT.

# XRT Sources / 2FGL region	# 2FGL Sources
0	39
1	55
> 1	57
Total observed sources	151

The numbers of coincidences between different approaches to obtain counterpart candidates for the 151 AGN candidates are listed in Table 4.2. At least one WISE counterpart candidate is contained in the 2FGL uncertainty region of 65 sources. Requiring an additional coincident counterpart, the number of sources reduces to 37. The combination of the WISE counterparts and the XRT counterpart candidates results in a number of 27 2FGL sources with coincidences. Adding the constraint of a further counterpart, 20 AGN candidates exhibit coincidences.

Table 4.2: Number of 2FGL sources with coincidences between different studies and wavelengths. The high-confidence AGN candidates from Doert and Errando, 2014 are denoted as 2FGL, and XRT denotes the X-ray point sources presented in this section. The WISE counterpart candidates, together with additional other counterparts, originate from F. Massaro et al., 2013c.

Coincidence	# 2FGL Sources
WISE / 2FGL	65
Other / WISE / 2FGL	37
XRT / WISE / 2FGL	27
Other / XRT / WISE / 2FGL	20

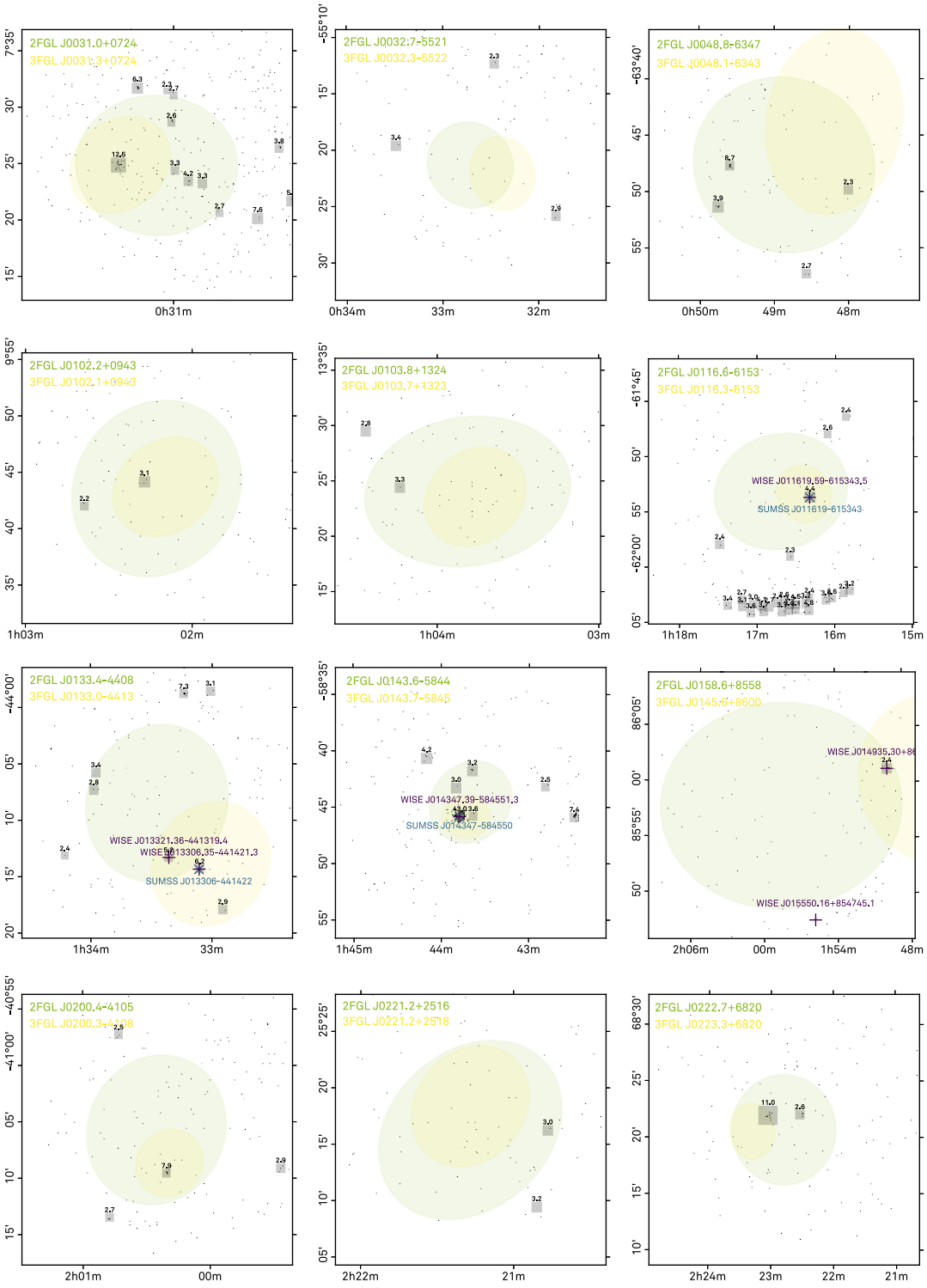

Figure 4.1: *Swift*-XRT image sections (0.4° × 0.4°) of the 151 observed high-confidence AGN candidates. The detected XRT point sources are visualized by gray boxes, and the corresponding number indicates the signal-to-noise ratio. The ellipses represent the 95% uncertainty regions provided by the 2FGL (green) and the 3FGL (yellow) catalog. Counterpart candidates determined by F. Massaro et al., 2013c are specified in purple (WISE candidates) and blue (other candidates).

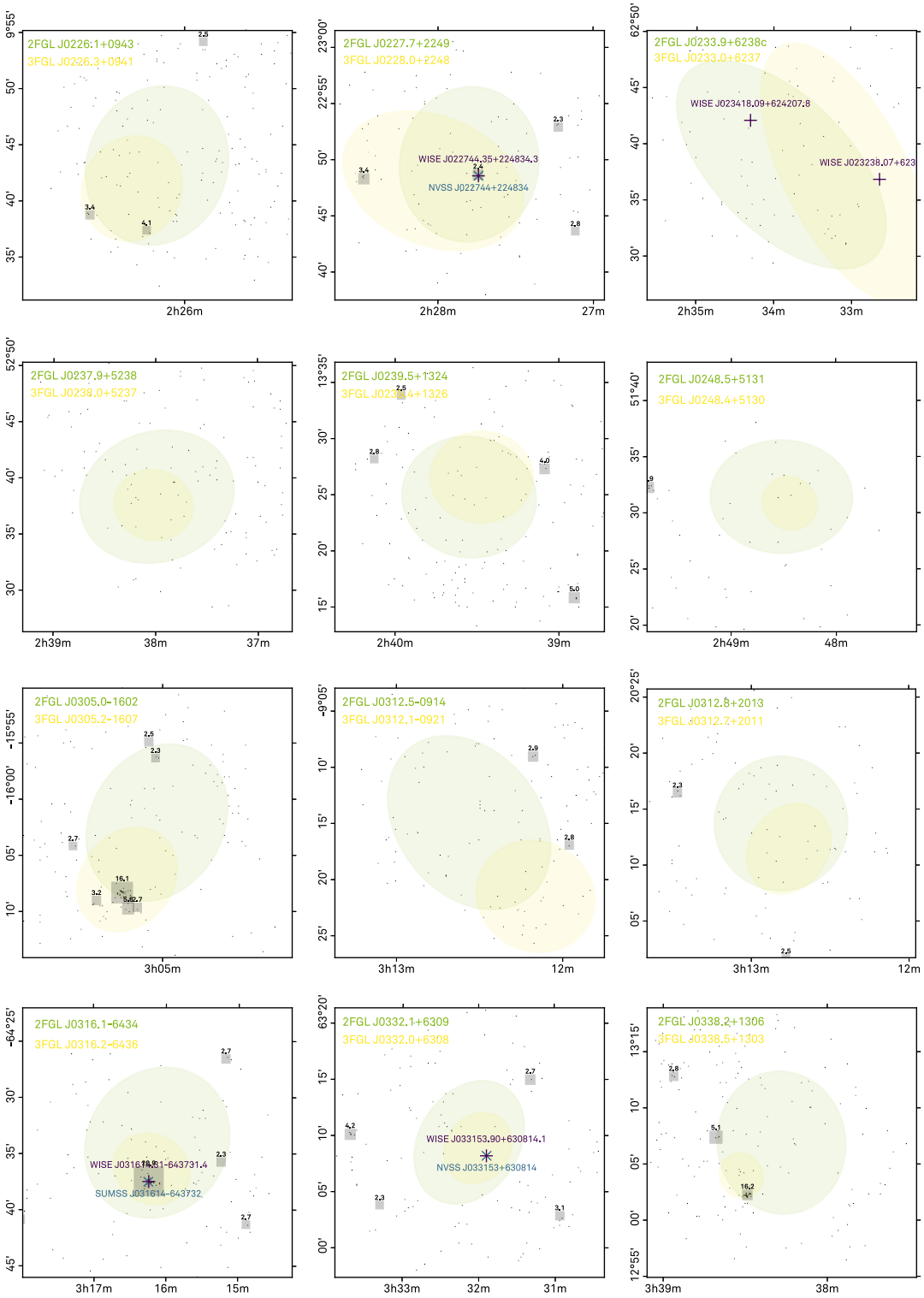

Figure 4.1: – continued from previous page

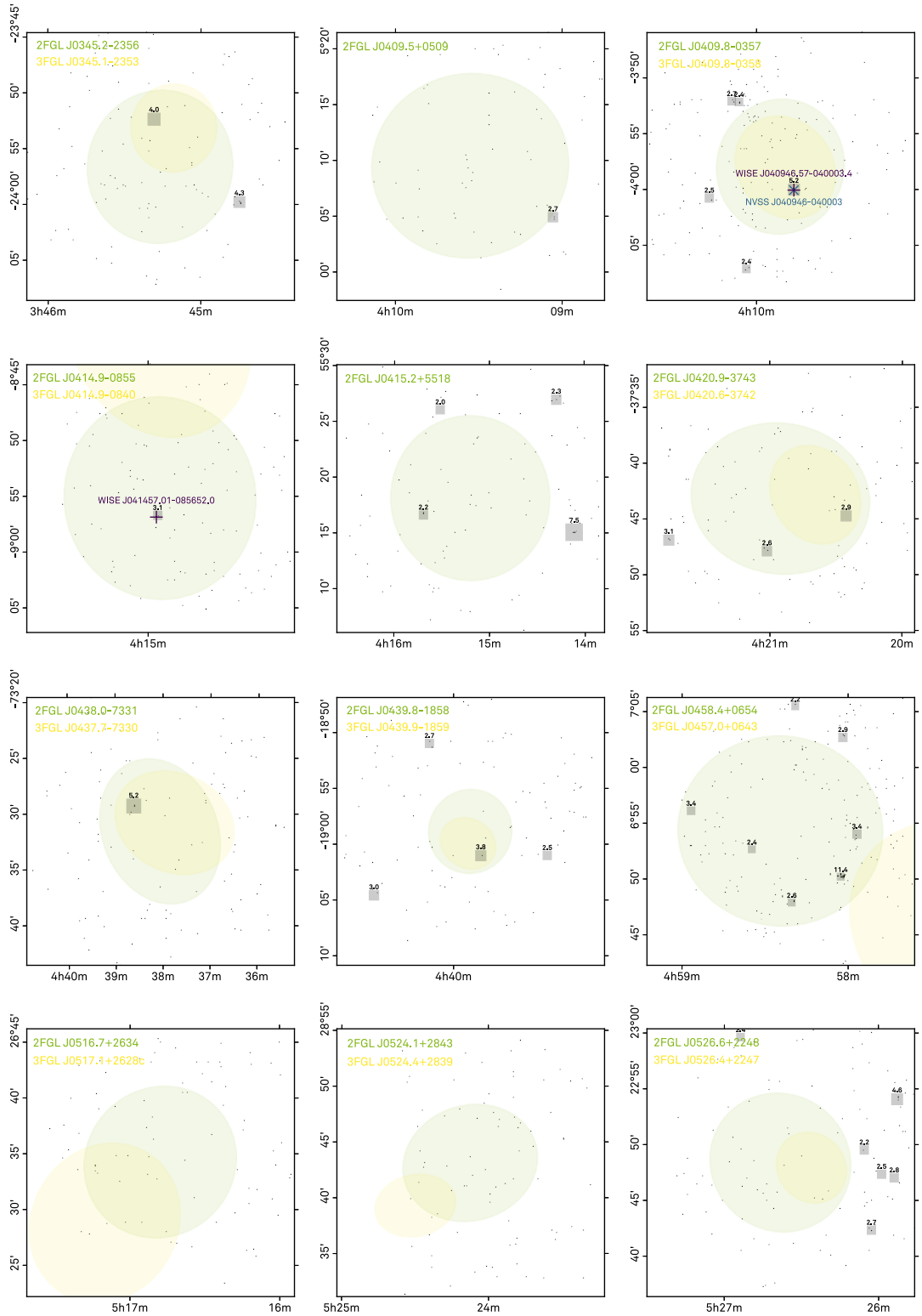

Figure 4.1: – continued from previous page

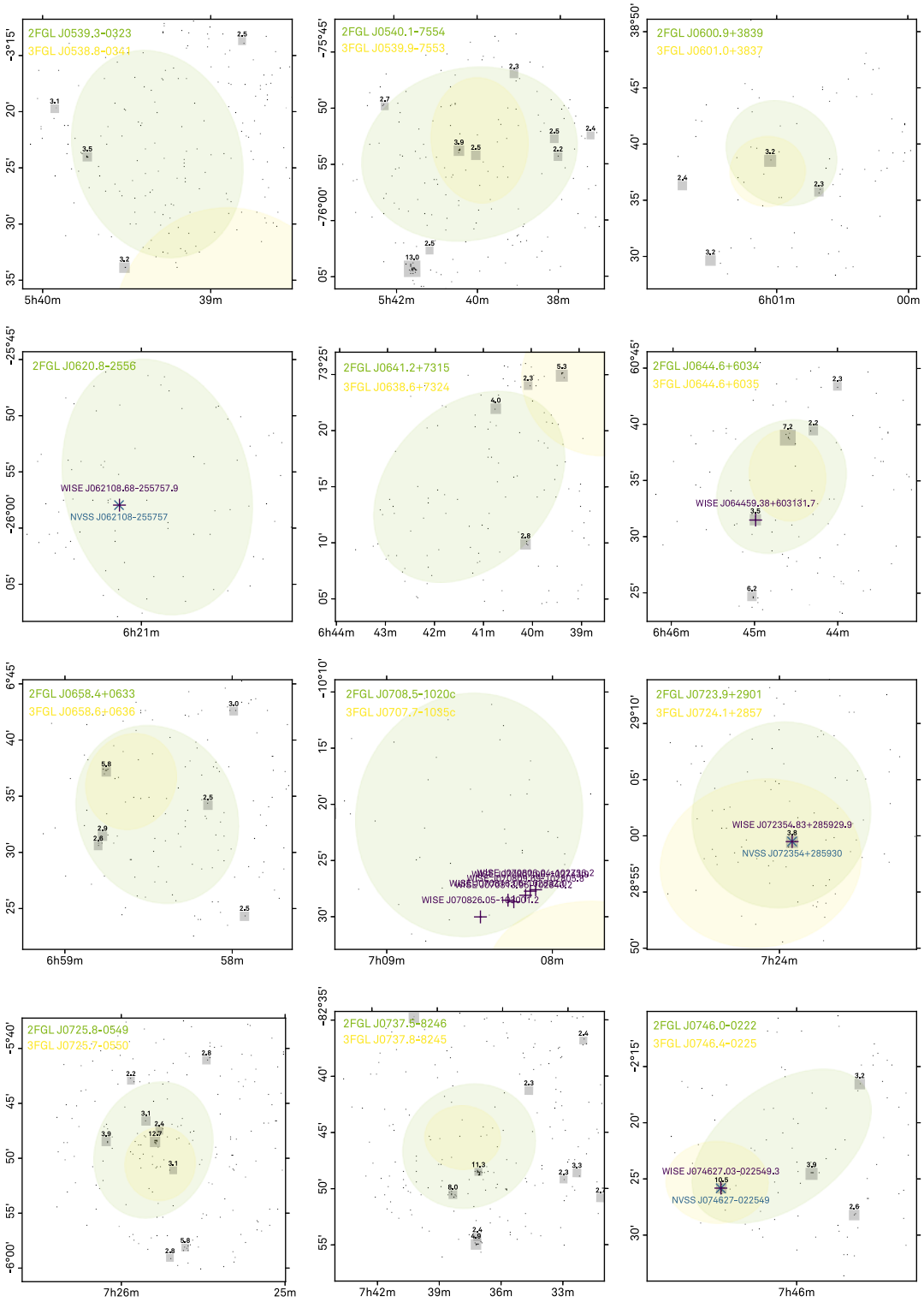

Figure 4.1: – continued from previous page

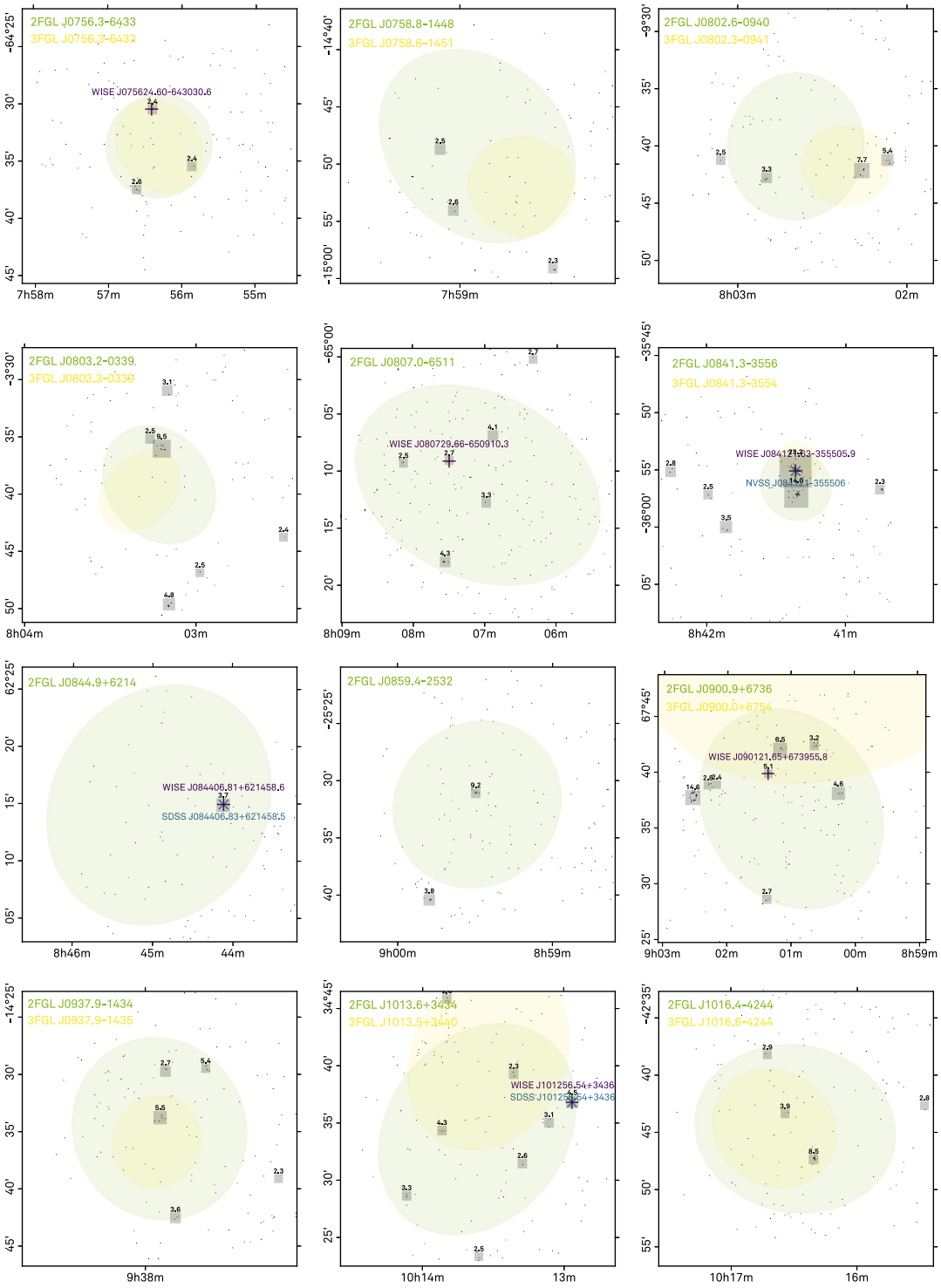

Figure 4.1: – continued from previous page

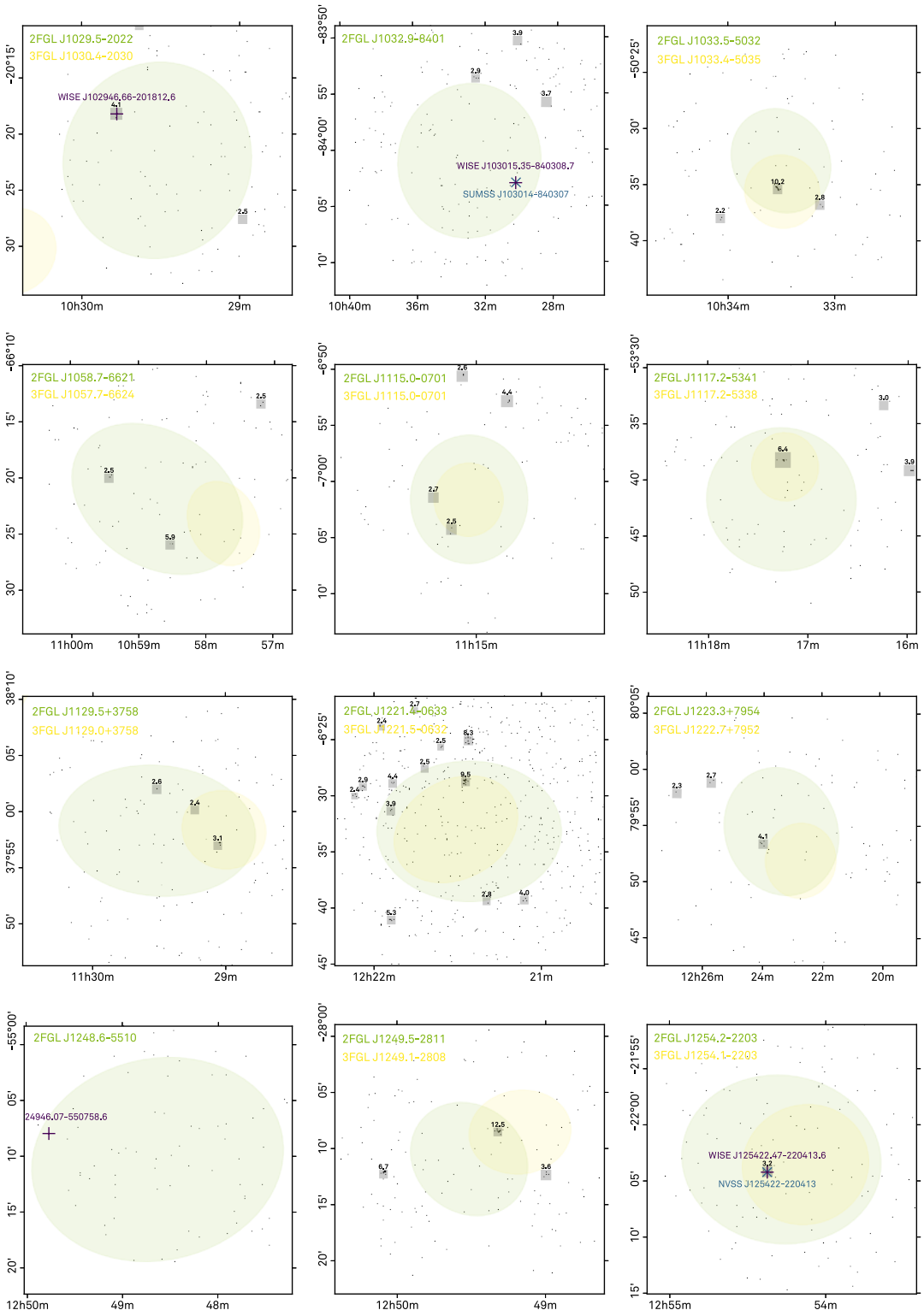

Figure 4.1: – continued from previous page

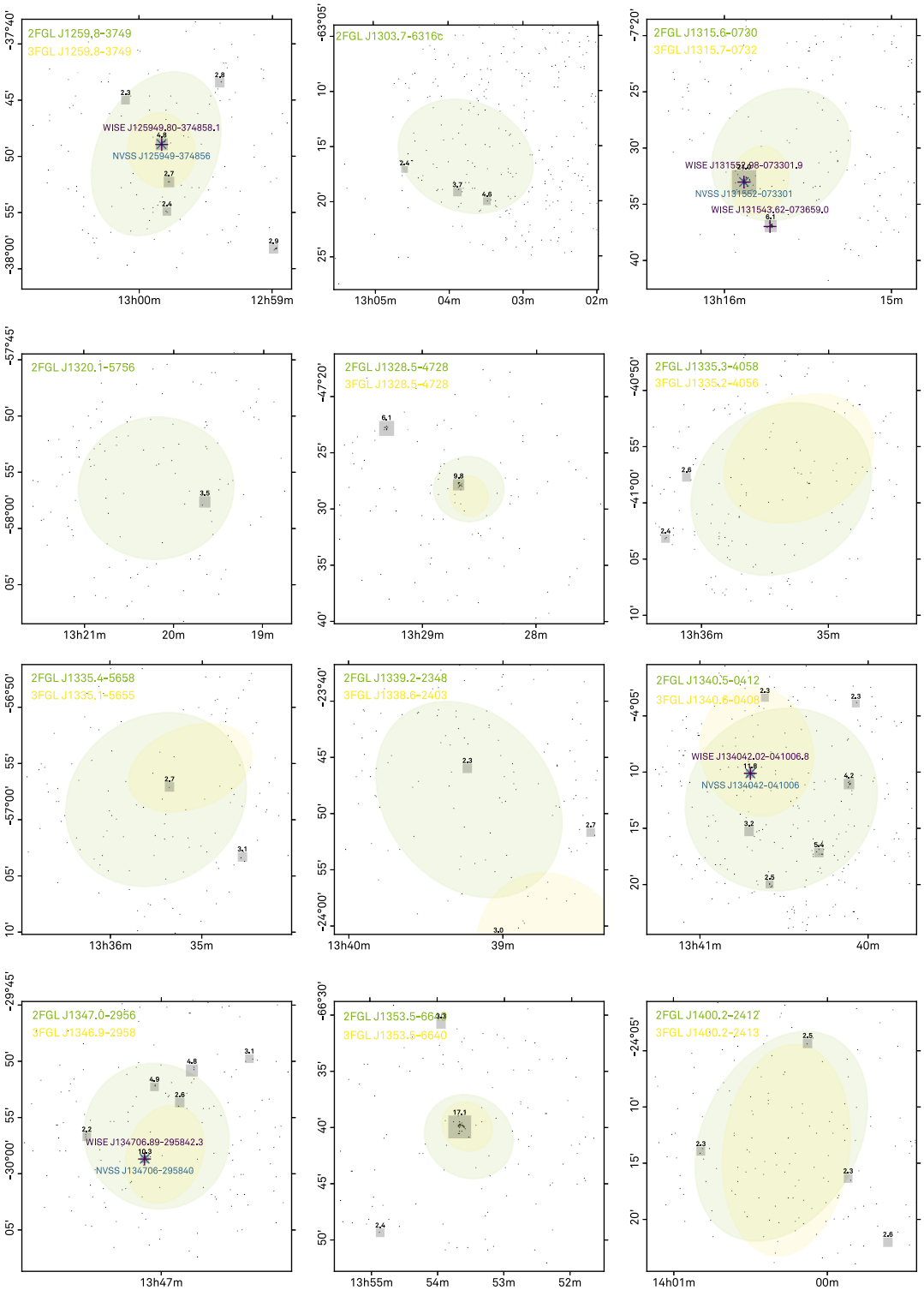

Figure 4.1: – continued from previous page

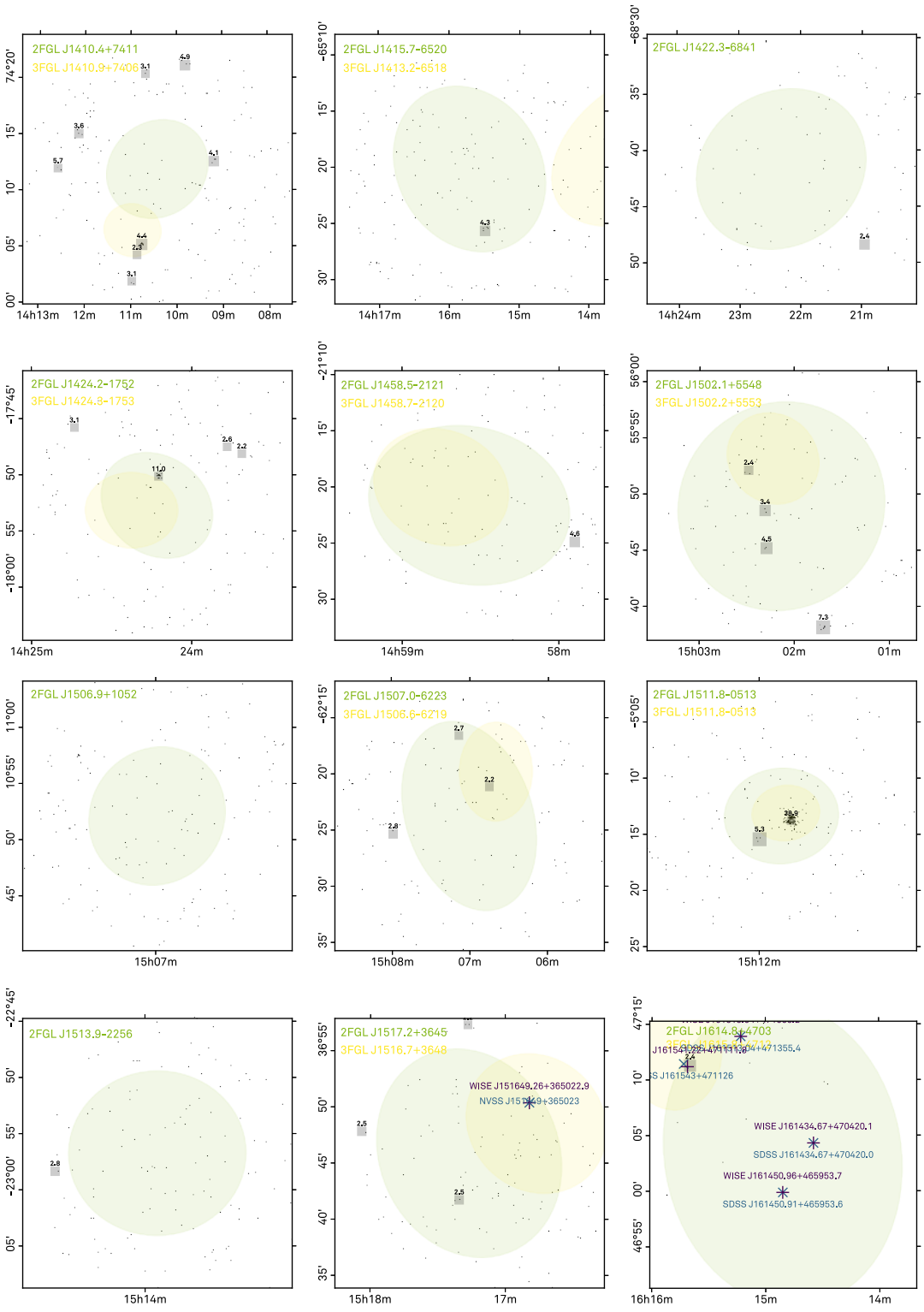

Figure 4.1: – continued from previous page

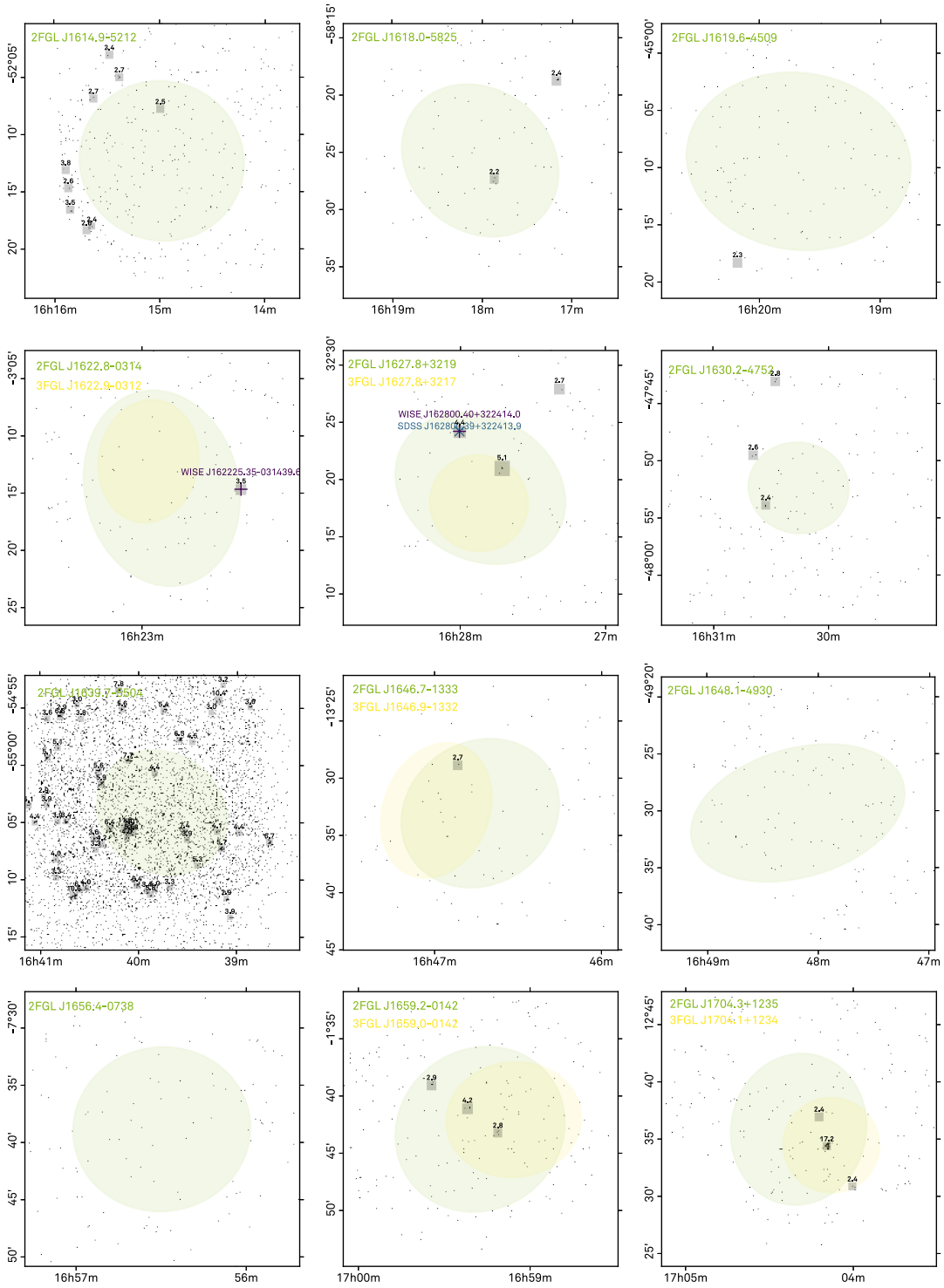

Figure 4.1: – continued from previous page

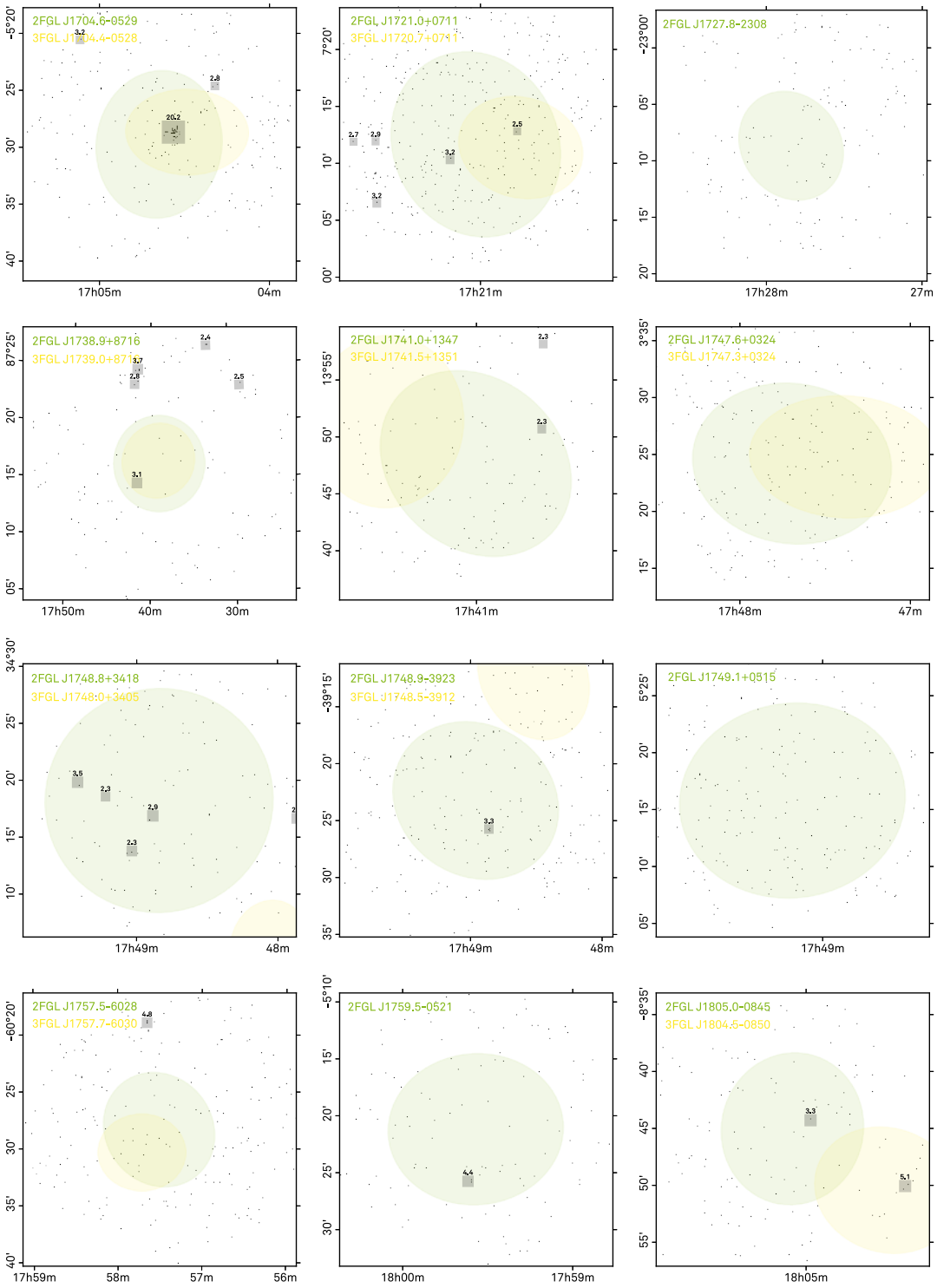

Figure 4.1: – continued from previous page

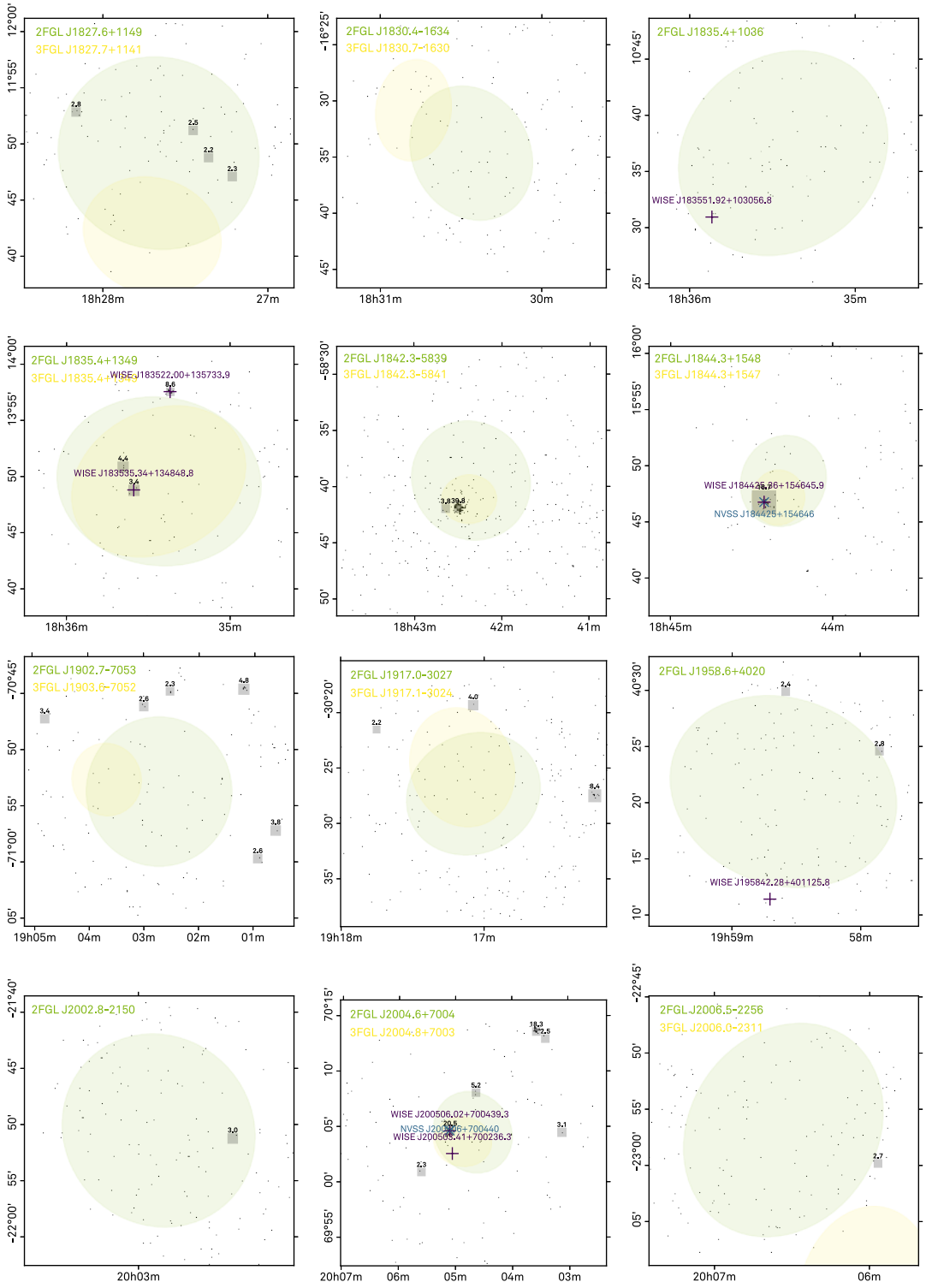

Figure 4.1: – continued from previous page

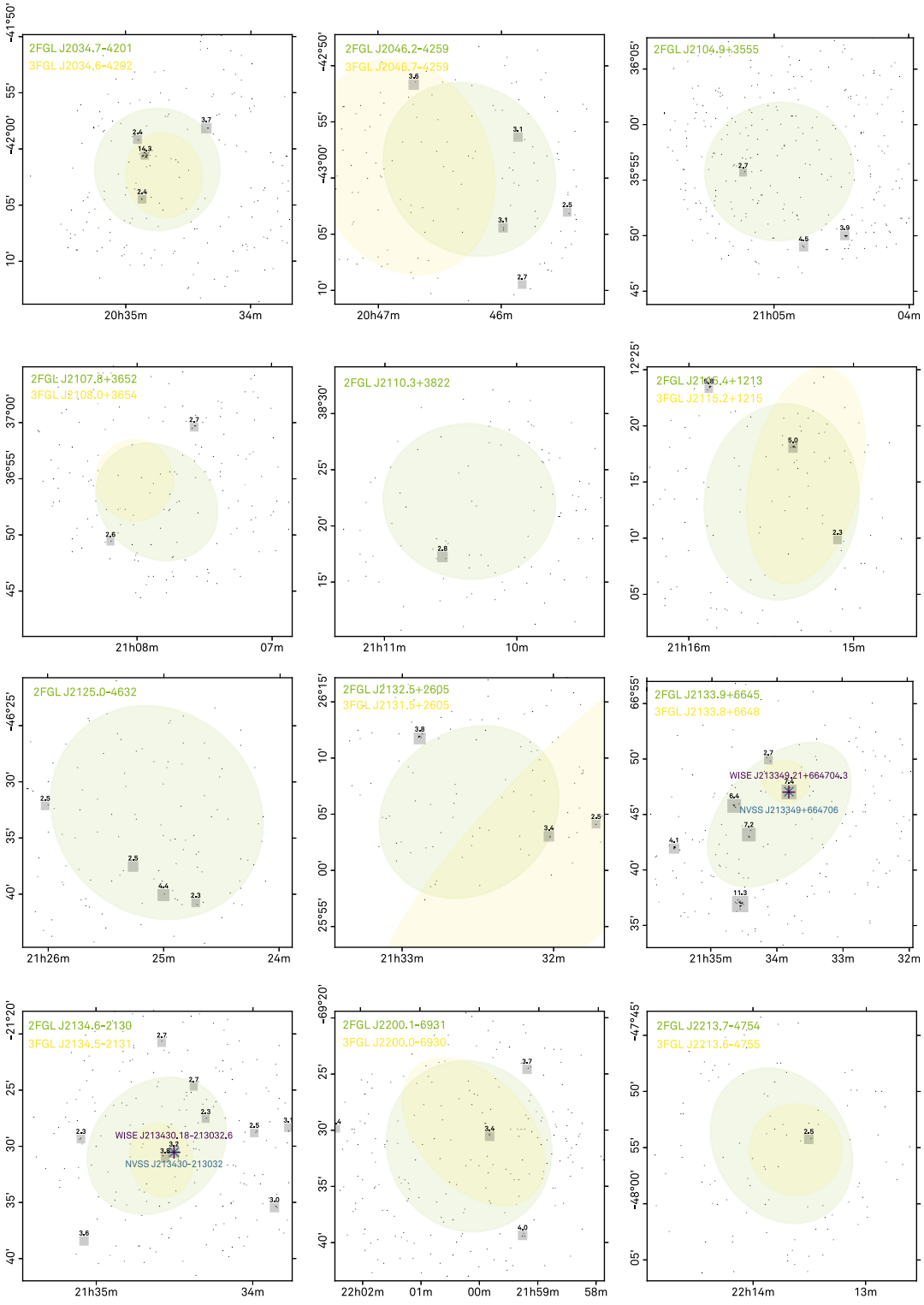

Figure 4.1: – continued from previous page

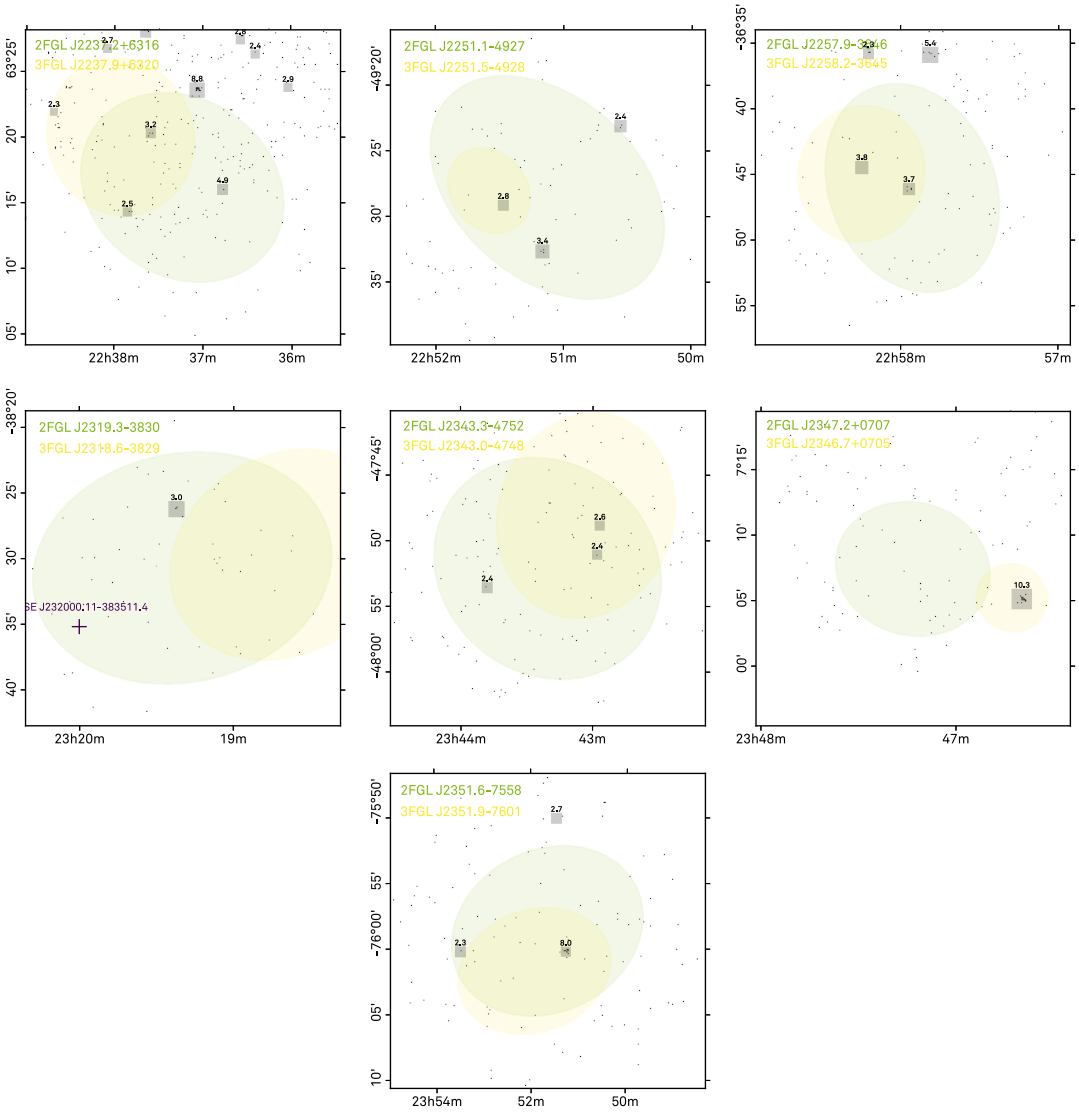

Figure 4.1: – continued from previous page

Table 4.3: Comprehensive list of 231 high-confidence AGN candidates. The columns LRF and LNN indicate the scores of the Random Forest classification and the neural network from the machine learning approach (Doert and Errando, 2014) to determine these candidates. F. Massaro et al., 2013c made use of the WISE blazar strip to associate blazar types (BZB = BL Lac object, BZQ = flat-spectrum radio quasar, UND = undetermined) and WISE counterpart candidates to 2FGL sources. In addition, further counterparts from other catalogs and surveys have been linked to the WISE candidates. The results of the *Swift*-XRT analysis presented in this chapter are listed as well: The exposure of the observations of 151 AGN candidates, and the number of detected XRT point sources in the 95% uncertainty region of the 2FGL source. The coincidence between an XRT point source and a WISE counterpart candidate is indicated by a check mark.

2FGL Name	LRF	LNN	Type	WISE Name	Other Name	XRT Exposure / s	#XRT Sources / 2FGL Region	Coincidence XRT/WISE
J0004.2+2208	0.98	0.97						
J0014.3−0509	1.00	1.00						
J0031.0+0724	0.99	1.00				18379	5	
J0032.7−5521	1.00	1.00				3982	0	
J0048.8−6347	0.90	0.92				3686	3	
J0102.2+0943	1.00	1.00				4032	2	
J0103.8+1324	0.96	1.00				4453	1	
J0116.6−6153	1.00	1.00				3330	1	
J0133.4−4408	0.99	1.00	BZQ	J013321.36−441319.4		4922	3	✓
J0143.6−5844	0.98	1.00	BZB	J014347.39−584551.3	SUMSS J014347−584550	4456	4	✓
J0158.4+0107	0.91	1.00	UND	J015836.25+010632.1	SDSS J015836.23+010632.0			
			UND	J015757.45+011547.8				
			UND	J015910.05+010514.7	SDSS J015910.05+010514.5			
J0158.6+8558	0.98	1.00	BZB	J014935.30+860115.3		2853	1	✓
J0200.4−4105	1.00	1.00				5468	1	
J0221.2+2516	0.98	1.00				3814	1	
J0222.7+6820	0.97	1.00				5213	2	
J0226.1+0943	0.92	1.00				7261	1	
J0227.2+6029c	0.93	1.00						
J0227.7+2249	0.99	1.00	BZB	J022744.35+224834.3	NVSS J022744+224834	3881	1	✓
J0233.9+6238c	0.98	1.00	BZB	J023418.09+624207.8		3204	0	
J0237.9+5238	0.93	0.99				4463	0	
J0239.5+1324	0.99	1.00				4348	0	
J0248.5+5131	1.00	1.00				2417	0	
J0253.9+5908	0.95	1.00						
J0305.0−1602	0.97	1.00				4802	2	
J0307.4+4915	0.99	1.00						
J0312.5−0914	0.97	1.00				5155	0	
J0312.8+2013	1.00	1.00				4112	0	
J0316.1−6434	0.99	1.00	BZB	J031614.31−643731.4	SUMSS J031614−643732	4217	2	✓
J0332.1+6309	0.94	1.00	BZB	J033153.90+630814.1	NVSS J033153+630814	5170	0	
J0338.2+1306	0.99	0.99				7519	1	
J0345.2−2356	0.93	0.97				4242	1	
J0404.0+3843	0.95	0.96						
J0404.6+5822	0.89	0.96						
J0409.5+0509	0.99	1.00				3405	0	
J0409.8−0357	1.00	1.00	BZB	J040946.57−040003.4	NVSS J040946−040003	4997	1	✓
J0414.9−0855	0.97	1.00	UND	J041457.01−085652.0	APMUKS 041232.66−090420.3	4265	1	✓
J0415.2+5518	0.94	0.98				3585	1	
J0416.0−4355	0.98	0.96	BZQ	J041605.81−435514.6	SUMSS J041605−435516			
J0420.9−3743	1.00	1.00				4378	2	
J0428.0−3845	0.97	0.97						
J0438.0−7331	0.96	1.00				3921	1	
J0439.8−1858	1.00	1.00				3992	1	
J0458.4+0654	1.00	1.00				6622	5	
J0515.0−4411	0.91	1.00						
J0516.7+2634	0.93	0.94				4651	0	
J0524.1+2843	0.97	1.00				3713	0	
J0526.6+2248	0.98	0.97				4077	0	
J0534.8−0548c	0.92	0.97						
J0539.3−0323	0.94	0.93				6682	1	
J0540.1−7554	0.97	0.99				4621	5	
J0555.9−4348	0.91	1.00	BZQ	J055531.59−435030.7				
			BZB	J055618.74−435146.1	SUMSS J055618−435146			
J0600.9+3839	0.97	0.96				4435	2	
J0602.7−4011	0.94	1.00	UND	J060237.10−401453.2				
J0620.8−2556	0.98	1.00	?	J062108.68−255757.9	NVSS J062108−255757	2944	0	
J0641.2+7315	0.98	1.00				5015	1	
J0644.6+6034	0.92	1.00	UND	J064459.38+603131.7		3260	3	✓
J0658.4+0633	0.90	0.93				4408	4	
J0708.5−1020c	0.89	0.98	BZB	J070807.98−102743.9		1951	0	
			BZQ	J070809.69−102805.8				
			BZB	J070806.04−102736.2				
			BZB	J070813.96−102840.2				
			BZQ	J070826.05−103001.2				
			BZB	J070816.06−102832.0				

Table 4.3: – continued from previous page

2FGL Name	LRF	LNN	Type	WISE Name	Other Name	XRT Exposure / s	#XRT Sources / 2FGL Region	Coincidence XRT/WISE
J0723.9+2901	1.00	1.00	BZQ	J072354.83+285929.9	NVSS J072354+285930	3826	1	✓
J0725.8–0549	1.00	1.00				4909	5	
J0737.5–8246	0.92	1.00				7938	2	
J0746.0–0222	1.00	1.00	BZB	J074627.03–022549.3	NVSS J074627–022549	4250	3	✓
J0753.2+1937	0.99	1.00						
J0756.3–6433	1.00	1.00	BZB	J075624.60–643030.6		4669	3	✓
J0758.8–1448	0.97	0.98				4368	2	
J0802.6–0940	0.97	1.00				4315	1	
J0803.2–0339	0.95	1.00				3964	2	
J0807.0–6511	0.87	1.00	BZQ	J080729.66–650910.3		8700	5	✓
J0841.3–3556	0.91	1.00	BZB	J084121.63–355505.9	NVSS J084121–355506	3751	2	✓
J0844.9+6214	0.90	0.99	BZB	J084406.81+621458.6	SDSS J084406.83+621458.5	2547	1	✓
J0846.0+2820	0.98	1.00						
J0859.4–2532	1.00	1.00				3392	1	
J0900.9+6736	1.00	1.00	UND	J090121.65+673955.8		7890	6	✓
J0902.8–4741c	0.93	0.97						
J0923.5+1508	1.00	1.00						
J0928.8–3530	0.96	0.92						
J0937.9–1434	0.99	1.00				4506	4	
J1013.6+3434	0.95	1.00	BZB	J101256.54+343648.8	SDSS J101256.54+343648.7	4681	6	✓
J1016.1+5600	0.99	0.99	BZQ	J101544.44+555100.7	NVSS J101544+555100			
J1016.4–4244	0.98	1.00				3966	3	
J1029.5–2022	1.00	1.00				4478	1	
J1032.9–8401	1.00	1.00	BZB	J103015.35–840308.7	SUMSS J103014–840307	8031	0	
J1033.5–5032	0.99	1.00				3874	1	
J1036.4–5828c	0.96	1.00						
J1038.6–5850c	0.98	1.00						
J1058.7–6621	0.95	0.99				3676	2	
J1059.3–6118c	0.97	0.99						
J1059.9–2051	1.00	1.00	BZB	J110025.72–205333.4	2MASX J11002568–2053333			
J1115.0–0701	0.98	1.00				3385	2	
J1117.2–5341	1.00	1.00				3397	1	
J1129.0–0532	0.94	0.99						
J1129.5+3758	0.95	0.99				4869	3	
J1208.6–2257	1.00	1.00	BZB	J120816.33–224921.9				
J1221.4–0633	0.99	0.99				15422	2	
J1223.3+7954	0.98	1.00				4180	1	
J1248.6–5510	1.00	1.00	UND	J124946.07–550758.6		3450	0	
J1249.5–2811	1.00	1.00				4127	1	
J1254.2–2203	0.96	1.00	BZB	J125422.47–220413.6	NVSS J125422–220413	3450	1	✓
J1259.8–3749	1.00	1.00	BZB	J125949.80–374858.1	NVSS J125949–374856	3673	4	✓
J1303.7–6316c	0.90	0.94				10558	3	
J1312.9–2351	1.00	1.00						
J1315.6–0730	1.00	1.00	BZB	J131552.98–073301.9	NVSS J131552–073301	6609	1	✓
J1320.1–5756	0.95	0.93				3530	1	
J1324.4–5411	1.00	1.00	BZQ	J132415.49–541104.4				
J1328.5–4728	1.00	1.00				2969	1	
J1329.5–3448	0.98	1.00						
J1335.3–4058	0.97	1.00				3949	0	
J1335.4–5658	0.86	0.94				4262	1	
J1339.2–2348	0.98	1.00				3212	1	
J1340.5–0412	1.00	1.00	BZB	J134042.02–041006.8	NVSS J134042–041006	7319	5	✓
J1345.8–3356	0.94	1.00	BZB	J134543.05–335643.3	NVSS J134543–335643			
J1347.0–2956	0.99	1.00	BZB	J134706.89–295842.3	NVSS J134706–295840	3610	4	✓
J1351.1–2749	0.99	1.00						
J1353.5–6640	0.95	1.00				3302	1	
J1400.2–2412	0.93	1.00				4443	2	
J1407.6–5937c	0.98	1.00						
J1410.4+7411	0.91	0.98				7279	0	
J1415.7–6520	1.00	1.00				4448	1	
J1417.5–4404	0.96	0.96						
J1417.7–5028	0.96	1.00						
J1422.3–6841	0.97	0.99				3435	0	
J1423.9–7842	0.95	1.00						
J1424.2–1752	0.98	1.00				5200	1	
J1427.6–6048c	0.87	1.00						
J1456.7–6247c	0.97	1.00						
J1458.5–2121	0.95	1.00				3387	0	
J1502.1+5548	0.99	1.00				4079	3	
J1502.4+4804	1.00	1.00						
J1506.9+1052	0.99	0.99				3320	0	
J1507.0–6223	1.00	1.00				3836	0	
J1511.8–0513	0.97	1.00				4282	2	
J1513.9–2256	1.00	1.00				3330	0	
J1517.2+3645	0.97	1.00	BZB	J151649.26+365022.9	NVSS J151649+365023	3064	1	
J1548.3+1453	0.89	1.00						
J1553.5–0324	0.97	1.00						
J1601.1–4220	0.92	0.96						
J1614.8+4703	0.99	1.00	BZQ	J161450.96+465953.7	SDSS J161450.91+465953.6	1191	1	
			BZQ	J161513.04+471355.2	SDSS J161513.04+471355.4			
			BZQ	J161434.67+470420.1	SDSS J161434.67+470420.0			
			BZB	J161541.22+471111.8	NVSS J161543+471126			✓

Table 4.3: – continued from previous page

2FGL Name	LRF	LNN	Type	WISE Name	Other Name	XRT Exposure / s	#XRT Sources / 2FGL Region	Coincidence XRT/WISE
J1614.9–5212	0.91	0.94				4232	1	
J1617.6–4219	0.95	1.00						
J1618.0–5825	1.00	1.00				3573	1	
J1619.6–4509	0.93	0.99				3901	0	
J1622.8–0314	0.95	0.96				3899	0	
J1627.8+3219	0.96	0.99				3455	2	
J1630.2–4752	0.96	0.96				4197	1	
J1636.6–0841	0.97	1.00						
J1639.7–5504	0.91	0.99				153249	12	
J1639.8–4921c	1.00	0.98	BZB	J163907.38–492605.6				
J1646.7–1333	0.99	1.00				2768	1	
J1647.0+4351	1.00	1.00	BZB	J164619.95+435631.0	NVSS J164619+435631			
J1648.1–4930	0.99	0.94				4107	0	
J1649.2–3004	0.96	0.99						
J1656.4–0738	0.99	1.00				3372	0	
J1659.2–0142	0.93	1.00				4155	3	
J1700.8–4912	0.98	0.98						
J1704.3+1235	0.94	1.00				4844	2	
J1704.6–0529	0.89	0.97				6391	1	
J1708.4+1003c	0.86	1.00						
J1717.3–2809	0.99	1.00						
J1721.0+0711	0.91	0.97				8329	2	
J1724.9–0508c	0.99	1.00						
J1727.6+0647	0.96	0.99						
J1727.8–2308	0.99	1.00				3618	0	
J1730.8+5427	0.89	0.97						
J1731.9–2703c	0.99	0.93						
J1733.2–2913c	0.94	0.97						
J1733.4–2812c	0.96	1.00						
J1738.9+8716	0.97	1.00				4140	1	
J1741.0+1347	0.97	1.00				4167	1	
J1742.0–2540c	0.98	1.00						
J1747.6+0324	0.95	1.00				3794	0	
J1748.8+3418	0.98	1.00				3380	4	
J1748.9–3923	0.99	1.00				3600	1	
J1749.1+0515	0.98	1.00				4280	0	
J1757.5–6028	0.95	1.00				4027	0	
J1759.4–2954	0.97	1.00						
J1759.5–0521	0.98	1.00				4092	1	
J1805.0–0845	0.99	1.00				3884	1	
J1827.6+1149	0.96	0.96				3588	4	
J1830.4–1634	0.87	0.96				3395	0	
J1835.4+1036	0.98	1.00				3377	0	
J1835.4+1349	0.99	1.00	BZB	J183535.34+134848.8		4272	2	✓
J1837.9+3821	0.87	0.97	UND	J183828.80+382704.3				
			BZB	J183837.16+381900.5				
J1842.3–5839	0.99	1.00				7314	2	
J1844.3+1548	0.86	1.00	BZB	J184425.36+154645.9	NVSS J184425+154646	4260	1	✓
J1902.3–1106	1.00	0.98						
J1902.7–7053	0.90	1.00				3270	0	
J1917.0–3027	0.97	1.00				3681	0	
J1919.5–7324	0.99	1.00						
J1925.7–7836c	0.98	1.00						
J1942.9–3528	0.98	1.00						
J1946.7–1118	0.98	0.94						
J1952.6–3252	0.99	1.00						
J1958.6+4020	0.99	1.00				2653	0	
J2002.8–2150	0.94	0.97				4383	1	
J2004.6+7004	0.99	1.00	BZQ	J200503.41+700236.3		4847	1	
			BZB	J200506.02+700439.3	NVSS J200506+700440			✓
J2006.5–2256	1.00	1.00				4749	0	
J2017.4–3215	0.97	1.00						
J2031.4–1842	0.90	1.00						
J2034.7–4201	0.95	1.00				6038	3	
J2036.0+4224c	0.91	0.98						
J2042.8–7317	0.99	1.00						
J2046.2–4259	0.97	1.00				4611	2	
J2053.2+1212c	0.99	1.00						
J2055.8+4754	0.95	1.00						
J2103.5–1112	0.92	0.99						
J2104.9+3555	1.00	0.98				8793	1	
J2107.8+3652	1.00	1.00				4884	0	
J2110.3+3822	1.00	1.00				3533	1	
J2112.3–4832	0.88	1.00						
J2114.1+5440	0.99	0.97	BZQ	J211508.92+544815.7				
J2115.4+1213	1.00	1.00				3801	2	
J2124.0–1513	0.99	1.00	BZQ	J212423.63–152558.2				
J2125.0–4632	0.98	0.99				4568	3	
J2131.0–5417	0.99	1.00						
J2132.5+2605	0.99	1.00				3270	1	
J2133.9+6645	0.92	1.00	BZB	J213349.21+664704.3	NVSS J213349+664706	7254	4	✓
J2134.6–2130	0.99	1.00	BZB	J213430.18–213032.6	NVSS J213430–213032	8883	4	✓

Table 4.3: – continued from previous page

2FGL Name	LRF	LNN	Type	WISE Name	Other Name	XRT Exposure / s	#XRT Sources / 2FGL Region	Coincidence XRT/WISE
J2200.1–6931	1.00	0.99				7118	1	
J2213.7–4754	0.98	1.00				3340	1	
J2237.2+6316	0.87	0.96				17729	3	
J2251.1–4927	0.92	0.97				3864	2	
J2251.8+4211	1.00	1.00						
J2257.9–3646	0.92	1.00				3525	2	
J2300.0–3553	0.99	1.00	BZQ	J230010.16–360159.9	6dF J2300101–360200			
J2319.3–3830	1.00	1.00	BZQ	J232000.11–383511.4	MRSS 347–103293	1916	1	
J2343.3–4752	0.98	1.00				4674	3	
J2347.2+0707	0.98	0.98				5017	0	
J2351.6–7558	0.97	0.99				3623	2	
J2354.2–6615	0.97	0.99						
J2356.0–5256	1.00	0.97						
J2358.4–1811	0.98	1.00	UND	J235828.61–181526.6				

4.4 Discussion

Multiple approaches to associate source classes with the UGS of the 2FGL catalog have been attempted. Doert and Errando, 2014 considered only the energy regime of high-energy gamma rays, while the strategy of F. Massaro et al., 2013c was based mainly on infrared wavelengths. As part of this chapter, the X-ray regime has been incorporated in addition to the gamma-ray regime, and the findings of the different procedures have been compared.

In the meantime, the third *Fermi*-LAT (3FGL) catalog (Acero et al., 2015) has been published, featuring twice the observation time, refined analysis methods, more gamma-ray point sources, and newly associated sources, previously contained in the UGS sample. In this way, a validation of the predictions of the presented procedures is enabled. The 95% uncertainty regions obtained from the 3FGL catalog are displayed in the above-mentioned images (Figure 4.1). For most of the sources, the size of the uncertainty region decreased, but for a few sources the most probable position changed to quite a large extent. In order to link 2FGL and 3FGL sources, an overlapping of their uncertainty regions is demanded in this validation. This is in contrast to the procedure applied for the 3FGL catalog, requiring a specific angular distance between the most probable positions dependent on the confidence of the semi-major axes of the uncertainty regions, but has been rated as more reasonable here, since the eccentricity of the uncertainty region is taken into account.

Out of the 231 high-confidence AGN candidates, 61 candidates are newly associated in the 3FGL catalog, and 54 of these are associated with AGNs. Comparing the resulting false-association rate of 11.5% to the predicted one of 11.2%, the validity of the deployed procedure is confirmed impressively, and the expected qualitative prospects are shown. The division into the AGN subclasses BL Lac object (BLL), flat-spectrum radio quasar (FSRQ) and AGN of uncertain type (BCU) of the predicted candidates is provided in Table 4.4. These values are also evaluated and provided for the other predictions, fulfilling the stated criteria. For high-confidence AGN candidates unified with a single or at least a single XRT counterpart, the false-association rate improves in both cases to 4.5%, but the number of candidates decreases due to the additional criteria. An even more constraining criterion is the need of a coincidence between a WISE candidate and an XRT point source, resulting in a false-association rate of 0.0%. Considering only the WISE candidates in combination with the AGN candidates, the false-association rate amounts to 3.8%, and is decreased to 0.0% as well, when an additional coincidence with a further wavelength is required. This shows the strength and necessity of multi-wavelength approaches to reliably associate the UGS.

The inspection of the division into subclasses indicates that more BLLs than FSRQs are apparently contained in the predicted samples, independent of the strategy and criteria. This is explained by the brightness of the AGN candidates and the nature of FSRQs: The high-confidence candidates are generally fainter than the associated 2FGL sources, and FSRQs tend to higher luminosities than BLLs. This is indicated for instance by the finding that the FSRQ sequence starts at higher luminosities than the BLL sequence in Ghisellini et al., 2017. Correspondingly, it is expected that the AGN candidate sample comprises less FSRQs than BLLs.

Table 4.4: Validation of AGN predictions for different criteria. AGN candidates refer to the high-confidence candidates from Doert and Errando, 2014, and XRT point sources to the *Swift*-XRT analysis presented in this chapter. The WISE counterpart candidates, together with additional other counterparts, originate from F. Massaro et al., 2013c. The validation is enabled by comparing the predictions based on the 2FGL catalog to the newly associated sources in the successive 3FGL catalog.

Criterion	# 2FGL Sources			False-Association Rate
	BLL	FSRQ	BCU	/ %
AGN candidate	11	6	37	11.5
AGN candidate ∩ (≥ 1 XRT)	10	6	26	4.5
AGN candidate ∩ 1 XRT	6	4	11	4.5
AGN candidate ∩ Coincidence WISE / XRT	5	3	11	0.0
AGN candidate ∩ WISE	7	3	15	3.8
AGN candidate ∩ Coincidence WISE / Other	6	3	11	0.0

5 Catalogs and Data Samples for the Search of Active Galactic Nuclei and Blazars

...it has great practical value

— Douglas Adams, *The Hitchhiker's Guide to the Galaxy*

The search for Active Galactic Nuclei (AGN) and blazars in the sample of Unassociated Gamma-ray Sources (UGS) is a challenging task. It requires information of multiple wavelengths and experiments to associate UGS with source classes and with their multi-wavelength counterparts, i. e. corresponding sources observed at other wavelengths.
For the analyses presented in chapter 6 and chapter 7, specific data samples have to be prepared. They are based on multiple catalogs from various wavelengths and experiments. Thus, the used catalogs are introduced, and the unification of the catalogs to create the data samples is explained.

5.1 Description of Catalogs

For simplified comparisons, Table 5.1 gives an overview of the different catalogs. Details of the individual catalogs are given in the following subsections.

Table 5.1: Summary of the properties of the different catalogs.

	3FGL	1SXPS	ALLWISE	FIRST
Qualifier	Gamma-ray	X-ray	Infrared	Radio
Time span years	4	8	1	18
Energy range	100 MeV – 300 GeV	0.3 – 10 keV	55 – 365 meV	6 eV
Energy bands	5	3	4	1
Number of sources	3033	151 524	747 634 026	946 432
Median localization uncertainty 90% c.l.	0.07°	5.5″	0.15″	$< 0.5''$ – 1″

5.1.1 3FGL: Third *Fermi*-LAT Catalog

The third *Fermi*-LAT (3FGL) source catalog (Acero et al., 2015) comprises information of 3033 point sources above 4σ significance, measured with the Large Area Telescope (LAT) aboard the *Fermi* satellite (see subsection 2.4.1) during the first four years of operation. In an energy regime between 100 MeV and 300 GeV, it is the deepest catalog so far. Compared to the 2FGL catalog (cf. chapter 4), the upper bound of the energy regime increased, due to twice the amount of data, improvements in the analysis, and a reprocessing of the data (Pass 7 processing).

In order to construct the catalog, several algorithms have been applied to each source, e. g. to determine their positions, or to reconstruct spectra and derive features describing the spectral properties.
The localization has been determined with a likelihood fit, where the likelihood consists of the product of two Gaussians of two orthogonal angular quantities. Like this, an uncertainty ellipse, described by the major and minor axes and an orientation angle, is derived for different levels of confidence.
Features, connected to the energy spectrum $\frac{\mathrm{d}N}{\mathrm{d}E}$ of a source, are determined by performing various spectral fits. The simplest representation of the spectrum is a power law,

$$\frac{\mathrm{d}N}{\mathrm{d}E} = K \left(\frac{E}{E_0} \right)^{-\gamma} , \tag{5.1}$$

with the flux normalization K, reference energy E_0, and spectral index γ. A more complicated representation is a log parabolic power law,

$$\frac{\mathrm{d}N}{\mathrm{d}E} = K \left(\frac{E}{E_0} \right)^{-a-\beta \ln(E/E_0)} , \tag{5.2}$$

with the spectral index a, and the curvature β. Another spectral shape is represented by an exponential cutoff power law,

$$\frac{\mathrm{d}N}{\mathrm{d}E} = K \left(\frac{E}{E_0} \right)^{-\Gamma} \exp\left(\left(\frac{E_0}{E_c} \right)^{b} - \left(\frac{E}{E_c} \right)^{b} \right) , \tag{5.3}$$

where Γ denotes the low-energy spectral slope, b the exponential index, and E_c the cutoff energy. The best fitting spectral representation is listed in the catalog. In addition, the spectral index of the power law is registered for every source. The spectral parameters, obtained using the full energy range, are also used to derive photon fluxes in five distinct energy bands by adapting the normalization in each band. The same procedure of adjusting the normalization for fixed spectral parameters is conducted for monthly time bins. Based on these monthly fluxes, a variability index of a source is calculated.
In case of specific issues during the determination of the described features, the source is flagged, and should be treated with caution. Exemplary issues are bad qualities during the spectral or localization fit, or unstable results when considering different models or methods.

Other approaches attempt to assign source classes to every source, such as pulsars, supernova remnants, BL Lac objects (BLLs), flat-spectrum radio quasars (FSRQs), or radio galaxies. In this context, a distinction between identifications and associations is made. For firm identifications, the detection of periodic variability, spatial morphology, or correlated variability is necessary. Associations are primarily based on close positional coincidences between the 3FGL source and possible counterparts at other wavelengths. An automated association procedure applies a Bayesian approach, estimating the probability of a chance coincidence by considering the positional coincidence and the expected number of accidental coincidences. Several catalogs have been consulted, such as catalogs of particular surveys, or catalogs with known source classes. The numbers of 3FGL sources for each source class are listed in Table 5.2. The most numerous sources classes are blazars, i. e. BLLs, FSRQs, and BCUs, with above 500 sources each, and pulsars with 167 sources. The total number of AGNs amounts to 1745, while only 278 non-AGN are contained in the catalog. Despite the large efforts to associate source classes and counterparts, still, a number of 1010 sources remains unassociated.

Table 5.2: Number of sources per source class in the 3FGL catalog. No differentiation between associations and identifications has been conducted.

Source Class	Number of Sources
BL Lac type of blazar	660
FSRQ type of blazar	484
Blazar candidate of uncertain type	573
Non-blazar active galaxy	3
Radio galaxy	15
Seyfert galaxy	1
Narrow line Seyfert	5
Compact steep spectrum quasar	1
Soft spectrum radio quasar	3
Total AGN	1745
Pulsar	167
Pulsar wind nebula / Supernova remnant	83
Globular cluster	15
High-mass binary	3
Binary	1
Nova	1
Star-forming region	1
Normal galaxy	3
Starburst galaxy	4
Total non-AGN	278
Unassociated	1010
Total	3033

In Figure 5.1 the sky distributions of different source classes are illustrated. While AGNs are distributed quite uniformly over the sky, non-AGNs populate mainly in the Galactic plane. This is due to the fact that most of the non-AGN source classes, such as supernova remnants are Galactic sources, and correspondingly are located within the Galactic plane. Non-AGNs outside the Galactic plane are mostly pulsars. Considering the sky distribution of blazars only, i. e. BLLs, FSRQs and BCUs, it is noticeable that only few blazars populate the Galactic plane. In this region, it is difficult to associate source classes and link counterparts, due to the bright Galactic foreground in radio and X-rays, and the optical extinction, and thus, e. g. incomplete catalogs. The sky distributions of unassociated sources and blazars of uncertain type are of special interest, since they are subject to the following sections. The unassociated sources are distributed over the complete sky, but they are in particular concentrated in the Galactic plane.
To examine this further, Figure 5.2 visualizes the Galactic latitude distribution for AGNs, and the associated and unassociated 3FGL sources. For AGNs an isotropic distribution is expected, but due to the above-stated difficulties, the distribution shows a drop at latitudes around 0°. Based on the average AGN density, approximately 100 AGNs are still uncovered in the band between $-5°$ and $5°$. The distribution of the unassociated sources suggests that these AGNs are hidden in the sample of unassociated sources.
A list of all features can be found in Appendix A.2. Version 16 of the 3FGL catalog has been downloaded from `http://fermi.gsfc.nasa.gov/ssc/data/access/lat/4yr_catalog/`.

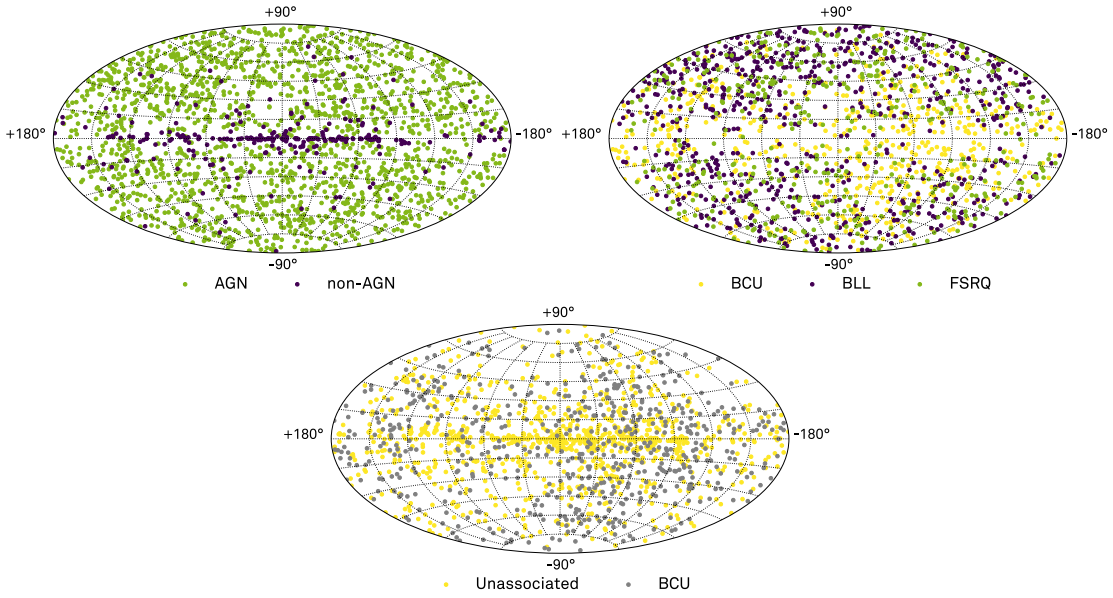

Figure 5.1: Sky maps in Galactic coordinates of the 3FGL catalog for different source classes.

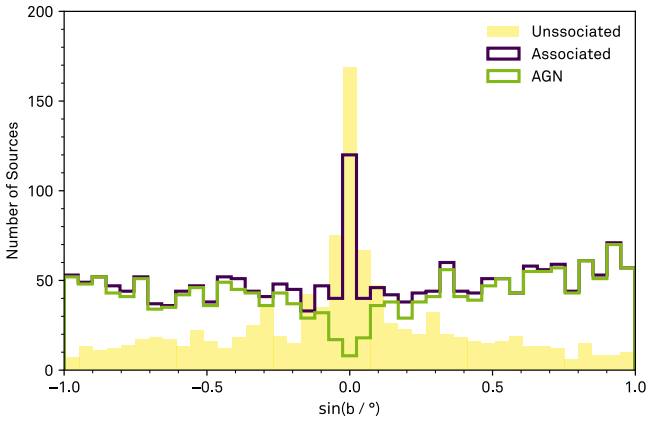

Figure 5.2: Galactic latitude b distribution for AGNs, associated and unassociated sources from the 3FGL catalog. A flat distribution indicates an isotropic distribution.

5.1.2 1SXPS: *Swift*-XRT Point Source Catalog

The deep *Swift*-XRT (1SXPS) point source catalog (Evans et al., 2014) comprises information of 151 524 X-ray point sources in the energy regime 0.3 – 10 keV, measured with the X-Ray Telescope (XRT) on board of the *Swift* satellite (cf. subsection 2.4.2) during the first 8 years of operation.
For the construction of this catalog, a new detection method has been developed to increase the sensitivity to detect faint sources, including an iterative process to determine the background map, a source determination using a point spread function fit, and a Likelihood test. The derived features are based on stacked images from observations in photon counting mode. The measurements are divided in snapshots, i. e. individual, continuous exposures, and observations, which combine all snapshots of one day. Variability measures have been calculated for different energy bands based on light curves, once binned in snapshots, and once binned in observations. The light curves of snapshots provide indications of variability times scales within one day, while light curves of observations indicate variability in time scales of more than one day. To convert the count rates to fluxes, energy conversion factors have been derived for two different models, namely an absorbed power law, and an absorbed optically thin thermal plasma model. Spectral properties are determined with three distinct methods – a spectral fit, an interpolation of hardness ratios, and a fixed spectral shape. Since individual methods for specific sources might fail, the best spectral results are stated as well. The catalog also provides flags to label sources to be handled cautiously.
A list of all features and their descriptions can be found in Table 7 of Evans et al., 2014. The 1SXPS catalog has been downloaded[1] from `http://www.swift.ac.uk/1SXPS`.

5.1.3 FIRST: Faint Images of the Radio Sky at Twenty cm

The final catalog (Helfand et al., 2015) of the survey of Faint Images of the Radio Sky at Twenty cm (FIRST) contains information about 946 432 radio sources at 1.4 GHz, measured between 1993 and 2011 with the Very Large Array (cf. subsection 2.4.3).
The catalog is constructed based on co-added images, i. e. images of multiple telescope pointings. However, no co-addition is performed between old VLA images and images of the expanded VLA, since the properties, such as the frequency response of the receiver or the noise level, changed. For the extraction of sources and their fluxes from the images, an adapted method, appropriate for extended sources as well, has been applied. In a first step, so-called islands, i. e. pixels exceeding a particular threshold plus surrounding pixels, are extracted. Up to four components in terms of two-dimensional Gaussians are fitted to these islands. If the fit results fulfill specific criteria, the elliptical properties of the components, and the corresponding fluxes are stated. In addition, a probability for a spurious source is determined with a combination of several decision tree classifiers. Often, such spurious sources are sidelobes of close bright sources.
A list of features and their descriptions can be found in White et al., 1997. The *14dec17* version has been downloaded from `https://heasarc.gsfc.nasa.gov/`.

[1] This work made use of data supplied by the UK Swift Science Data Centre at University of Leicester.

5.1.4 ALLWISE: Wide-field Infrared Survey Explorer Catalog

The ALLWISE catalog, derived from measurements with the Wide-field Infrared Survey Explorer at 3.4, 4.6, 12, 22 μm, comprises details about 746 634 026 sources in the mid-infrared regime. The catalog is based on co-added images from two complete sky coverage epochs. It provides basic information about the position of the sources, detected with a multiband source detection algorithm, which has been applied to a combined signal-to-noise image from the co-added images of all bands. The estimation of the position of the detected sources is performed with a so-called *proper motion fit*, taking into account the motion of the explorer. Magnitudes and fluxes are obtained separately for each band by conducting profile-fit photometry. Information about the quality of the measurement and the reliability of the source detection and properties are available as flags, e. g. a contamination and confusion flag, an extended source flag, or a variability flag.
Description of parameters and the ALLWISE catalog in general can be found at `http://wise2.ipac.caltech.edu/docs/release/allwise/`. The catalog has been obtained[2] from `http://irsa.ipac.caltech.edu` using the interface[3] to query specific entries of the catalog.

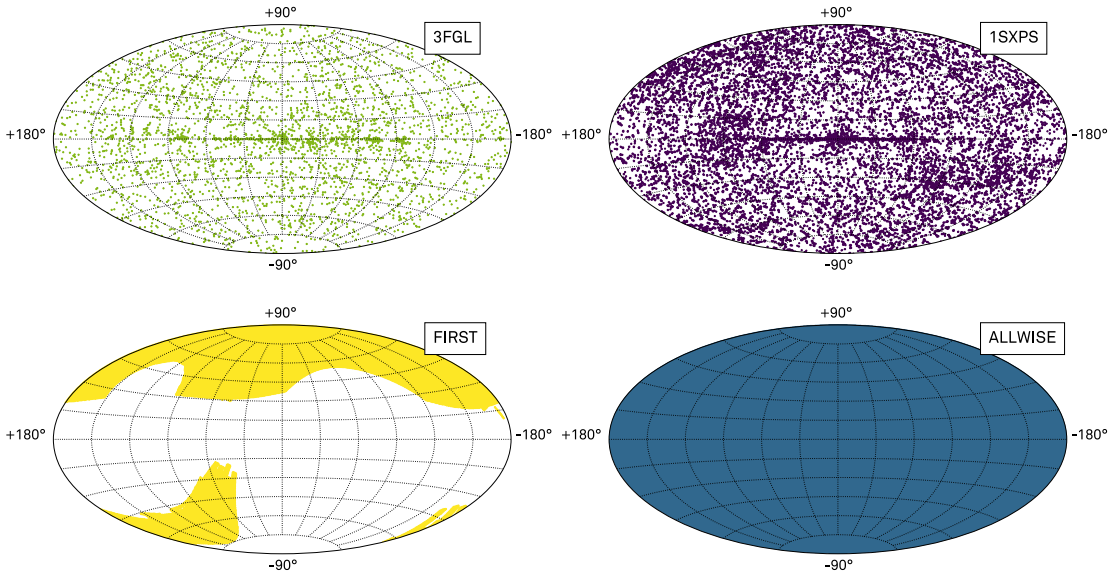

Figure 5.3: Sky maps in Galactic coordinates for different catalogs. The ALLWISE catalog contains an extreme number of sources (747 634 026) and covers the complete sky. Thus, the distribution is very dense. The FIRST catalog comprises a huge number of sources (946 432) as well, but covers only the Galactic caps. The 3FGL and 1SXPS catalogs contain fewer sources, i. e. 3033 and 151 524, respectively.

[2] This publication makes use of data products from the Wide-field Infrared Survey Explorer, which is a joint project of the University of California, Los Angeles, and the Jet Propulsion Laboratory/California Institute of Technology, and NEOWISE, which is a project of the Jet Propulsion Laboratory/California Institute of Technology. WISE and NEOWISE are funded by the National Aeronautics and Space Administration.

[3] This research has made use of the NASA/IPAC Infrared Science Archive, which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

5.2 Preparation of Data Samples

The first step to prepare the data samples is the unification of the individual catalogs. Since the intended goal is to associate UGS, i. e. gamma-ray sources, the 3FGL catalog is merged with one additional catalog at a time. Accordingly, three data samples are prepared: One sample based on information from the 3FGL and the 1SXPS catalogs, another sample comprising the 3FGL and the FIRST catalog, and a third sample relating the 3FGL catalog with the ALLWISE catalog. The unification is performed by searching for all 1SXPS/FIRST/ALLWISE sources – the counterpart candidates – within the 95% uncertainty regions of the localization of the 3FGL sources. 25 extended 3FGL sources (for a list see Table 1 in Acero et al., 2015) and 3 other sources (Crab nebula, Crab pulsar, PSR J1513-5908) are excluded from the search, since no uncertainty regions are defined for these sources. The 1SXPS and FIRST counterpart candidates are searched for in a circular region with a radius equal to the semi-major axis of the elliptic uncertainty region of the 3FGL sources. Using this approximation, a few more sources are potentially found, but this is no major issue for the further analysis. Since the counterparts feature localization uncertainties as well, they are considered in the search. In case of the ALLWISE and the FIRST catalog, the semi-axes and the radius of the 3FGL regions are enlarged by 0.15 and 1 arcsecond (the corresponding uncertainty of 90% confidence level), respectively. The uncertainties in the 1SXPS catalog are quite large compared to the two other counterpart catalogs and range from 2.4 (1% quantile) to 11.5 (99% quantile) arcseconds, and thus, are considered for each source individually. The sources of the counterpart catalogs are evaluated by comparing the angular distance between the 3FGL source and the counterpart with the (enlarged) radius of the ellipse or circle. The angular distance d between two positions with right ascension RA and declination Dec is defined by

$$d = \arccos\left[\sin(Dec_1)\sin(Dec_2) + \cos(Dec_1)\cos(Dec_2)\cos(RA_1 - RA_2)\right] . \tag{5.4}$$

This evaluation is performed on the complete 1SXPS and FIRST catalogs. However, the ALLWISE catalog is very large and cannot be downloaded as one single file, but it offers the possibility to query only parts according to user-specific criteria. Apart from the criterion that the counterparts are localized within the enlarged ellipse, the counterparts of the ALLWISE catalog need to exhibit a signal-to-noise ratio larger than 2 in all bands. Measurements with smaller ratios result in flux upper limits, which are problematic during the feature generation and further analysis.

Table 5.3: Statistics of found counterpart candidates in the different catalogs.

Catalog	# Counterparts in 3FGL Unc. Region	# 3FGL Sources with Counterparts	# Counterparts per 3FGL Source
1SXPS	3 866	1 207	3.2
FIRST	4 561	663	6.9
ALLWISE	117 145	2 883	40.6

Table 5.3 gives an overview of the statistics of the found counterpart candidates in the different catalogs. While on average only approximately 3 1SXPS and 7 FIRST counterparts are located in the uncertainty region of 3FGL sources, the number of about 40 ALLWISE counterparts per region is much larger. This makes the association of the correct counterpart even more important, but also more difficult, as there are more counterparts to choose from. On the other hand, the training sample is larger, which facilitates the creation of a suitable model.

The next step is the choice of features from the individual catalogs for the subsequent analysis, and a generation of new features thereof. In the following, the features considered for the feature generation and selection are described for each catalog.
From the 3FGL catalog, information regarding the fluxes in five energy bands and the integral flux is considered, as well as spectral properties in terms of the spectral index, the pivot energy, and the curvature. In addition, features stating the variability of the flux over time, the significance of the source on average and in the peak interval, and the significance of the curvature of the spectrum are available.
From the 1SXPS catalog, features specifying the fluxes in three energy bands and the integral flux, and the corresponding unabsorbed fluxes, derived with different absorption models, are included. Moreover, features characterizing the flux variability between observations, the number of observations with a detection, the number of observations, and the results of spectral fits, i. e. the spectrum type, the spectral index and parameters of the absorption model.
From the FIRST catalog, features indicating the flux at one frequency, the integral flux in the band, the properties of the source extension, and the probability for the presence of spurious sources are incorporated.
From the ALLWISE catalog, features denoting the magnitude in four bands, the Stetson K variability indices for each band, and the correlation coefficients between two bands are included.

For the final pre-selected feature set, the inclusion of features correlated to the brightness, such as fluxes, of a source are avoided. These features can introduce a bias, since the associated sample tends to be brighter than the sample of unassociated sources. Instead, these features are adapted within the feature generation, e. g. by calculating flux ratios or normalizing features to the integral flux.
Following the same approach as in section 4.1, hardness ratios $\mathtt{HR}_{\mathtt{ij}}$ of fluxes $\mathtt{F}_{\mathtt{i}}$, $\mathtt{F}_{\mathtt{j}}$ and mean energies $\mathtt{E}_{\mathtt{i}}$, $\mathtt{E}_{\mathtt{j}}$ for consecutive bands *i* and *j* are derived:

$$\mathtt{HR}_{\mathtt{ij}} = \frac{F_i E_i - F_j E_j}{F_i E_i + F_j E_j} \,. \tag{5.5}$$

Based on these ratios, `hardness slopes` for consecutive bands are determined:

$$\mathtt{hardness\ slope}_{\mathtt{ijkl}} = \mathtt{HR}_{\mathtt{ij}} - \mathtt{HR}_{\mathtt{kl}} \tag{5.6}$$

and

$$\mathtt{hardness\ slope}_{\mathtt{ijk}} = \mathtt{HR}_{\mathtt{ij}} - \mathtt{HR}_{\mathtt{jk}} \,. \tag{5.7}$$

Variabilities, significances, signal-to-noise ratios and fluxes in individual bands are normalized to the integral flux of one of the catalogs. Additionally, hardness ratios between different catalogs are divided by each other, and ratios between fluxes of distinct catalogs are generated. From the 1SXPS features, the ratio between the number of detected observations and total observations, normalized to the integral flux is determined. The FIRST features are used to derive the shape and an area normalized to the integral flux to describe the extension of a source. Based on the ALLWISE features, the colors, i. e. the differences, of the magnitudes of distinct consecutive bands are derived, as well as the ratios thereof. Moreover, the angular distance between a counterpart and the most probable position of the 3FGL source is calculated for every counterpart catalog.

Table 5.4: Statistics of features from the different catalogs of the pre-selected feature sets.

	1SXPS	FIRST	ALLWISE
# Features from Counterpart Catalog	25	8	25
# Features from 3FGL Catalog	17	17	17
# Features from Unified Catalogs	13	9	8
Total	55	34	50

The resulting feature set is checked for values which can cause problems in the subsequent analysis steps, as e. g. *NaN* or infinity values. One possibility to cope with this is the complete removal of features with many problematic values or the removal of the specific counterpart from the sample. This way, information is lost, and the samples become even more sparse. As an alternative, the problematic values are set to values outside the range of the corresponding feature. Like this, no loss of information occurs, and models, such as Random Forest classifiers, are still able to treat the corresponding counterparts separately. Another issue for some analysis steps are categorical features, and thus, they are transformed with a *one-hot* encoding.
The process of selecting and generating feature sets as described above serves as a feature pre-selection for the analyses in chapter 6 and chapter 7. The following naming convention has been introduced: Features suited for the subsequent analyses receive a catalog tag placed in front of the feature name, such as `3FGL_`, `1SXPS_` or `3FGL_1SXPS_`. Table 5.4 gives an overview of the features from the different catalogs of the pre-selected feature sets.

5.3 Creation of Training and Application Samples

The training and application samples are based on the above-described samples. For the analyses in chapter 6 and chapter 7, two different training samples derived from each counterpart catalog are needed: One sample for the search for AGNs, and one for blazars. The AGN training sample receives a label with two categories – *AGN* and *non-AGN* – while the blazar training sample has a label with three categories – *BLL*, *FSRQ* and *non-blazar*. These training samples comprise all associated 3FGL sources and their corresponding counterparts. For the blazar training sample, all 3FGL sources associated with a blazar of uncertain type and their counterpart candidates are removed, since they do not fall in any of the three categories. Moreover, all 3FGL sources not associated with a BLL or a FSRQ and their counterparts are labeled as *non-blazars*. Analogously, all 3FGL sources not associated with an AGN (cf. Table 5.2 for all AGN sub-classes) and their corresponding counterpart candidates are labeled as *non-AGN* for the AGN sample. To label the remaining entries of the training samples, the positions of the associated sources are needed. By definition, associated sources are mostly assigned to particular counterparts with more precise localizations. The positions have been obtained using name resolvers from the Canadian Astronomy Data Centre[4], the NASA Extragalactic Database[5] and the SIMBAD Astronomical Database[6]. These positions, the 3FGL and associated source names are listed in Appendix A.1.

[4]This research used the facilities of the Canadian Astronomy Data Centre operated by the National Research Council of Canada with the support of the Canadian Space Agency.
[5]This research has made use of the NASA/IPAC Extragalactic Database (NED) which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.
[6]This research has made use of the SIMBAD database, operated at CDS, Strasbourg, France.

Counterparts in a circular region with a specific radius around the associated position are labeled as *AGN*, or *BLL* and *FSRQ*, respectively, while the remaining counterparts are labeled as *non-AGN* or *non-blazar*. Figure 5.4 illustrates this procedure for the creation of the AGN training sample.

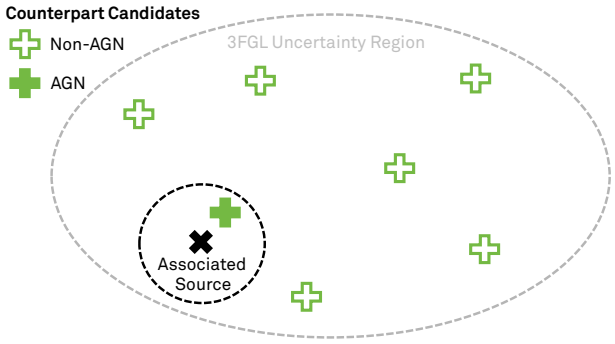

Figure 5.4: Sketch to illustrate the labeling procedure using the example of the AGN training sample. Considering all counterparts within the 3FGL uncertainty region, the counterpart in a specific region around the associated source is labeled as *AGN*, while all remaining counterparts are labeled as *non-AGN*.

The choice of the radius of the circular region is dependent on the catalog, for example on the source localization accuracy of the experiment and the source density in the measured waveband. It is also a trade-off between not linking the corresponding counterpart and linking spurious counterparts. Thus, a method has been applied to adapt the radius for each counterpart catalog separately.
Dependent on the radius R, the number of counterpart associations in annuli $N(R) - N(R - \Delta R)$ of specific widths ΔR around positions p is counted, according to the following equation:

$$\Delta N(R) = \sum_{p} \left(N_p(R) - N_p(R - \Delta R) \right) . \tag{5.8}$$

The width of the annulus has been chosen in relation to the localization accuracy of the corresponding catalog and amounts to $\Delta R_{\mathrm{1SPXS}} = 0.2''$, $\Delta R_{\mathrm{FIRST}} = 0.05''$, and $\Delta R_{\mathrm{ALLWISE}} = 0.025''$. Once, this is performed for the positions of the associated sources, and another time for random shifts of these positions by 40 to 60 arcseconds. The resulting distributions are shown in Figure 5.5 for each catalog. All distributions of the random associations are flat, while the total associations show an excess for small radii. A Gaussian kernel has been used to smooth these distributions to determine the minimum radius at which the distributions of the total associations and the random associations are equal. A radius of 8.2 arcseconds has been obtained for the 1SXPS catalog, a radius of 2.5 arcseconds for the FIRST catalog, and a radius of 3.5 arcseconds for the ALLWISE catalog. Applying these radii to the above-described labeling procedure, no multiple associations for individual sources have been created, while finding associations for a large fraction of the sources.
It has to be noted that these training samples rely on the class and counterpart associations in the 3FGL catalog and on the correct association of counterpart candidates from the considered catalogs as previously described. Thus, they can introduce additional uncertainties to the models.

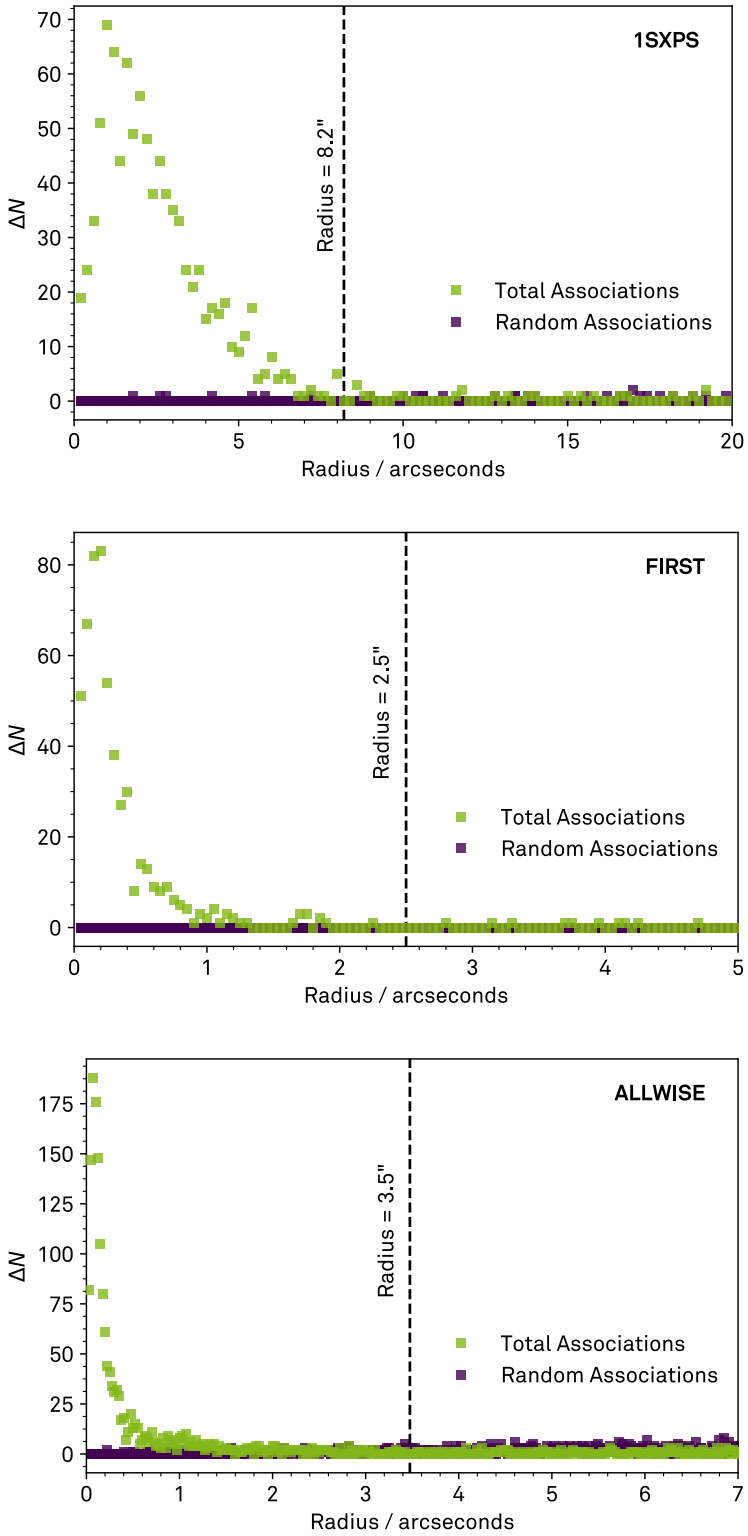

Figure 5.5: Distributions of total and random associations to determine the radius to label counterpart candidates. The parameter ΔN denotes the number of counterpart associations in an annulus, summed up over multiple positions. For the total associations, these positions are equal to the positions of the associated sources, while for the random associations these positions are randomly shifted by 40-60 arcseconds.

Besides the labeled training samples, different application samples, i. e. unlabeled samples, are created. The unassociated sample comprises all unassociated 3FGL sources together with their counterpart candidates. The uncertain blazar sample contains all 3FGL sources, associated with blazars of uncertain type, but only with the associated counterpart candidate, since this type of classified sources exhibits already an associated counterpart. Moreover, for testing purposes, the same sample is produced, in addition including the corresponding counterpart candidates. Table 5.5 gives an overview of the created samples and the corresponding number of entries.

Finally, the range of each feature has been compared for each counterpart catalog between the training and application samples to ensure that the range of the application samples is covered by the training samples.

Table 5.5: Statistics of entries in the created samples.

	1SXPS	FIRST	ALLWISE
AGN Training Sample			
# AGN	887	514	1 297
# Non-AGN	1 485	2 807	14 779
Blazar Training Sample			
# BLL	399	277	507
# FSRQ	313	169	409
# Non-blazar	1 233	2 398	7 851
Application Samples			
# Unassociated	766	964	88 396
# Blazars of Uncertain Type	163	59	360
# Test	367	404	6 868

6 Search for High-Confidence Active Galactic Nucleus Candidates and their Multi-Wavelength Counterparts in the 3FGL Catalog

The Answer to the Great Question...of Life, the Universe and Everything...

— Douglas Adams, *The Hitchhiker's Guide to the Galaxy*

As of yet, the deepest all-sky survey in the gamma-ray regime has been accomplished by the Large Area Telescope (LAT) aboard the *Fermi* satellite. Despite sophisticated methods and great efforts, the Third *Fermi*-LAT Source Catalog (3FGL) comprises 1010 Unassociated Gamma-ray Sources (UGS) out of 3033 point sources detected above 4σ (Acero et al., 2015). The assignment of source classes, such as AGNs and non-AGNs, and the linkage of counterparts at other wavelengths to these UGS are of great interest and importance to refine our knowledge of gamma-ray emitting objects. Multiple strategies, ranging from physical to statistical approaches, have been followed to complete these tasks. For instance, Doert and Errando, 2014 and Mirabal et al., 2012 utilized machine learning methods to discriminate between AGNs and non-AGNs, and between AGNs and pulsars, respectively, and proved capability of statistical procedures. F. Massaro et al. (i. a. D'Abrusco et al., 2013, F. Massaro et al., 2013c, F. Massaro et al., 2013b, Paggi et al., 2013) investigated extensively multi-wavelength data to assign blazar classes and to link multi-wavelength counterparts to the UGS. In chapter 4, a combination of these statistical and physical procedures has been conducted, and the prospects of this combination have been examined. For a list of AGN candidates, selected by machine learning methods applied to the 2FGL catalog, a series of X-ray analyses and a search for correlations with further wavebands have been performed to find counterparts, and thus, to associate sources with a class. A validation revealed that the accuracy of the association increased when considering multiple wavelengths. However, quite some manual work was necessary, complicating an update with more recent data and an extension to further wavelengths in the future. In recent years, the field of data science and analytics has grown immensely and provides state-of-the-art techniques to deal with big data and to extract knowledge from it. It has also become an integral part in the exploration of astronomical data. For these reasons, a new method to assign source classes and to link counterparts using machine learning methods and multi-wavelength data has been developed within the framework of this thesis.

The point source localization uncertainty of *Fermi*-LAT measurements is in the order of several arc-minutes. As a result, several hundreds of possible counterparts might be located within this region, making the association of a 3FGL source with a counterpart at another wavelength ambiguous, and the extension of the 3FGL catalog data to multi-wavelength data difficult. To circumvent this difficulty at this point, every possible combination of a specific 3FGL source with all candidates of one counterpart catalog, i. e. counterparts in the corresponding 3FGL uncertainty region, has been considered. These combinations feature the requested multi-wavelength information, and are utilized to improve the class affiliation models. For the creation of these models, a Random Forest classifier has been applied.

This classifier assigns scores closely related to a class affiliation to these combinations. Combinations with high scores imply not only a specific class affiliation, but also the corresponding counterpart. Like this, source classes can be assigned and counterparts linked to 3FGL sources at the same time. The procedure to merge the 3FGL catalog with an X-ray, an infrared and a radio catalog for the counterpart candidates is explained in section 5.2. Proceeding from this, the generation of samples to build the classification models is described in section 5.3.

Within the scope of this chapter, various objectives have been pursued. First, a classification between AGNs and non-AGNs – which corresponds to a 2-class problem – has been conducted, and second, the association with multi-wavelength counterparts. Initially, the focus has been on the proof of concept of this novel approach, using the example of one catalog per X-ray, infrared and radio waveband. For each of these three wavebands, a list of high-confidence AGN candidates and their counterparts has been compiled. Subsequently, the results of the individual counterpart catalogs have been merged to fully exploit the power of multi-wavelength studies. Based on that, physical conclusions have been drawn from the features that turned out to be important in the classification and from the capability of the individual wavebands for the assigned tasks.

The developed method is in some aspects quite different compared to typical machine learning applications in particle physics. On the one hand, no simulations are used to train the classification model. Instead, the training sample has been compiled based on the associations provided in the 3FGL catalog and further assumptions regarding the corresponding counterpart (cf. section 5.3). However, these associations and assumptions are not necessarily correct. Moreover, the training sample does not originate from the same probability density function as the sample to be classified, since the training sample comprises sources that are moderate to associate, while the other sample includes sources difficult to associate. On the other hand, some restrictions apply to the cross validation procedure. A particular 3FGL source is multiplied by the number of counterparts in its uncertainty region, and multiple counterpart combinations contain the same information. To avoid the introduction of a bias, all derived combinations have to end up in the same cross validation subsample. Since the number of counterpart candidates per 3FGL source varies, every subsample exhibits different signal-to-background ratios, influencing e. g. the performance evaluation.
The developed method has been realized using the machine learning library *scikit-learn* (Pedregosa et al., 2011).

6.1 AGN Candidates and Infrared Counterparts

F. Massaro et al., 2012 discovered a strip in the infrared color-color space, containing mostly blazars – a subclass of AGNs. Consequently, the utilization of infrared measurements to associate UGS is very promising. In this section, the generation of two samples of AGN candidates and their infrared counterparts is described. One sample provides high-confidence candidates that can be further processed directly. The second sample contains medium-confidence candidates and is utilized for further processing in section 6.4, together with the results obtained for the X-ray and radio catalogs. The infrared counterpart candidates originate from the ALLWISE catalog, which has been introduced in subsection 5.1.4. The classification model is built with 1 297 AGNs and 14 779 non-AGNs, and is applied to 88 396 counterpart combinations, which are located in 1 006 3FGL uncertainty regions.

6.1.1 Feature Selection

The selection of features the classification model is built upon, is an integral part of the classification procedure. In section 5.2, a set of features, appropriate for the classification, has been pre-selected. This feature set still comprises highly correlated features, which are to be removed. An initial Random Forest classification model has been built, using the complete pre-selected feature set. The correlation has been calculated according to the Spearman's rank correlation coefficient, detecting monotonic dependencies between features, instead of only linear dependencies as it is the case for the Pearson correlation. Here, two features have been defined as highly correlated, if the absolute correlation exceeds 0.95. To decide which of the two features is to be removed, the feature importance of the Random Forest classification model has been used as the criterion. The feature with the smaller importance has been removed.
To select a set of features from the remaining ones for the final classification, a recursive backward elimination has been performed (see subsection 3.1.3 for different feature selection approaches). An advantage of the approach of backward elimination is that the classification algorithm, which is to be applied for the final classification model, is taken into account. By construction, it also considers multivariate dependencies between the features to find an optimal feature set. For this purpose, a Random Forest classification has been set up with the following hyperparameters:

Number of trees 200
Split selection criteria Information gain
Number of features considered at each node $\sqrt{\text{Number of available features}}$
Minimum number of samples for split 5

The first classification model is created with all available features, and the feature with the smallest resulting feature importance is removed. With this new feature set, another classification model is generated, and the procedure is recursively repeated, until only one feature is left. The performance of each feature set is evaluated in addition within a 10-fold cross validation, measured with the Area Under the Curve (AUC) score (cf. section 3.3). The mean and the standard uncertainty of the mean of the AUC scores over the 10 cross validation steps are depicted in Figure 6.1 as a function of the feature set.

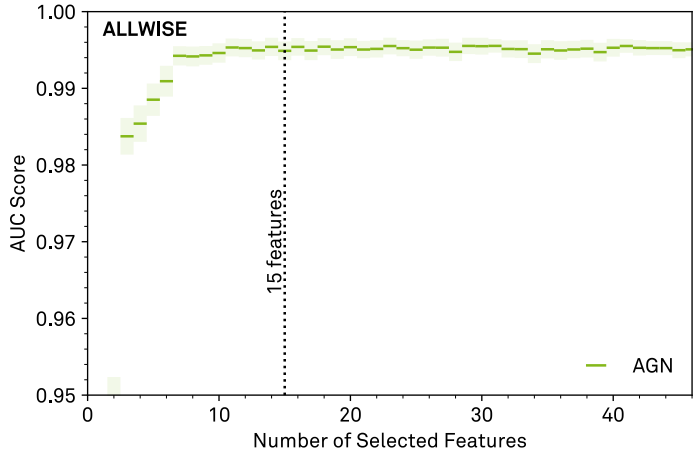

Figure 6.1: Performance of different numbers of selected features, determined with a recursive backward elimination for the ALLWISE catalog in a 10-fold cross validation. The performance is quantified by the mean and the standard uncertainty of the mean of the AUC score, and is saturating above a specific number of features. A number of 15 features has been chosen for the further analysis.

The score is rising with an increasing number of features in the set. However, the score is saturating starting from a specific number of features, and no maximum is formed. An explanation might be that most of the features contain the same information, and do not provide new information to the classification model, and are thus commutable. Due to that, the set with minimum features, once the score is saturating, has been chosen for the subsequent analysis. The 15 selected features are listed in Figure 6.2. Out of these features, 12 are based solely on the ALLWISE catalog, 1 on the 3FGL catalog, and 2 have been generated based on both catalogs.

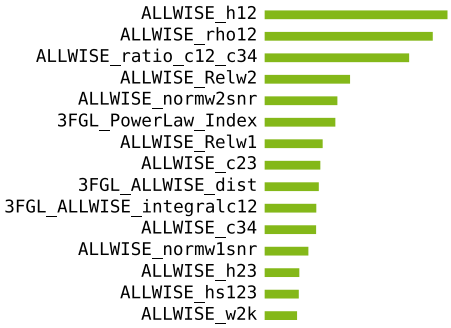

Figure 6.2: List of selected features and their relative feature importance. Out of these, 12 features are based on the ALLWISE catalog, 1 on the 3FGL catalog, and 2 on both catalogs.

6.1.2 Optimization and Performance of the Classification Model

The greatest optimization potential has already been exploited through the feature generation and selection. Further optimization has been performed by tuning the hyperparamters of the Random Forest classifier, such as the number of features considered at each node or the split selection criterion. Increasing the number of trees in the forest does not necessarily lead to a better performance, it just has to be ensured that the number is large enough to reach a saturating performance. To ensure a sufficiently accurate performance, a number of 200 trees has been chosen. The minimum number of samples in a node required for a split is a parameter to steer the pre-pruning of a decision tree. Pre-pruning is a method to abort the growth of a tree by a stop criterion, and helps to prevent overfitting. The number of features considered for a split at a node influences the randomness and diversity of the trees. A common general choice of this hyperparameter is the square root of the total number of features in the set, but it might be advisable to adapt the choice to the properties of the feature set. For sets with features comprising very similar information, smaller values might perform better, while for sets with many irrelevant features larger values are useful. The number of features considered at a node and the minimum number of samples in a node for a split, together with the split selection criterion, have been optimized by scanning a range of appropriate values, rating them according to the AUC score, obtained within a 10-fold cross validation. The maximum AUC score has been achieved with the following hyperparameters:

Number of trees 200
Split selection criteria Information gain
Number of features considered at a node 4
Minimum number of samples for split 5

With these optimized hyperparameters, the final model has been created. The corresponding feature importance is visualized in Figure 6.2. Three features exhibit a noticeable larger importance than the others: The hardness ratio between two infrared wavebands, the correlation coefficient between two wavebands, and the ratio between two infrared colors. Thus, the information from the ALLWISE catalog is crucial for the classification and counterpart search.

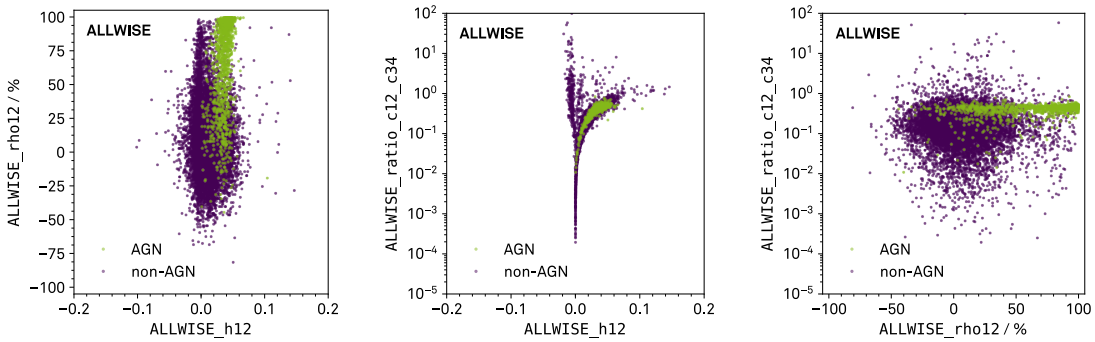

Figure 6.3: Two-dimensional distributions between the three features with the highest feature importance for AGNs and non-AGNs. The AGNs populate a distinctive region in every distribution.

Figure 6.3 shows two-dimensional distributions of these three features. In every distribution, the AGNs populate a distinctive region, and a large amount of non-AGNs can be discarded by simple cuts. It is noticeable that AGNs exhibit a quite strong correlation between flux measurements in two wavebands, while other infrared sources feature small correlations and anti-correlations. This is related to the variability of the sources. AGNs are known to be variable, while typical infrared sources like stars feature stationary spectra. For stationary sources, the fluxes fluctuate solely statistically around a mean value, independently in each observation. Thus, the correlation between different measurements thereof scatters around zero. In contrast, different measurements of fluxes of variable sources fluctuate around diverse values, varying simultaneously in similar wavebands for the assumption of the same underlying process, which causes the variability. Correspondingly, the correlation between the flux measurements are correlated – the stronger the flux variation, the stronger the correlation. Moreover, it can be seen that the AGNs feature positive hardness ratios in the infrared regime, i. e. the source brightens with increasing energy, which is consistent with typical AGN spectra. Non-AGNs feature both positive and negative hardness ratios. Another observation is that the ratio between the infrared colors, related to the curvature of the spectrum, takes only values in a very specific range for the AGNs. In summary, the following physical conclusions can be drawn: In the infrared regime, AGNs are variable, feature a characteristic curvature, and a positive and distinctive slope.

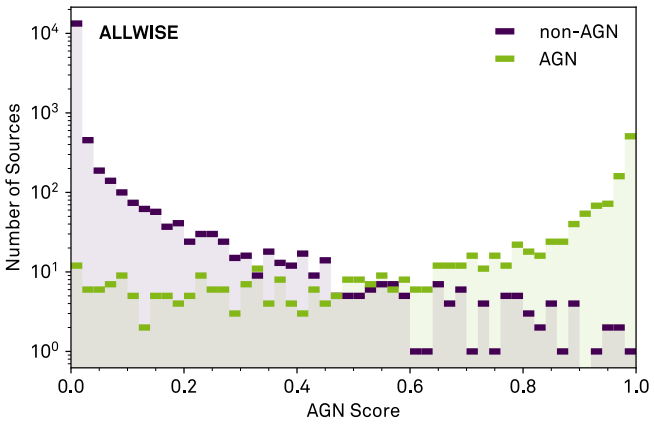

Figure 6.4: Distribution of the AGN scores of the training sample, obtained in a 10-fold cross validation. The scores have been derived from the Random Forest classification model, created with the 3FGL and the ALLWISE catalogs. While the AGN class features a population at large AGN scores, the non-AGN class shows a population at small AGN scores.

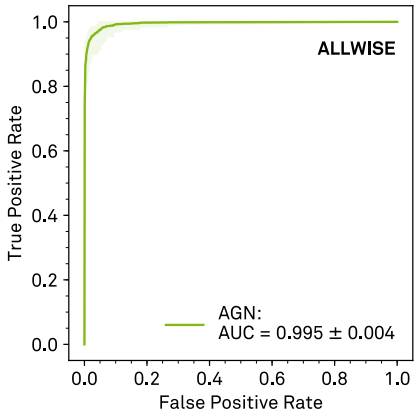

Figure 6.5: ROC curve of the Random Forest classification model for the ALLWISE catalog. The curve has been averaged over 10 curves from the 10-fold cross validation. The shaded area illustrates the minimum and maximum scores. The corresponding AUC and its uncertainty in terms of the standard deviation amounts to 0.995 ± 0.004.

The distribution of the AGN score is shown in Figure 6.4 for the training sample. The counterpart combinations labeled as *non-AGN* exhibit a clear excess at small AGN scores, while the combinations labeled as *AGN* have an excess at large AGN scores, indicating a successful classification. To quantify the performance of the classification, multiple performance measures have been calculated.
In Figure 6.5, the Receiver Operating Characteristic (ROC) curve, i. e. the true positive rate dependent on the false positive rate, is depicted. The curve has been averaged over the 10 curves from the cross validation, and the shaded area visualizes the minimum and maximum true positive rate for each false positive rate. The area under this curve denotes the previously mentioned AUC score. A score of 0.995 ± 0.004 has been achieved.
Further performance measures like precision and recall have been calculated as a function of the threshold of the AGN score. So far, no threshold has been chosen to define the class assignment and corresponding counterpart of a 3FGL source, based on the predicted score. The choice of the threshold is a trade-off between high precisions and high recalls, and depends on the intended use of the resulting data sample. Both performance measures are presented in Figure 6.6 for different thresholds of the AGN score. The shapes of the scores of the individual cross validation steps are partly rather different. This is i. a. related to the different signal-to-background ratios in the particular cross validation subsamples.

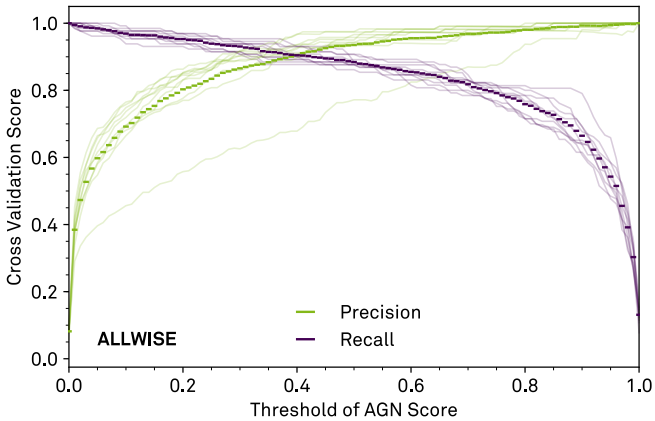

Figure 6.6: Precision and recall for different thresholds of the AGN score of the Random Forest classification model for the ALLWISE catalog. The averaged precision and recall scores have been determined in a 10-fold cross validation. The individual scores of each cross validation subsample are shown as well.

6.1.3 Application of the Classification Model to the Unassociated Sample

Applying the final model to the unassociated sample, a distribution of AGN scores, as shown in Figure 6.7, has been generated. This distribution is rather different compared to the overall distribution of the training sample. This is due to the following: The training sample has been created from associated sources, based on specific and rather obvious properties of the source. On the contrary, complicated and uncertain sources have not been associated, and thus, are not included in the training sample, but are now subject to the classification.

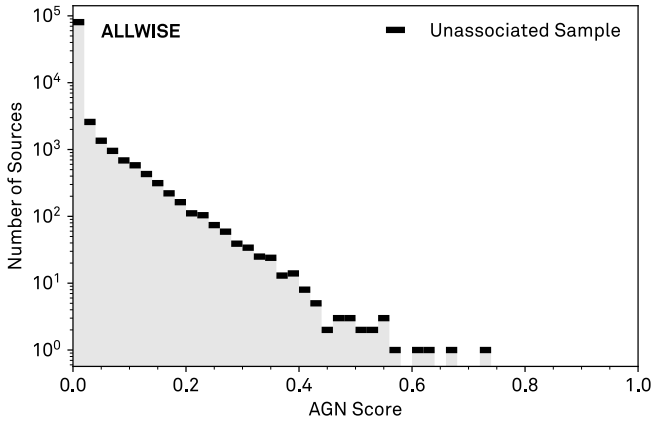

Figure 6.7: Distribution of the AGN scores of the unassociated sample. The scores have been determined with the Random Forest classification model, created with the 3FGL and the ALLWISE catalogs. Only few sources feature large AGN scores.

Still, as stated previously, a threshold of the AGN score has to be chosen to define the predicted class, depending on the intended use. Two samples for different purposes and fulfilling diverse requirements have been created. The purpose of one sample – the high-confidence ALLWISE sample – is to provide directly AGN and corresponding counterpart candidates of high-confidence, requiring a high precision. The other sample – the medium-confidence ALLWISE sample – aims at a large number of candidates, dealing with the compromise that the precision is only moderate. The purpose of this sample is to obtain high-confidence candidates by considering medium-confidence samples of other wavelengths together, requiring small distances between the medium-confidence candidates of the different wavelengths. Hence, a precision of 0.9 has been chosen for one sample, and a precision of 0.6 for the other one, corresponding to thresholds of the AGN score of 0.39 and 0.05, respectively.
Depending on the threshold, multiple counterpart candidates for single 3FGL sources might be predicted, making the association ambiguous. Figure 6.8 illustrates the multiplicities of the counterparts. For a precision of 0.6, multiplicities of up to 127 medium-confidence counterpart candidates for a single 3FGL source exist, while for the high precision, the association of a high-confidence candidate is unambiguous with one exception. The high-confidence sample comprises 38 counterpart candidates, while the medium-confidence sample contains 4486 candidates in total for 650 3FGL sources. Table 6.1 lists information of all high-confidence candidates, i. e. the 3FGL and the WISE name, the AGN score, the Galactic latitude and 3FGL analysis flags. Out of the 38 high-confidence candidates, 22 exhibit no 3FGL analysis flag and 20 are located outside the Galactic plane. Candidates fulfilling both criteria of no analysis flag and outside the Galactic plane amount to 17.

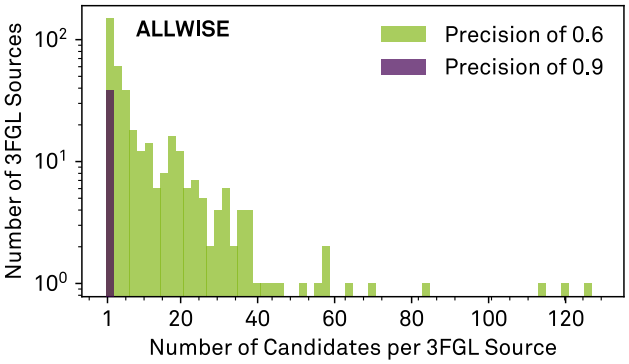

Figure 6.8: Number of 3FGL sources with specific numbers of candidates per 3FGL source, depending on the threshold of the AGN score. For small thresholds, the model associates counterpart candidates to many 3FGL sources, but often multiple candidates for a single 3FGL source. Thus, the association is ambiguous. For large thresholds, the association is definite.

Table 6.1: List of 38 high-confidence AGN candidates and their corresponding ALLWISE counterparts. Out of these, 17 candidates exhibit neither 3FGL analysis flags, nor is located within the Galactic plane.

3FGL Name	$\lvert b\rvert$ / deg	3FGL Flags	WISE Name	AGN Score
J0216.0+0300	54	0	J021600.45+030011.9	0.60
J0221.2+2518	33	0	J022126.96+251433.6	0.41
J0248.4+5130	7	0	J024834.21+513118.2	0.50
J0327.4+5828c	2	33	J032812.08+582939.0	0.40
J0420.4–6013	42	0	J042021.42–601403.0	0.42
J0420.6–3742	45	0	J042025.10–374445.0	0.66
J0538.9+1646	8	0	J053855.10+164611.5	0.64
J0634.1+0424	2	4	J063422.04+042146.4	0.40
J0650.9+6524	24	1	J065131.19+652704.6	0.43
J0725.7–0550	5	0	J072548.08–055338.8	0.47
J0731.8–3010	5	0	J073151.19–301058.9	0.50
J0740.6–5230	14	0	J074029.54–522918.1	0.41
J0802.3–0941	11	0	J080215.90–094210.9	0.56
J0954.8–3948	11	0	J095458.33–394654.9	0.41
J1027.8+8253	33	0	J102841.96+825340.5	0.41
J1049.7+1548	60	0	J104939.35+154837.6	0.47
J1050.4+0435	53	0	J105030.76+043054.9	0.41
J1130.7–7800	16	0	J113032.06–780105.4	0.42
J1209.9+7607	41	0	J120930.27+760911.9	0.49
J1548.4+1455	47	0	J154824.38+145702.8	0.40
J1635.3+4257	42	0	J163549.47+430017.9	0.40
J1748.0+2701	25	1	J174816.75+270433.3	0.45
J1748.5–3912	6	4	J174840.89–391340.5	0.49
J1757.8–3126	4	8	J175753.78–312608.8	0.43
J1829.3–0135c	4	52	J182917.11–013358.5	0.44
J1915.9+1112	0	12	J191611.38+111343.5	0.39
J1922.2+2313	4	4	J192212.25+231520.4	0.52
J1951.8–1102	18	16	J195129.23–110219.9	0.53
J2024.6+3747	0	17	J202433.45+375131.4	0.39
J2034.4+3833c	1	2100	J203427.12+383633.2	0.42
J2044.0+1035	19	0	J204351.62+103407.2	0.56
J2053.9+2922	10	1	J205350.72+292314.5	0.54
J2058.0+4347	1	1	J205741.48+434658.3	0.40
J2147.2+4730c	5	56	J214725.45+473158.7	0.42
			J214726.29+473312.4	0.48
J2223.3+0103	45	0	J222329.57+010226.7	0.57
J2254.1+6229	3	4	J225336.50+623334.3	0.46
J2321.3+5113	9	0	J232127.44+511115.1	0.50
J2358.6–1809	75	0	J235836.83–180717.4	0.73

6.2 AGN Candidates and X-ray Counterparts

In this section, the compilation of two samples of AGN candidates and their X-ray counterparts is described, analogous to the previous section. It is based on X-ray counterpart candidates from the 1SXPS catalog, which has been summarized in subsection 5.1.2. The classification model is created with 887 AGNs and 1 485 non-AGNs, and is applied to 766 counterpart combinations, which are located in 166 3FGL uncertainty regions.

6.2.1 Feature Selection

The same procedure as described in subsection 6.1.1 has been applied for the feature selection in this section. The resulting AUC scores in terms of the mean and the standard uncertainty of the mean are depicted in Figure 6.9. As the number of features in the set increases, the score increases as well, but no maximum is formed, instead, the curve is saturating starting from a particular number of features. This behavior is analogous to the one for the ALLWISE catalog, and the same explanation applies. The 15 selected features are listed in Figure 6.10. Out of these features, 8 are based solely on the 1SXPS catalog, 4 on the 3FGL catalog, and 3 have been generated based on both catalogs.

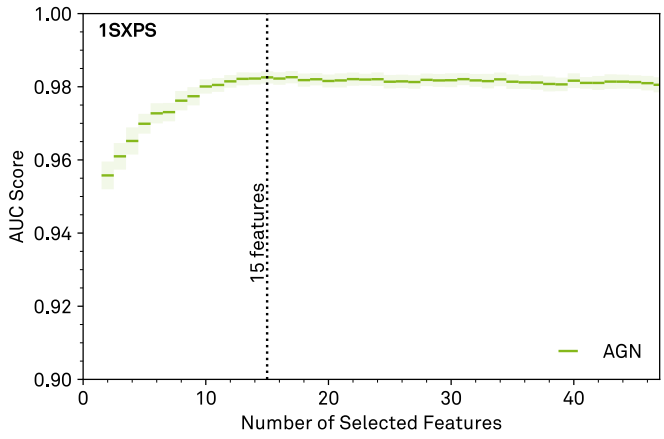

Figure 6.9: Performance of different numbers of selected features, determined with a recursive backward elimination for the 1SXPS catalog in a 10-fold cross validation. The performance is quantified by the mean and the standard uncertainty of the mean of the AUC score, and is saturating above a specific number of features. A number of 15 features has been chosen for the further analysis.

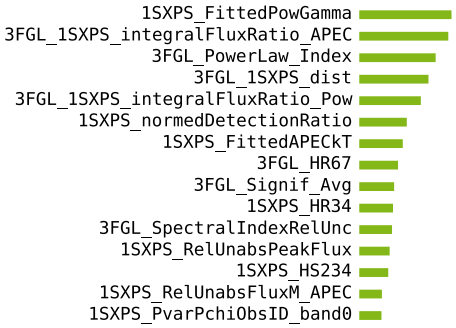

Figure 6.10: List of selected features and their relative feature importance. Out of these, 8 features are based on the 1SXPS catalog, 4 on the 3FGL catalog, and 3 have been generated from both catalogs.

6.2.2 Optimization and Performance of the Classification Model

Analogous to the previous section, the classification model for the 1SXPS catalog has been optimized. The resulting hyperparameters of this optimization are the following:

Number of trees 200
Split selection criteria Information gain
Number of features considered at a node 3
Minimum number of samples for split 8

With these optimized hyperparameters, the final model has been created. Figure 6.10 illustrates the corresponding feature importance. The three most important features are the ratio between the integral X-ray and gamma-ray fluxes, and the power law index of the X-ray spectrum and the one of the gamma-ray spectrum. Thus, the information from both catalogs is crucial for the classification and counterpart search, and especially the spectral indices and the integral fluxes of both catalogs are important.

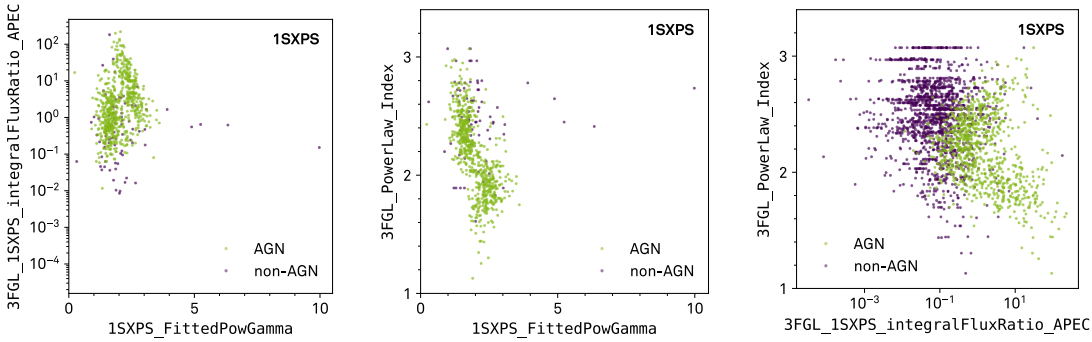

Figure 6.11: Two-dimensional distributions between the three features with the highest feature importance for AGNs and non-AGNs. The feature `1SXPS_FittedPowGamma` contains a lot of invalid values, which are not included in the distributions. In the distribution of the remaining features, the AGNs populate a distinctive region in every distribution.

Figure 6.11 shows two-dimensional distributions of these three features. The fit of the spectral slope of the X-ray spectrum mostly fails for non-AGNs, but this is also a valuable information for the classification. However, the spectral slope of the gamma-ray spectrum, dependent on the ratio of the integral X-ray and gamma-ray flux, indicates a class affiliation. In general, the ratio of these fluxes tend to be higher for AGNs. In summary, the following physical conclusions can be drawn: In the X-ray regime, the connection between properties in the X-ray and gamma-ray regime are important, and AGNs exhibit flux ratios of the two energy regimes in a distinctive range .

The AGN score distributions for the two classes in the training sample are shown in Figure 6.12. Again, one class features mainly large AGN scores, while the other one contains many small AGN scores. The performance has been quantified with the ROC curve, as shown in Figure 6.13. The AUC score, i. e. the area under this curve, amounts to 0.981 ± 0.006. The performance measures precision and recall are presented in Figure 6.14 for various thresholds of the AGN score.

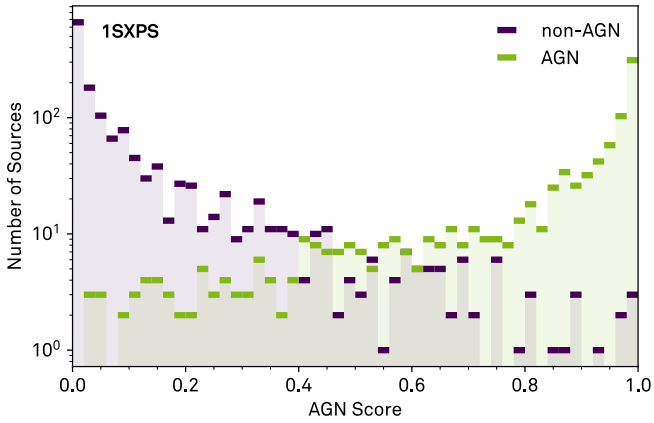

Figure 6.12: Distribution of the AGN scores of the training sample, obtained in a 10-fold cross validation. The scores have been derived from the Random Forest classification model, created with the 3FGL and the 1SXPS catalogs. While the AGN class features a population at large AGN scores, the non-AGN class shows a population at small AGN scores.

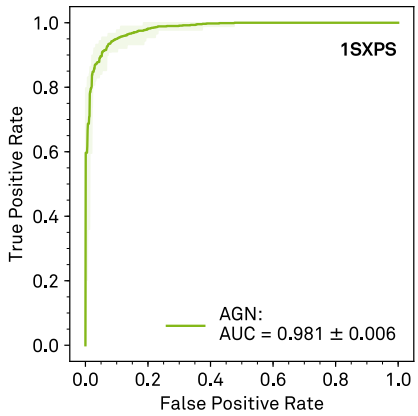

Figure 6.13: ROC curve of the Random Forest classification model for the 1SXPS catalog. The curve has been averaged over 10 curves from the 10-fold cross validation. The shaded area illustrates the minimum and maximum The corresponding AUC and its uncertainty in terms of the standard deviation amounts to 0.981 ± 0.006.

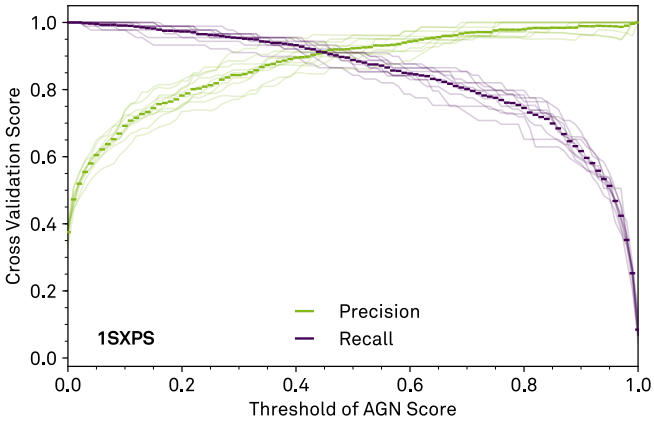

Figure 6.14: Precision and recall for different thresholds of the AGN score of the Random Forest classification model for the 1SXPS catalog. The averaged precision and recall scores have been determined in a 10-fold cross validation. The individual scores of each cross validation subsample are shown as well.

6.2.3 Application of the Classification Model to the Unassociated Sample

The application of the final model to the unassociated sample leads to the distribution of AGN scores, as shown in Figure 6.15. Analogously, this distribution is rather different compared to the overall distribution of the training sample for the same reasons explained previously.

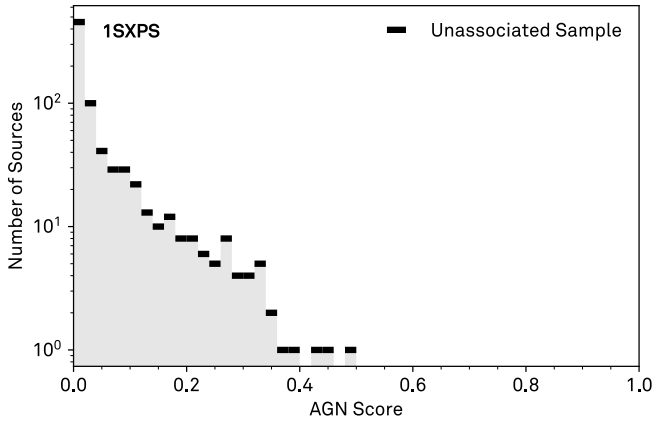

Figure 6.15: Distribution of the AGN scores of the unassociated sample. The scores have been determined with the Random Forest classification model, created with the 3FGL and the 1SXPS catalogs. Only few sources feature medium or large AGN scores.

In order to create the high- and medium-confidence samples, the same precisions of 0.9 and 0.6 as before have been requested, leading to thresholds of the AGN score of 0.42 and 0.05, respectively. The number of counterpart candidates per 3FGL source is illustrated in Figure 6.16. Multiplicities of up to 13 medium-confidence candidates occur for a single 3FGL source for the precision of 0.6. For the precision of 0.9, the association of a high-confidence candidate is unambiguous. The medium-confidence sample contains 185 candidates in total for 112 3FGL sources. The high-confidence sample comprises 3 counterpart candidates. Information regarding this candidate, i. e. the 3FGL and the 1SXPS name, the AGN score, the Galactic latitude and 3FGL analysis flag is listed in Table 6.2.

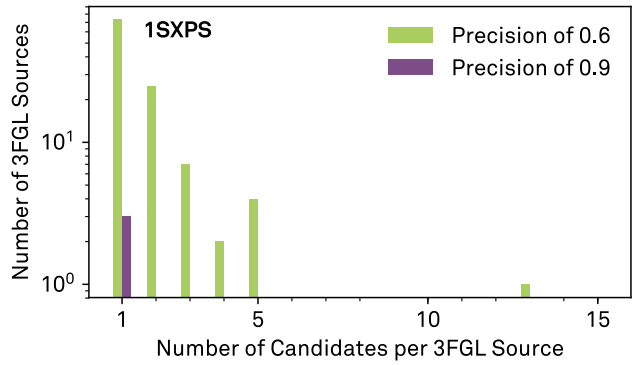

Figure 6.16: Number of 3FGL sources with specific numbers of candidates per 3FGL source, depending on the threshold of the AGN score. For small thresholds, the model associates counterpart candidates to many 3FGL sources, but often multiple candidates for a single 3FGL source. Thus, the association is ambiguous. For large thresholds, the association is definite.

Table 6.2: List of 3 high-confidence AGN candidates and their corresponding 1SXPS counterpart candidates.

3FGL Name	$\|b\|$ / deg	3FGL Flags	1SXPS Name	AGN Score
J0721.5−0221	6	256	1SXPS J072113.8−022053	0.45
J1132.0−4736	13	1	1SXPS J113209.5−473857	0.48
J2300.0+4053	17	0	1SXPS J230012.4+405225	0.44

6.3 AGN Candidates and Radio Counterparts

Relativistic electrons produce synchrotron radio emission, which is capable to upscatter photons to high energies, suggesting a correlation between radio and gamma-ray fluxes. Ackermann et al., 2011 detected such a correlation for gamma-ray emitting AGNs, making the utilization of radio measurements to associate UGS very promising. In this section, the compilation of two samples of AGN candidates and their radio counterparts is described, analogous to the previous sections. It is based on radio counterpart candidates from the FIRST catalog, which has been summarized in subsection 5.1.3. The classification model is created with 514 AGNs and 2 807 non-AGNs, and is applied to 964 counterpart combinations, which are located in 103 3FGL uncertainty regions.

6.3.1 Feature Selection

In this section, the same feature selection procedure as in the previous sections has been conducted. The mean and the standard uncertainty of the mean of the resulting AUC scores are shown in Figure 6.17. With an increasing number of features, the AUC increases first, and then saturates. The curves of the AUC scores of the ALLWISE and the 1SXPS catalog have similar shapes, and the same explanation applies here. The 15 selected features are listed in Figure 6.18. Out of these features, 6 have been extracted from the FIRST catalog, 1 from the 3FGL catalog, and 8 are based on both catalogs.

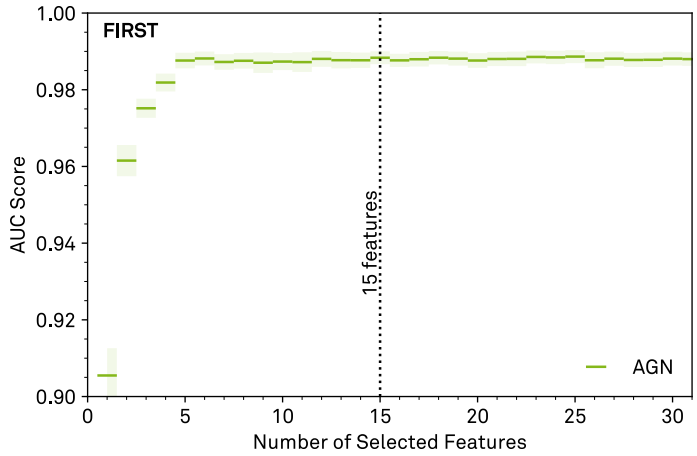

Figure 6.17: Performance of different numbers of selected features, determined with a recursive backward elimination for the FIRST catalog in a 10-fold cross validation. The performance is quantified by the mean and the standard uncertainty of the mean of the AUC score, and is saturating above a specific number of features. A number of 15 features has been chosen for the further analysis.

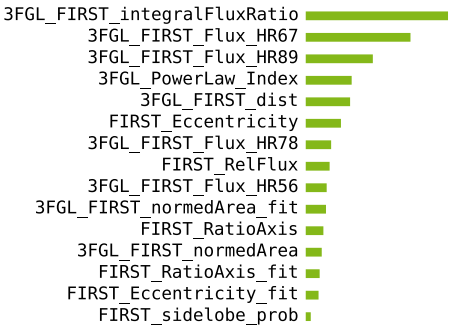

Figure 6.18: List of selected features and their relative feature importance. Out of these, 6 features are based on the FIRST catalog, 1 on the 3FGL catalog, and 8 features on both catalogs.

6.3.2 Optimization and Performance of the Classification Model

The classification model for the FIRST catalog has been optimized analogously to the procedures followed in the previous sections, resulting in the following hyperparameters:

Number of trees 200
Split selection criteria Information gain
Number of features considered at a node 7
Minimum number of samples for split 7

Subsequently, the final model has been generated with the stated hyperparameters. The corresponding feature importances are depicted in Figure 6.18. The three most important features are the ratio between the integral radio and gamma-ray fluxes, and the ratios between hardness ratios of the 3FGL catalog and the radio flux. Thus, the information from both catalogs is crucial for the classification and counterpart search, and especially the flux ratio of both catalogs is important.

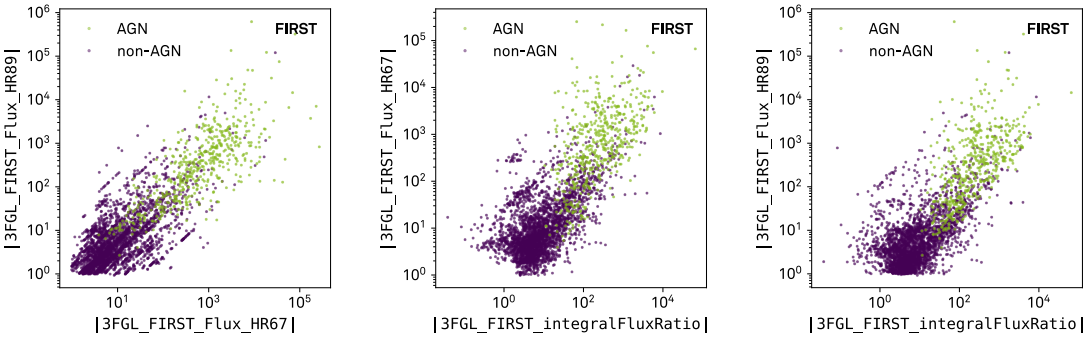

Figure 6.19: Two-dimensional distributions between the three features with the highest feature importance for AGNs and non-AGNs. For illustrative purposes, the absolute values of the features are shown. The AGNs and non-AGNs populate diverse regions in every distribution.

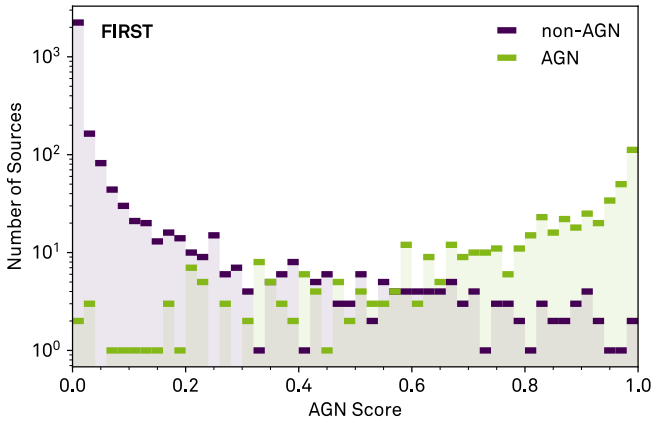

Figure 6.20: Distribution of the AGN scores of the training sample, obtained in a 10-fold cross validation. The scores have been derived from the Random Forest classification model, created with the 3FGL and the FIRST catalogs. While the AGN class features a population at large AGN scores, the non-AGN class shows a population at small AGN scores.

Figure 6.19 shows two-dimensional distributions of these three features. In general, AGNs exhibit both large ratios between radio and gamma-ray fluxes, as well as large gamma-ray hardness ratios compared to non-AGNs. In summary, the following physical conclusions can be drawn: The larger the ratio between radio and gamma-ray flux of an AGN, the softer the spectrum in the gamma-ray regime.

The AGN score distributions for the two classes in the training sample are shown in Figure 6.20. The shape of the distributions is rather similar to the ones obtained for the infrared and the X-ray regime. An AUC score of 0.988 ± 0.006 have been obtained from the ROC curve in Figure 6.22. Additionally, the performance is quantified by the precision and recall dependent on a threshold of the AGN score, which is depicted in Figure 6.21.

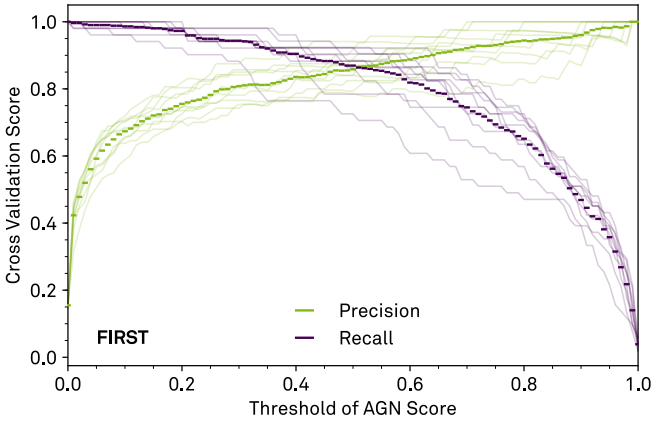

Figure 6.21: Precision and recall for different thresholds of the AGN score of the Random Forest classification model for the FIRST catalog. The averaged precision and recall scores have been determined in a 10-fold cross validation. The individual scores of each cross validation subsample are shown as well.

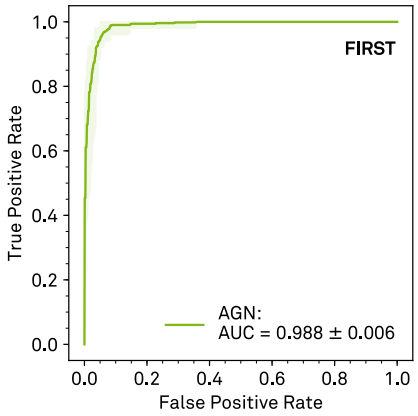

Figure 6.22: ROC curve of the Random Forest classification model for the 1SXPS catalog. The curve has been averaged over 10 curves from the 10-fold cross validation. The shaded area illustrates the minimum and maximum The corresponding AUC and its uncertainty in terms of the standard deviation amounts to 0.988 ± 0.006.

6.3.3 Application of the Classification Model to the Unassociated Sample

The final classification model has been applied to the unassociated sample. The resulting distribution of AGN scores is shown in Figure 6.23, which contains only counterpart combinations with scores smaller than 0.3.

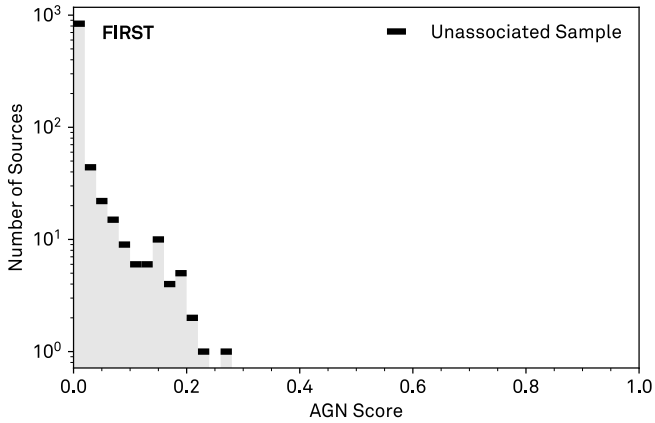

Figure 6.23: Distribution of the AGN scores of the unassociated sample. The scores have been determined with the Random Forest classification model, created with the 3FGL and the FIRST catalogs. All sources feature rather small AGN scores.

For the preparation of the high- and medium-confidence samples, thresholds of the AGN score of 0.64 and 0.05 have been applied to achieve precisions of 0.9 and 0.6, respectively. The number of counterpart candidates per 3FGL source is illustrated in Figure 6.24. For the precision of 0.9, no candidates have been found. The medium-confidence sample comprises 72 candidates in total for 52 3FGL sources. However, multiplicities of up to 5 medium-confidence candidates arise for a single 3FGL source.

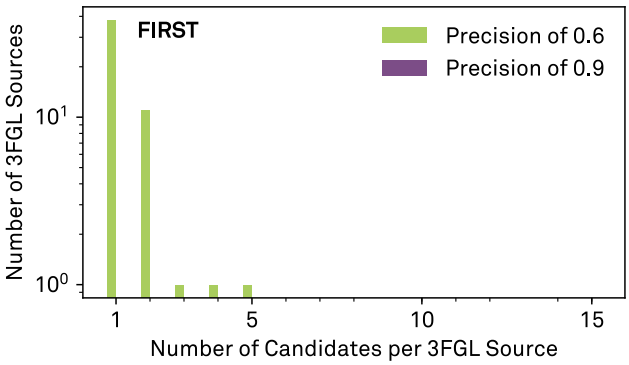

Figure 6.24: Number of 3FGL sources with specific numbers of candidates per 3FGL source, depending on the threshold of the AGN score. For small thresholds, the model associates counterpart candidates to many 3FGL sources, but often multiple candidates for a single 3FGL source. Thus, the association is ambiguous. For large thresholds, no candidates have been found.

6.4 High-Confidence AGN Candidates and their Multi-Wavelength Counterparts

In the previous sections, medium-confidence counterpart candidates have been determined for each of the counterpart catalogs 1SXPS, FIRST and ALLWISE. The thresholds of the AGN scores to obtain these medium-confidence samples have been chosen by demanding a precision of 0.6. To achieve a high-confidence sample, the medium-confidence samples have been joined. The additional criterion for an AGN candidate and its corresponding counterparts is that the distance between the counterparts is smaller than a defined value, i. e. the counterparts are positionally coincident. The maximum distances have been set to the sum of the association radii of the corresponding catalogs (cf. section 5.3), i. e. the maximum allowed distance amounts to 11.7 arcseconds between 1SXPS and ALLWISE counterparts, to 6.0 arcseconds between FIRST and ALLWISE counterparts, and to 10.7 arcseconds between FIRST and 1SXPS counterparts.
Based on the training samples, the performance of this procedure has been determined in terms of the precision. The uncertainty of the precision has been estimated with 1000 bootstrap samples. Table 6.3 summarizes the results for different combinations of the counterpart catalogs. The resulting precisions of more than 0.9 impressively prove the successful creation of a high-confidence sample from medium-confidence samples by taking into account the positional coincidence between the counterparts.

Table 6.3: Precisions for different combinations of counterpart catalogs. The medium-confidence candidates of the individual counterpart catalogs have been merged, and high-confidence candidates have been obtained by requiring a specific positional distance between the candidates of different catalogs. The uncertainty has been determined with 1000 bootstrap samples.

Counterpart Catalogs	Precision
1SXPS / ALLWISE	0.909 ± 0.010
FIRST / ALLWISE	0.976 ± 0.008
FIRST / 1SXPS	0.952 ± 0.012
FIRST / 1SXPS / ALLWISE	0.996 ± 0.004

The resulting high-confidence AGN candidates and their corresponding counterparts are listed in Table 6.4, Table 6.5, Table 6.6 and Table 6.7 for the different combinations of catalogs.

Table 6.4: List of 26 high-confidence AGN candidates and their corresponding 1SXPS and ALLWISE counterpart candidates.

3FGL Name	$\|b\|$ / deg	3FGL Flags	1SXPS Name	WISE Name	Dist. / ″	1SXPS Score	WISE Score
J0240.0–0253	54	0	J024001.7–024323	J024001.89–024321.8	3.1	0.12	0.14
J0529.2+3822	2	5	J052939.6+382325	J052939.39+382327.1	3.7	0.12	0.11
J0641.1+1004	2	1	J064059.3+095521	J064059.34+095520.1	1.4	0.06	0.10
J0704.3–4828	18	0	J070421.6–482645	J070421.81–482647.5	3.0	0.37	0.12
J0746.4–0225	11	0	J074627.1–022550	J074627.02–022549.3	1.3	0.22	0.19
J0802.3–0941	11	0	J080206.8–094111	J080206.74–094116.9	6.3	0.17	0.32
			J080216.1–094207	J080215.90–094210.9	4.7	0.30	0.56
J0826.3–6400	15	0	J082628.0–640414	J082627.86–640415.4	1.9	0.33	0.24
J0838.8–2829	8	0	J083842.4–282830	J083842.77–282830.8	3.2	0.06	0.17
J0900.0+6754	37	256	J090110.9+674202	J090110.67+674203.6	2.9	0.16	0.05
			J090122.1+673952	J090121.66+673955.7	4.2	0.14	0.14
J0919.5–2200	19	0	J092002.8–215837	J092002.74–215834.9	4.2	0.31	0.24
J1038.0–2425	29	0	J103748.2–242840	J103748.10–242845.5	6.0	0.18	0.16
J1132.0–4736	13	1	J113209.5–473857	J113209.26–473853.3	5.0	0.48	0.16
J1240.3–7149	9	0	J124021.3–714857	J124021.21–714857.7	1.1	0.24	0.21
J1249.1–2808	35	0	J124919.2–280832	J124919.31–280834.4	1.8	0.29	0.27
J1251.0–4943	13	0	J125058.9–494450	J125058.64–494453.8	5.7	0.18	0.09
J1315.7–0732	55	0	J131553.0–073301	J131552.97–073302.0	0.3	0.20	0.18
J1626.2–2428c	17	32	J162608.1–242741	J162607.63–242741.8	6.6	0.22	0.17
J1811.3–1927c	0	288	J181131.0–192705	J181130.85–192702.6	2.5	0.26	0.13
J1829.2+3229	19	0	J182937.1+322953	J182937.43+322953.3	3.1	0.14	0.13
J1829.2+2731	17	256	J182913.8+272905	J182913.96+272902.8	3.2	0.15	0.27
J1904.7–0708	6	4	J190444.5–070739	J190444.57–070740.1	3.6	0.19	0.17
J2223.3+0103	45	0	J222329.5+010227	J222329.57+010226.7	1.0	0.09	0.57
J2247.2–0004	50	0	J224710.3–000507	J224710.36–000507.3	0.2	0.10	0.07
J2254.1+6229	3	4	J225456.9+623414	J225455.60+623406.4	11.5	0.07	0.12
			J225456.9+623414	J225456.32+623415.3	3.4	0.07	0.25
J2258.2–3645	64	0	J225814.7–364428	J225815.00–364434.3	8.7	0.13	0.30
J2358.6–1809	75	0	J235837.1–180714	J235836.83–180717.4	7.4	0.34	0.73

Table 6.5: List of 1 high-confidence AGN candidate and its corresponding 1SXPS and FIRST counterpart candidates.

3FGL Name	$\|b\|$ / deg	3FGL Flags	FIRST Name	1SXPS Name	Dist. / ″	FIRST Score	1SXPS Score
J1315.7–0732	55	0	J131552.9–073301	J131553.0–073301	0.1	0.15	0.20

Table 6.6: List of 7 high-confidence AGN candidates and their corresponding FIRST and ALLWISE counterpart candidates.

3FGL Name	$\|b\|$ / deg	3FGL Flags	FIRST Name	WISE Name	Dist. / ″	FIRST Score	WISE Score
J0020.9+0323	59	0	J002050.2+032358	J002050.24+032357.9	0.4	0.11	0.11
J1049.7+1548	60	0	J104939.3+154837	J104939.35+154837.6	0.0	0.12	0.47
J1115.0–0701	49	0	J111511.6–070241	J111511.74–070240.0	1.4	0.16	0.12
J1234.7–0437	58	512	J123444.2–043622	J123444.22–043622.4	0.1	0.24	0.39
J1315.7–0732	55	0	J131552.9–073301	J131552.97–073302.0	0.1	0.15	0.18
J1548.4+1455	47	0	J154824.3+145702	J154824.38+145702.8	0.0	0.13	0.40
J1635.3+4257	42	0	J163534.4+430144	J163534.46+430145.0	0.7	0.15	0.06

Table 6.7: List of 1 high-confidence AGN candidate and its corresponding FIRST, 1SXPS and ALLWISE counterpart candidates.

3FGL Name	FIRST Name	1SXPS Name	WISE Name	Dist. FIRST/ 1SXPS	Dist. FIRST/ WISE	Dist. 1SXPS/ WISE	FIRST Score	1SXPS Score	WISE Score
J1315.7–0732	J131552.9–073301	J131553.0–073301	J131552.97–073302.0	0.1″	0.1″	0.3″	0.15	0.20	0.18

Reviewing the number of medium-confidence candidates at the beginning, and the number of resulting high-confidence candidates, the amount of candidates reduced from several hundreds to a few ten. Simply by the fact that not every catalog comprises observations of the same 3FGL source, the number of candidates reduces. The 1SXPS catalog observed the complete sky on the one hand, but on the other hand the sampling is rather sparse. In contrast, the FIRST catalog's sampling is quite dense, but the sky is only partly covered. These properties of the catalogs are also illustrated in Figure 5.3. Especially, combinations with the FIRST catalog result in only few AGN candidates. This is also the reason that the FIRST / 1SXPS combination – and consequently the FIRST / 1SXPS / ALLWISE combination – resulted in only one AGN candidate – the source 3FGL J1315.7-0732. To possibly confirm the result obtained with the newly developed method, a literature search has been performed.
The source is associated with 2FGL J1315.6–0730 and 1FHL J1315.7-0730, both based on *Fermi*-LAT observations. Landi, R. et al., 2015 studied *Swift*-XRT counterparts of 1FHL sources, and list several counterparts associated with 1FHL J1315.7-0730, reported by further authors. In the radio regime, it is associated with NVSS J131552–073301, and has been detected at multiple frequencies. Based on these measurements, a flat radio spectrum has been determined, according to both Schinzel et al., 2017 and Petrov et al., 2013. Moreover, they stated the association with WISE J131552.98-073301.9. Both the WISE and the NVSS counterpart are positionally coincident with the counterparts obtained in this section, emphasizing the validity of this approach. The source has been further included in searches for blazar or AGN candidates in the UGS sample, i. a. in machine learning approaches by Ackermann et al., 2012a, Doert and Errando, 2014, and Parkinson et al., 2016, all of them suggesting an AGN nature of the source. Lefaucheur and Pita, 2017 provide a list of their brightest found blazar candidates with no analysis flag and the Galactic latitude $|b| > 10^{\circ}$, 3FGL J1315.7-0732 being one of them. They assigned the type BL Lac to the source, as well as Mirabal et al., 2012 and F. Massaro et al., 2013c did.
To further probe this, the so-called *Fermi* blazar sequence has been considered here. Ghisellini et al., 2017 averaged multiple spectral energy distributions of blazars in bins of their gamma-ray luminosity – not only for all blazars, but also for BL Lac objects (BLL) and flat-spectrum radio quasars (FSRQ), separately. These distributions have been described by a simple phenomenological function – a power law, representing a flat radio spectrum, and two broken power laws that can be associated with the synchrotron hump and the inverse Compton hump. The BLL and FSRQ sequences are shown in Figure 6.25, compared to the luminosities, based on the 3FGL, 1SXPS, ALLWISE, and FIRST catalogs, for different redshifts z. The luminosities for a redshift of $z = 0.3$ are compatible with the spectral energy distribution obtained by fitting the phenomenological function to the averaged BLLs with luminosities between 10^{45} and 10^{46} erg s^{-1}. This finding confirms the BLL nature of the source, and additionally provides a redshift estimate. However, it has to be noted that the observations related to the catalogs are not simultaneous, and that the conclusions drawn from the figure must be considered with caution. The observations for the FIRST catalog have been conducted in 1993/94, for the ALLWISE catalog in 2010/11, for the 1SXPS in 2010, and for the 3FGL catalog between 2008 and 2012.

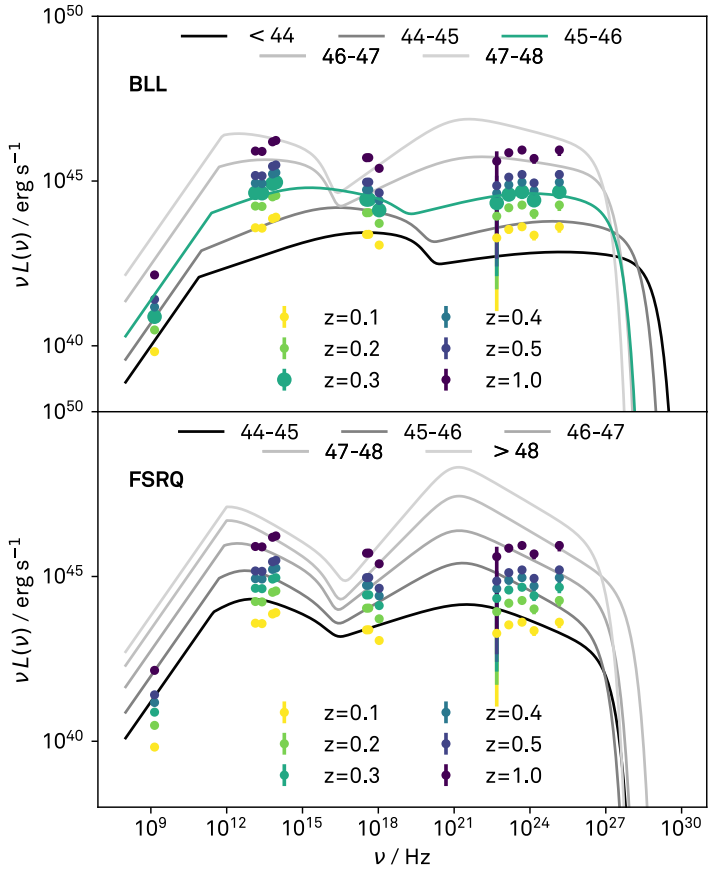

Figure 6.25: Luminosities of the AGN candidate 3FGL J1315.7-0732, compared to the BLL (top) and FSRQ (bottom) sequences. The similarity between the highlighted luminosities (large turquoise dots) and spectral energy distribution (turquoise line) suggests a BLL nature and a redshift of $z = 0.3$.

6.5 Conclusion

In this chapter, a new method to find AGN candidates in the UGS sample and to associate them with counterparts has been developed, and its capability has been proven. For this purpose, a machine learning procedure has been adapted, and differences to typical applications have been presented and discussed. The incorporation of multi-wavelength information has boosted the performance, compared to procedures using only *Fermi*-LAT data. It has also provided counterparts, necessary for an association according to Acero et al., 2015, enabling multi-wavelength studies, such as the creation of spectral energy distributions e. g. to probe emission models. For the determination of the counterparts, catalogs with observations in three different wavebands – radio, infrared and X-ray – have been utilized, and the appropriateness of the individual waveband for the assigned task has been investigated. It turned out that all wavebands are suitable, and that the resulting models achieve high precisions. Based on the individual models for each waveband, samples have been generated for each counterpart catalog, providing high-confidence AGN candidates and (mostly) unambiguous counterparts. These samples are ready to use, but they only comprise information of two wavebands – the gamma-ray waveband and the waveband of the considered counterpart catalog – possibly insufficient e. g. for the obtainment of a spectral energy distribution. In addition, and based on the same models, samples providing a large number of medium-confidence AGN candidates and partly ambiguous counterparts have been generated. These medium-confidence samples have been merged with each other, and the additional restriction of positional coincidence between counterparts of different wavebands has lead to a high-confidence sample of AGN candidates and their counterparts in more than only one waveband. The obtained results are summarized in Table 6.8, revealing impressively the power of exploiting multi-wavelength information.
The source with found counterparts in all wavebands has been investigated further, confirming its AGN nature. Moreover, it has been compared with the BLL and FSRQ sequences, specifying the AGN nature to be a BL Lac object and estimating a redshift of $z = 0.3$.

Table 6.8: Overview of the obtained results. Depending on the considered counterpart catalog, different numbers of AGN candidates are achieved with different precisions, depending on the AGN score threshold. Both with solely one counterpart catalog, as well as with multiple catalogs precisions of more than 0.9 can be obtained.

Counterpart Waveband	Number of AGN Candidates	Precision
Infrared	650	0.598 ± 0.083
X-ray	112	0.605 ± 0.028
Radio	52	0.592 ± 0.044
Infrared	38	0.901 ± 0.081
X-ray	3	0.898 ± 0.032
Radio	0	0.901 ± 0.054
X-ray / Infrared	26	0.909 ± 0.010
Radio / Infrared	7	0.976 ± 0.008
Radio / X-ray	1	0.952 ± 0.012
Radio / X-ray / Infrared	1	0.996 ± 0.004

7 Search for High-Confidence Blazar Candidates and their Multi-Wavelength Counterparts in the 3FGL Catalog

Forty-two.

— Douglas Adams, *The Hitchhiker's Guide to the Galaxy*

Apart from 1010 Unassociated Gamma-ray Sources (UGS), the Third *Fermi*-LAT Source Catalog (3FGL) comprises 573 sources associated to Blazar Candidates of Uncertain type (BCU) (Acero et al., 2015). Assigning blazar classes to the UGS and BCU samples, and associating the UGS with counterparts, extends the sample of blazars available to refine our knowledge of gamma-ray emitting objects and high-energy astroparticle physics. An enlarged blazar sample will improve blazar population studies, and will support the determination of the extragalactic background light.

Extensive work to find blazar candidates and their counterparts among the UGS has been conducted by F. Massaro et al. (e. g. D'Abrusco et al., 2013, Paggi et al., 2013), investigating various wavelengths with multiple procedures. For an overview regarding these studies, the reader is referred to chapter 4. Furthermore, Hassan et al., 2012 and Lefaucheur and Pita, 2017 applied multivariate classifiers to assign blazar classes to the BCU samples, illustrating the prospects of such methods. However, both machine learning procedures provided only blazar classes, and not the corresponding counterparts.

In chapter 6, the capabilities to find AGN candidates and their corresponding multi-wavelength counterparts have already been proven. Moreover, the comparison of the catalog's fluxes of all wavebands with the sequences (Ghisellini et al., 2017) of BL Lac objects (BLL) and flat-spectrum radio quasars (FSRQ) indicated that also the unveiling of the AGN's subclass is possible. Thus, it seems obvious to attempt to refine the classification task – from a 2-class problem, i. e. the discrimination between AGN and non-AGN, to a 3-class problem, i. e. the discrimination between BLL, FSRQ and non-blazar.

As pointed out in the previous chapter, the localization uncertainty of *Fermi*-LAT measurements is in the order of several arcminutes, resulting in hundreds of possible counterparts. Within the developed method, every possible combination between a 3FGL source and all counterparts inside the corresponding 3FGL uncertainty region has been considered. A Random Forest classifier assigns class affiliation scores to these combinations. High scores imply not only a particular class affiliation, but also a combination with a likely counterpart. In addition to the gamma-ray band of the 3FGL catalog, counterpart catalogs of three different wavebands – infrared, X-ray and radio – have been used. For each counterpart catalog, an individual classification model has been created. In the end, the individual models have been joined, and restrictions regarding the distance between the counterparts have been imposed to ensure a positional coincidence and to increase the precision of the predictions. The procedure to join the various catalogs and to generate samples for the developed method are described in chapter 5.

Within the scope of this chapter, similar objectives than in chapter 6 have been pursued. In addition to the UGS sample, also the BCU sample has been classified. Likewise, some uncommon aspects arise that need to be taken into account when interpreting the results. Here, an additional specialty occurs, since a 3-class problem is considered, and most performance measures are binary. The developed method has been realized using the machine learning library *scikit-learn* (Pedregosa et al., 2011).

7.1 Blazar Candidates and Infrared Counterparts

The infrared counterpart candidates, considered in this section, originate from the ALLWISE catalog (cf. subsection 5.1.4). Compared to the training sample in section 6.1, the sample used in this section is smaller – it contains 507 BLLs, 409 FSRQs, and 7851 non-blazars. The BCUs, i. e. blazars of uncertain type, make up a large fraction of the AGNs in the 3FGL catalog. Since they are associated with blazars, but neither unambiguously with BLLs, nor with FSRQs, they cannot be assigned any of the class types and had to be removed from the sample. AGNs, not belonging to the blazar subclass, are now contained in the non-blazar subsample. The generated classification models have been applied to 88 396 counterpart combinations, located in 1 006 3FGL uncertainty regions of the UGS sample, and to 360 sources of the BCU sample.

7.1.1 Feature Selection

The following feature selection is based on the feature set, pre-selected in section 5.2. In a first step, features highly correlated to others have been removed. For this purpose, an initial Random Forest classification model has been generated with the complete pre-selected feature set, and the importance of each feature has been determined. If two features had an absolute Spearman rank correlation coefficient larger than 0.95, the least important feature is removed.
A recursive backward elimination (cf. subsection 3.1.3) has been applied to select the final feature set from the remaining features. For this procedure, a Random Forest classification has been set up with the following hyperparameters:

Number of trees 200
Split selection criteria Information gain
Number of features considered at each node $\sqrt{\text{Number of available features}}$
Minimum number of samples for split 5

For the first model, all available features are considered, and the least important feature is removed. The next model is generated with this new feature set, and the described procedure is repeated recursively until only one feature remains. The performance of each feature set is quantified by the Area Under the Curve (AUC) score (cf. section 3.3) in a 10-fold cross validation (cf. section 3.3). The AUC or the true and false positive rates, respectively, are defined for binary classifications. However, here a 3-class problem is considered. Thus, a *one-vs-rest* strategy (cf. section 3.3) is applied, i. e. each class is validated against the remaining classes, resulting in performance measures for each class separately. The mean and the standard uncertainty of the mean of the AUC scores over the 10 cross validation steps are shown in Figure 7.1, dependent on the number of features in the set the model is built upon.

For each class, the score is rising for an increasing number of features up to a saturation value. This might be due to many similar informative features. As long as no new information, rather than no irrelevant information is added, the AUC score performs rather stable. However, the scores for the FSRQ class are significantly worse than the others, and the classification of FSRQs seems more difficult than the one of BLLs. Nevertheless, the scores of all classes indicate a very good performance of the classification model. Since no obvious score maximum has been formed, a number of 20 features has been selected, achieving a rather stable score for all classes. Figure 7.2 lists the selected features. Out of these features, 15 features have been derived solely from the ALLWISE catalog, 3 from the 3FGL catalog, and 2 features have been calculated with information from both catalogs.

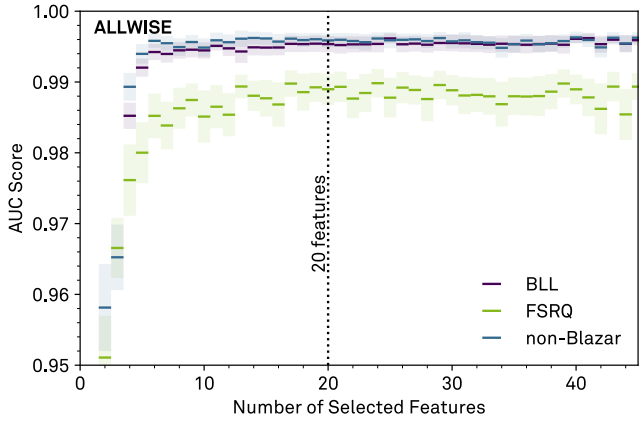

Figure 7.1: Performance of different numbers of selected features for each of the three classes, determined with a recursive backward elimination for the ALLWISE catalog in a 10-fold cross validation. The performance is evaluated with the *one-vs-rest* strategy, and is quantified by the mean of the AUC score, and is saturating above a specific number of features. A number of 20 features has been chosen for the further analysis.

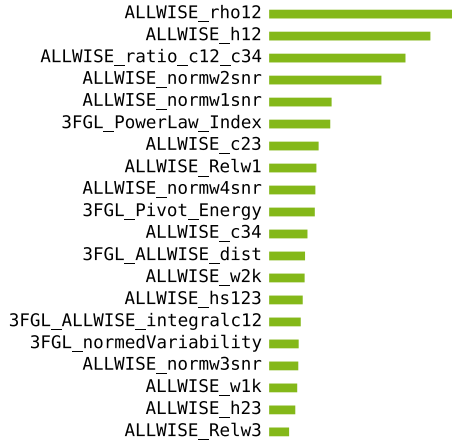

Figure 7.2: List of selected features and their relative feature importance. Out of these, 15 features are based on the ALLWISE catalog, 3 on the 3FGL catalog, and 2 on both catalogs.

7.1.2 Optimization and Performance of the Classification Model

Subsequent to the optimization of the classification model through the feature generation and selection, the hyperparameters of the Random Forest have been optimized, analogous to the optimization presented in subsection 6.1.2. The number of trees in the Random Forest has been set to 200, since it turned out that this number is sufficiently large for the assigned tasks. The split selection criterion, the minimum number of samples in a node required for a split (steering the growth of the tree), and the number of features considered for a split at a node (influencing the randomness and diversity of the trees) have been optimized in a grid search. To quantify the performance of a specific set of hyperparameters, the AUC score have been determined in a 10-fold cross validation, following the *one-vs-rest* strategy. The maximum AUC has been achieved with the following set of hyperparameters:

Number of trees 200
Split selection criteria Information gain
Number of features considered at each node 4
Minimum number of samples for split 3

Based on these parameters and the selected features, the final classification model has been generated. The selected features with their corresponding feature importance are depicted in Figure 7.2. Four features are remarkably more important than the others: The hardness ratio between two infrared wavebands, the correlation coefficient between two wavebands, the ratio between two infrared colors, and the signal-to-noise ratio of one infrared waveband, normalized to the flux in the waveband. The three most important ones are the same as in section 6.1. Thus, the same two-dimensional distributions are shown in Figure 7.3. Similar to the AGNs, the blazars populate a distinctive region in every distribution, enabling to remove large amounts of non-blazars by simple cuts. However, only small differences between the BLL and FSRQ distributions are distinguishable. The physical conclusions that can be drawn are the following: In the infrared regime, blazars are variable, exhibit characteristic curvatures, and positive and distinctive slopes. On average, the spectra of FSRQs appear steeper and less curved than the ones of BLLs.

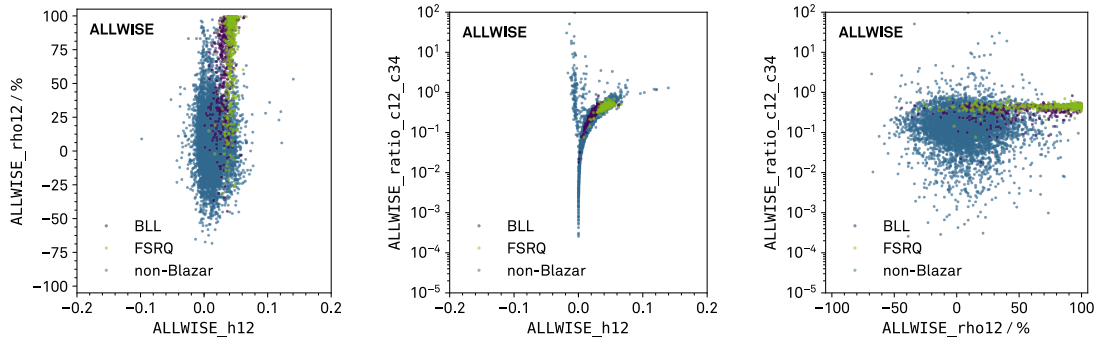

Figure 7.3: Two-dimensional distributions between the three features with the highest feature importance for BLLs, FSRQs and non-blazars. The blazars populate a distinctive region in every distribution.

In Figure 7.4 the distributions of the BLL and FSRQ scores are shown for the training sample. The distribution of the BLL scores feature a maximum at a score of 1 for the BLL class, while the remaining classes have maxima at a score of 0. The FSRQs exhibit large FSRQ scores in contrast to the other classes, but the maximum of the distribution is at a score smaller than 1 and less prominent as the one of the BLLs. In conclusion, the classification between BLLs, FSRQs and non-blazars has been successful. The corresponding performance has been quantified by different measures.

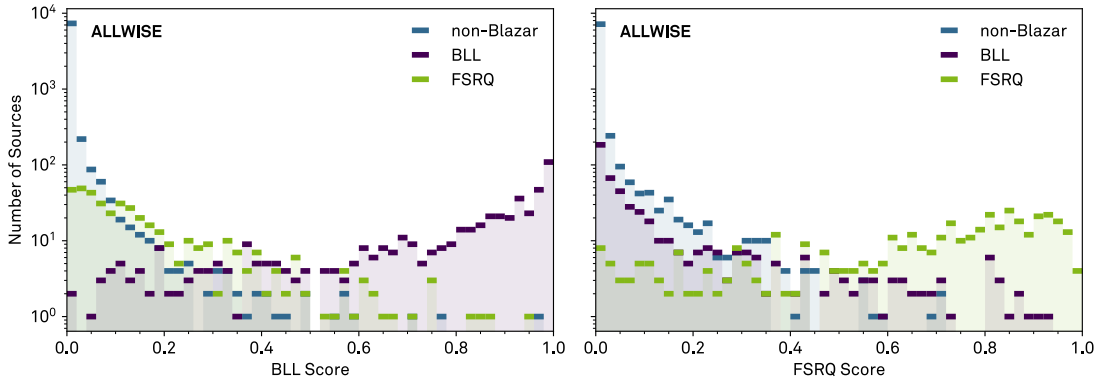

Figure 7.4: Distribution of the BLL (left) and FSRQ (right) scores of the training sample, obtained in a 10-fold cross validation. The scores have been derived from the Random Forest classification model, created with the 3FGL and the ALLWISE catalogs. While the BLL (FSRQ) class features a population at large BLL (FSRQ) scores, the respective remaining classes exhibit small scores.

The Receiver Operating Characteristic (ROC) curve, i. e. the true positive rate against the false positive rate, has been calculated for each class following the *one-vs-rest* strategy, and the averages, minima and maxima of the 10 cross validation curves are shown in Figure 7.5. The Area Under this Curve (AUC) amounts to 0.995 ± 0.004 for BLLs and 0.988 ± 0.008 for FSRQs. The corresponding mean and standard deviation have been derived from the cross validation.

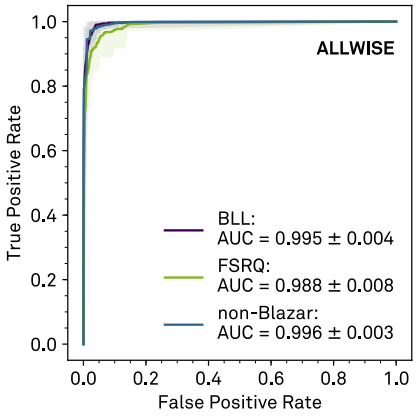

Figure 7.5: ROC curves of the Random Forest classification model for the ALLWISE catalog. The curve has been averaged over 10 curves from the 10-fold cross validation. The shaded area illustrates the minimum and maximum rates. The corresponding AUC and its uncertainty in terms of the standard deviation amounts to 0.995 ± 0.004 for BLLs and 0.988 ± 0.008 for FSRQs.

The further performance measures precision and recall have been determined as a function of the BLL and FSRQ score thresholds, respectively, and are depicted in Figure 7.6. Compared to the precision and recall of the AGN classification in the previous chapter, the measures are less stable, which might be due to the smaller number of sources in a class in the training sample. Moreover, the shapes of the performance measures of the individual cross validation steps are in some cases quite different. This is i. a. related to the different signal-to-background ratios of the cross validation subsamples. Both measures perform slightly better for the BLLs, which is compatible to the score distributions, where the BLL distribution had a sharp maximum at BLL scores of 1, and the FSRQ distribution had a broad maximum at FSRQ scores below 1. This is partly related to the smaller number of FSRQs in the training sample. Another explanation is the similarity between FSRQs and low-frequency peaked BLLs, degrading the performance of the FSRQ classification. The BLL performance has not obviously worsened, since the intermediate-frequency and high-frequency peaked BLLs are discriminable from the other classes.

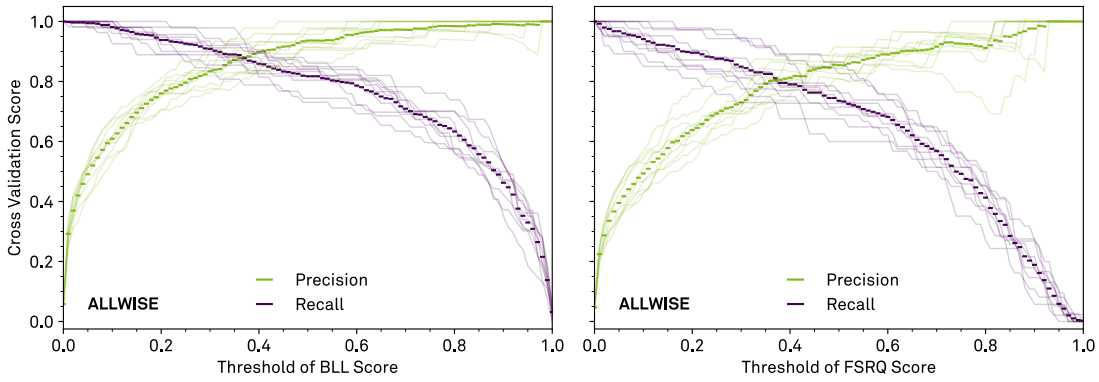

Figure 7.6: Precision and recall for different thresholds of the BLL (left) and FSRQ (right) scores of the Random Forest classification model for the ALLWISE catalog. The averaged precision and recall scores have been determined in a 10-fold cross validation. The individual scores of each cross validation subsample are shown as well.

The BCUs offer the possibility for further tests. The BCUs are blazars, associated with a certain counterpart, i. e. the classification model only predicts the source type. Still, the classification model can also be applied to the counterparts within the 3FGL uncertainty region of a BCU, referred to as the *unassociated BCU sample*. The corresponding test sample contains 6 868 possible ALLWISE counterparts for 360 3FGL sources. Accordingly, the unassociated BCUs should exhibit large non-blazar scores, while the associated BCUs should feature small scores, which is the case. The obtained non-blazar score distribution (Figure 7.7) shows these properties, and thus, proves the capability of the model.

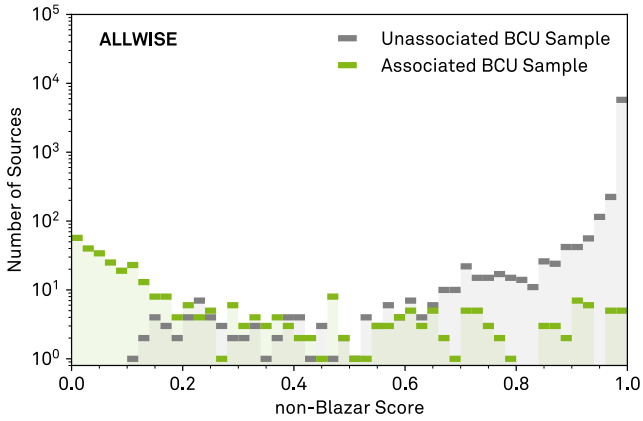

Figure 7.7: Distribution of the non-blazar scores for 3FGL sources associated with BCUs (associated BCU sample), and the counterpart candidates in the corresponding 3FGL uncertainty region (unassociated BCU sample). The shape of the distributions proves the capability of the method.

7.1.3 Application of the Classification Model to the Unassociated Sample and the Uncertain Blazar Sample

The application of the final classification model to the unassociated (UGS) and the uncertain blazar (BCU) samples leads to the BLL and FSRQ score distributions as shown in Figure 7.8. Compared to the training sample, the shapes of the BLL and FSRQ score distributions appear rather different. The score distributions of the UGS sample feature only a very prominent maximum at 0, while those of the BCU sample are roughly uniformly distributed. The differences between the samples are on the one hand due to the fact that the BCU sample contains only associated sources, while the training and the UGS samples comprise multiple possible counterparts. On the other hand, the training sample has been derived from associated sources, based on specific and quite obvious characteristics of a source, which the unassociated sources do not feature, and the classification model cannot consider.

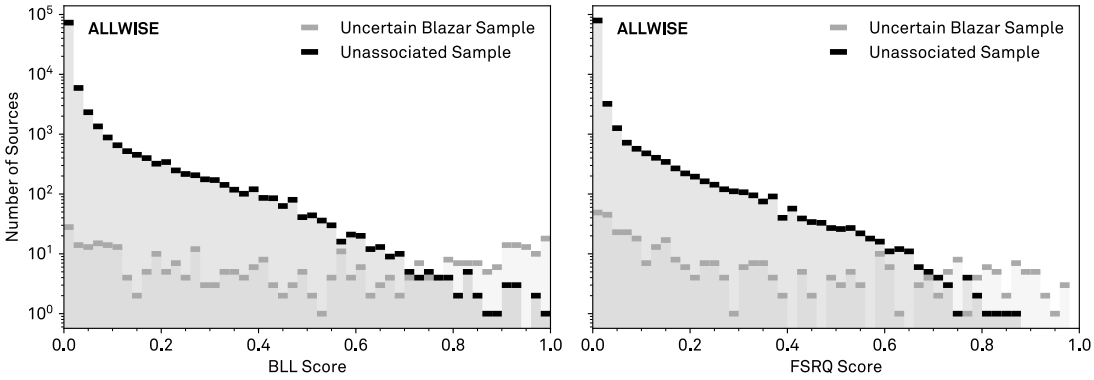

Figure 7.8: Distribution of the BLL (left) and FSRQ (right) scores of the unassociated (UGS) and the uncertain blazar (BCU) sample. The scores have been determined with the Random Forest classification model, created with the 3FGL and the ALLWISE catalog. Multiple sources from both the UGS and the BCU sample feature large BLL and FSRQ scores, respectively.

So far, no thresholds of the predicted scores have been chosen to define the class assignment and the corresponding counterpart. The choice of the thresholds is a trade-off between high precisions and high recalls, and depends on the intended use of the resulting data sample. Two samples for different purposes and with different demands have been generated for each the UGS and the BCU samples. The purpose of one sample – the high-confidence ALLWISE sample – is to provide BLLs and FSRQs and corresponding counterparts of high confidence, which are ready to use for further investigations. This demands a high precision. The second sample – the medium-confidence ALLWISE sample – demands a high recall, i. e. a large number of candidates, leading to a smaller precision than for the first sample. This sample is to be used in combination with similar samples derived from the other counterpart catalogs, and the precision is to be increased by demanding a positional coincidence between the counterparts of different wavelengths.

To obtain an unambiguous prediction for a class affiliation, a minimum score threshold of 0.5 is required. Thresholds of 0.5 achieve precisions of at least 0.85. Thus, thresholds of 0.5 are chosen for the medium-confidence samples, leading to a precision of 0.94 for predicting BLLs, and 0.85 for FSRQs. For the high-confidence samples, precisions of at least 0.9 are demanded, resulting in a threshold of the BLL score of 0.5 and a precision of 0.94, and a threshold of the FSRQ score of 0.63.

Depending on the threshold, multiple counterpart candidates for individual 3FGL sources might be predicted, making the association ambiguous. The number of counterparts per 3FGL source is depicted in Figure 7.9 for the UGS sample. For both the medium- and the high-confidence sample a maximum number of 9 counterparts per BLL candidate, and a maximum number of 6 and 14 counterparts per FSRQ candidate for precisions of 0.90 and 0.85, respectively, have been obtained. The high-confidence sample contains 253 BLL counterpart candidates for 147 3FGL sources, and 45 FSRQ counterpart candidates for 29 sources. The medium-confidence sample comprises 253 BLL counterpart candidates for 147 3FGL sources, and 172 FSRQ counterparts for 77 sources. Table 7.1 lists the high-confidence BLL and FSRQ candidates and the names of their corresponding counterparts for the UGS sample, and gives additional information such as the Galactic latitude and the analysis flag.

Applying the thresholds to the BCU sample, the high-confidence ALLWISE sample contains 167 BLL and 63 FSRQ candidates, and the medium-confidence sample 167 BLL and 91 FSRQ candidates. The high-confidence BLL and FSRQ candidates of the BCU sample are listed in Table 7.2.

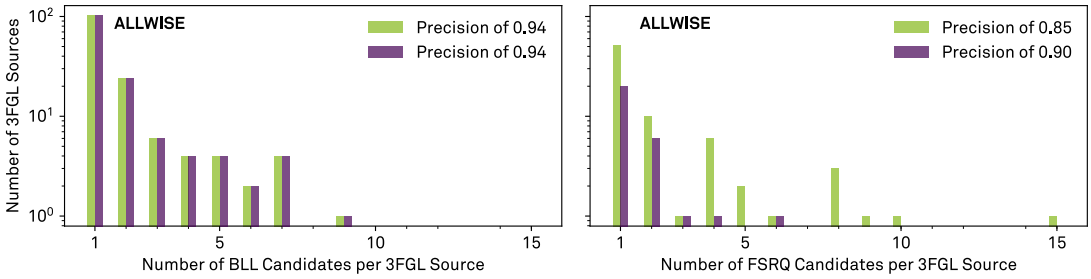

Figure 7.9: Number of 3FGL sources with specific numbers of BLL (left) and FSRQ (right) candidates per 3FGL source, depending on the BLL and FSRQ score, respectively. For the chosen threshold, the models associated multiple counterparts to individual 3FGL sources, making the association ambiguous.

Table 7.1: List of 147 high-confidence BLL and 29 FSRQ candidates and their corresponding ALLWISE counterparts. Out of these, 31 BLL and 4 FSRQ candidates exhibit neither 3FGL analysis flags, nor are located within the Galactic plane.

3FGL Name	$\|b\|$ / deg	3FGL Flags	WISE Name	BLL Score	FSRQ Score
BLL Candidates					
J0004.2+6757	5	25	J000346.72+675811.8	0.76	0.22
J0039.3+6256	0	0	J003920.34+625814.8	0.61	0.01
J0138.5−4612	69	0	J013833.91−461416.9	0.69	0.01
			J013831.31−461023.1	0.55	0.01
J0216.0+0300	54	0	J021600.45+030011.9	0.81	0.13
J0221.2+2518	33	0	J022126.96+251433.6	0.60	0.10
J0224.0+6235	2	16	J022431.45+623334.4	0.55	0.34
J0248.4+5130	7	0	J024834.21+513118.2	0.89	0.04
J0307.3+4916	8	0	J030727.21+491510.6	0.83	0.07
J0312.7−2222	58	0	J031235.70−222117.2	0.96	0.01
J0342.3+3148c	18	48	J034204.35+314711.4	0.66	0.16
			J034158.53+314855.5	0.54	0.28
			J034219.29+314326.8	0.61	0.21
J0420.4−6013	42	0	J042021.42−601403.0	0.77	0.03
J0420.6−3742	45	0	J042025.10−374445.0	0.61	0.28
J0426.3+3510	10	0	J042625.87+351507.6	0.66	0.24
J0429.8+3611c	8	56	J043001.53+360733.0	0.64	0.20
			J043022.20+360435.7	0.69	0.13
J0430.1−3103	43	0	J042958.93−305935.2	0.64	0.03
J0506.3−0357	25	16	J050649.87−035433.1	0.61	0.08
			J050608.27−035021.8	0.68	0.09
J0524.5−6937	33	1	J052407.32−693732.2	0.60	0.10
J0526.4+2247	7	0	J052622.04+224802.0	0.58	0.13
J0529.2+3822	2	5	J052859.47+382053.8	0.56	0.30
J0535.7−0617c	20	49	J053527.91−061415.1	0.51	0.16
			J053547.64−062136.0	0.67	0.20
J0538.9+1646	8	0	J053855.10+164611.5	0.84	0.08
J0541.1+3553	3	1	J054040.99+355641.6	0.58	0.27
J0605.0−0000	10	0	J060458.42+000043.2	0.52	0.00
J0611.5+1957	1	1	J061121.53+200111.6	0.55	0.13
J0634.1+0424	2	4	J063417.05+042735.1	0.52	0.17
J0707.7−1035c	1	32	J070813.44−103218.3	0.51	0.24
J0725.7−0550	5	0	J072545.14−055113.0	0.66	0.15
			J072549.40−054927.6	0.69	0.17
			J072543.69−055205.8	0.81	0.12
			J072537.94−055006.0	0.76	0.14
			J072535.46−055047.3	0.55	0.13
			J072536.35−055146.9	0.82	0.10
			J072546.64−055036.2	0.54	0.04
			J072547.02−055314.7	0.72	0.17
			J072548.08−055338.8	0.93	0.04
J0731.8−3010	5	0	J073151.19−301058.9	0.82	0.04
J0740.6−5230	14	0	J074029.54−522918.1	0.60	0.07
J0746.4−0225	11	0	J074627.02−022549.3	0.86	0.09
J0802.3−0941	11	0	J080206.74−094116.9	0.67	0.17
			J080215.90−094210.9	0.63	0.09
J0823.3−4205c	3	32	J082336.00−420752.8	0.74	0.11
J0832.8−4321	2	1	J083241.10−432034.3	0.59	0.33
J0847.4−4327	0	4	J084733.75−432905.7	0.63	0.19
J0857.6−4258	2	0	J085819.69−425229.8	0.63	0.25
J0903.2−4744	1	16	J090312.06−474935.4	0.60	0.24
J0914.5−4736	1	4	J091442.73−473417.7	0.61	0.21
J0935.1−1736	25	0	J093514.78−173658.8	0.88	0.03
J1002.5−5709	1	20	J100255.94−570638.2	0.51	0.19
J1021.9−5815c	1	49	J102150.74−581638.1	0.55	0.34
			J102138.63−581633.7	0.54	0.30
J1027.8+8253	33	0	J102841.96+825340.5	0.80	0.04
J1031.4−5855	1	0	J103147.69−585338.7	0.57	0.26
J1034.0−5731c	1	304	J103404.03−572346.2	0.51	0.27
J1039.5+7324	41	0	J103939.49+732557.0	0.50	0.38
J1100.2−2044	35	0	J110025.72−205333.3	0.58	0.31
J1101.9−6053	1	20	J110141.85−605406.4	0.56	0.07
J1128.7−6232	1	12	J112926.47−622819.8	0.62	0.18
J1130.7−7800	16	0	J113032.06−780105.4	0.57	0.01
J1132.0−4736	13	1	J113209.26−473853.3	0.74	0.03
J1151.8−6108	1	517	J115148.62−610813.4	0.52	0.24
J1209.9+7607	41	0	J120930.27+760911.9	0.65	0.30
J1240.3−7149	9	0	J124021.21−714857.7	0.91	0.01
J1244.3−5515	8	1	J124415.68−551200.3	0.53	0.21
J1249.1−2808	35	0	J124919.31−280834.4	0.62	0.00
J1315.7−0732	55	0	J131552.97−073302.0	1.00	0.00
J1335.1−5655	5	0	J133451.10−565517.3	0.64	0.17
J1410.9+7406	42	512	J141046.00+740511.2	0.54	0.11
J1441.5−5955c	0	36	J144127.09−595719.2	0.53	0.27
			J144129.39−595824.9	0.55	0.10
			J144205.28−595702.9	0.52	0.07
			J144128.97−595452.5	0.57	0.10
J1457.6−6249	3	0	J145843.47−625341.6	0.64	0.17
			J145837.85−625337.1	0.66	0.20

Table 7.1: – continued from previous page

3FGL Name	$\|b\|$ / deg	3FGL Flags	WISE Name	BLL Score	FSRQ Score
J1513.3–3719	17	0	J151318.66–372011.5	0.91	0.03
J1518.2–5232	4	0	J151807.33–523431.6	0.55	0.06
J1545.0–6641	9	0	J154458.88–664146.9	0.76	0.03
J1548.4+1455	47	0	J154824.38+145702.8	0.74	0.07
J1552.8–5330	0	64	J155307.69–532908.8	0.51	0.15
J1554.4–5315c	0	560	J155434.07–531533.3	0.65	0.08
			J155424.90–531335.2	0.52	0.13
J1626.2–4911	0	0	J162619.83–491506.3	0.68	0.12
			J162626.88–491236.7	0.53	0.04
			J162641.02–491358.0	0.51	0.14
J1626.2–2428c	17	32	J162607.03–242724.5	0.51	0.31
J1628.2–2431c	17	48	J162816.50–243658.3	0.55	0.18
			J162813.78–243249.6	0.60	0.23
J1628.9–4852	0	8	J162859.32–485321.1	0.60	0.13
			J162909.22–485234.3	0.59	0.06
			J162906.39–485104.2	0.54	0.24
			J162901.43–484939.1	0.54	0.21
			J162855.98–484906.0	0.56	0.20
			J162848.55–485602.6	0.57	0.23
			J162908.78–484707.6	0.52	0.19
J1630.3–4346	3	8	J163019.31–434254.1	0.54	0.03
J1636.2–4709c	0	884	J163626.31–471547.8	0.51	0.08
			J163547.75–471313.5	0.61	0.21
			J163630.43–471048.7	0.64	0.17
			J163602.08–471327.8	0.53	0.16
			J163555.50–471329.0	0.54	0.21
J1638.6–4654	0	0	J163849.95–465513.2	0.52	0.09
J1640.4–4634c	0	32	J164030.26–463405.8	0.59	0.10
			J164034.92–463453.0	0.56	0.10
			J164035.07–463535.5	0.56	0.08
			J164026.02–463359.5	0.73	0.06
			J164028.09–463322.3	0.51	0.11
J1646.9–1332	20	0	J164651.78–132848.2	0.57	0.03
J1650.3–4600	1	16	J165032.87–460119.5	0.60	0.18
J1651.5–4626	1	16	J165147.54–462432.3	0.59	0.13
J1701.3–5216	6	16	J170123.25–521722.1	0.50	0.15
J1704.1+1234	29	0	J170409.58+123421.7	0.65	0.25
J1714.6–3327	3	0	J171438.16–332800.3	0.60	0.20
J1715.4–4028	1	0	J171517.89–402652.5	0.52	0.10
			J171516.50–402901.9	0.55	0.09
J1723.6–3646	0	4	J172340.22–364519.1	0.51	0.21
J1726.6–3530c	0	36	J172629.92–353201.9	0.54	0.13
			J172626.84–352803.5	0.56	0.19
J1727.7–2637	5	0	J172737.29–263946.8	0.51	0.34
J1734.7–2930	2	4	J173446.84–292559.1	0.51	0.28
J1736.5–2839	2	2048	J173651.60–284315.3	0.50	0.26
			J173627.38–283907.3	0.62	0.16
J1737.3–3214c	0	36	J173729.95–321621.2	0.70	0.06
			J173720.19–321914.7	0.50	0.07
			J173718.67–321924.1	0.55	0.21
			J173720.12–322003.7	0.58	0.18
			J173718.76–322047.5	0.70	0.10
			J173726.34–320952.1	0.78	0.14
			J173710.59–321427.8	0.71	0.09
J1739.9–3124c	0	60	J173954.98–312318.4	0.54	0.20
			J173959.01–312525.5	0.51	0.20
J1740.5–2843	1	4	J174022.06–284215.3	0.57	0.11
J1740.5–2726	2	0	J174033.34–272942.5	0.50	0.32
			J174020.89–272717.6	0.50	0.20
J1740.5–2642	2	4	J174025.50–264417.4	0.51	0.09
			J174036.22–264550.6	0.56	0.15
			J174049.19–264118.6	0.55	0.18
			J174035.14–264800.2	0.59	0.14
			J174018.79–264310.9	0.69	0.10
			J174013.40–264104.5	0.55	0.17
			J174020.00–263833.4	0.56	0.09
J1744.7–3043	1	28	J174428.92–304237.3	0.53	0.30
J1745.1–3011	1	20	J174521.44–301235.6	0.53	0.18
			J174500.98–301001.0	0.53	0.19
			J174519.47–300911.3	0.52	0.20
			J174513.60–301155.2	0.51	0.30
J1748.3–2815c	0	2084	J174824.43–281630.2	0.56	0.09
			J174819.08–281410.7	0.60	0.09
			J174827.61–281630.8	0.52	0.04
			J174826.46–281629.1	0.54	0.09
			J174823.32–281542.9	0.52	0.06
			J174823.99–281515.5	0.61	0.12
J1749.2–2911	1	68	J174918.21–291442.5	0.62	0.15
			J174922.58–290756.7	0.52	0.16
J1754.0–2930	2	4	J175402.55–293238.4	0.52	0.18
J1758.8–2402	0	20	J175840.19–240558.2	0.62	0.14
			J175901.88–240314.2	0.68	0.10
			J175854.59–240309.6	0.71	0.07

Table 7.1: – continued from previous page

3FGL Name	$\|b\|$ / deg	3FGL Flags	WISE Name	BLL Score	FSRQ Score
J1758.8–2346	0	16	J175903.40–234609.7	0.67	0.05
			J175901.55–234605.6	0.55	0.02
			J175846.46–234538.1	0.64	0.06
			J175845.17–234619.4	0.54	0.03
			J175847.75–234528.5	0.52	0.05
J1759.2–3848	7	0	J175903.30–384739.3	0.51	0.22
J1759.6–2141	1	1	J175949.01–214026.9	0.54	0.15
			J175942.73–214030.7	0.58	0.18
J1800.8–2402	0	0	J180059.22–240302.9	0.62	0.22
J1801.6–2213	0	4	J180146.15–221438.8	0.51	0.06
			J180128.83–221106.5	0.69	0.08
J1806.7–2451	2	8	J180640.98–245404.2	0.50	0.22
			J180648.11–244654.2	0.55	0.13
J1809.2–2725	4	8	J180913.34–271938.2	0.59	0.16
J1810.1–1910	0	20	J180950.41–190836.3	0.68	0.11
			J180949.39–190942.7	0.51	0.28
			J181027.15–190755.1	0.54	0.10
			J180950.11–191056.8	0.58	0.08
			J181023.65–190923.6	0.53	0.06
			J181011.09–190916.4	0.53	0.12
			J181020.04–190614.9	0.59	0.10
J1810.7+5335	27	0	J181037.98+533501.5	0.93	0.04
J1811.3–1927c	0	288	J181133.49–192309.5	0.54	0.24
			J181140.59–192705.9	0.54	0.24
			J181117.55–192949.0	0.52	0.17
J1817.2–1739	1	1	J181704.75–173923.1	0.52	0.07
J1819.5–1345	1	9	J181933.42–134641.0	0.53	0.32
J1820.3–1529c	0	48	J182039.27–152955.9	0.53	0.21
			J182025.98–153013.5	0.51	0.29
J1823.2–1339	0	0	J182317.73–133727.9	0.60	0.19
J1824.3–0620	3	516	J182417.03–061831.8	0.52	0.18
J1829.2–1504	2	1	J182922.48–150759.6	0.54	0.09
J1829.3–0135c	4	52	J182917.11–013358.5	0.62	0.14
J1829.7–1304	1	1	J182933.08–130123.2	0.53	0.24
J1829.9–0203c	4	288	J182927.72–020531.6	0.60	0.23
			J182937.10–015834.8	0.61	0.15
			J182933.76–020638.3	0.55	0.16
			J183013.18–020308.3	0.65	0.19
			J182958.23–015805.3	0.51	0.31
J1831.7–0230	3	0	J183144.01–022824.2	0.53	0.12
J1831.7–0157c	3	48	J183145.49–015702.8	0.50	0.09
J1833.9–0711	1	276	J183339.78–070949.1	0.72	0.14
			J183355.25–071537.2	0.67	0.13
			J183358.69–071130.0	0.64	0.12
			J183347.79–071452.1	0.65	0.17
			J183355.32–071506.0	0.75	0.15
			J183402.56–071126.6	0.59	0.12
J1834.6–0659	1	0	J183437.53–065816.4	0.73	0.05
J1834.8–0630c	1	32	J183455.83–063014.2	0.59	0.04
			J183434.08–063124.4	0.54	0.12
J1837.6–0717	0	25	J183729.70–071715.0	0.67	0.10
			J183724.12–071913.8	0.64	0.12
			J183729.53–072002.9	0.68	0.12
			J183739.93–072137.5	0.51	0.05
J1838.9–0646	0	0	J183838.12–064655.6	0.53	0.12
			J183904.95–064929.3	0.65	0.05
			J183854.52–064906.9	0.63	0.10
J1840.1–0412	1	0	J184015.49–041518.4	0.53	0.15
J1843.0–0359c	0	49	J184259.24–040006.3	0.52	0.09
			J184314.42–040223.7	0.57	0.16
J1844.3–0344	0	0	J184423.38–034256.6	0.61	0.20
			J184430.74–034224.7	0.71	0.10
J1848.4–0141	0	4	J184831.83–013935.4	0.53	0.31
J1849.5–0124c	0	316	J184934.85–012425.2	0.52	0.16
			J184950.42–012516.6	0.52	0.29
J1857.2+0059	1	20	J185705.40+005949.6	0.58	0.14
J1900.3+0411c	0	293	J190039.17+041250.6	0.52	0.18
J1902.3–3702c	18	32	J190141.64–365953.2	0.59	0.23
J1922.2+2313	4	4	J192212.25+231520.4	0.80	0.08
J1934.2+6002	18	0	J193419.62+600139.5	0.97	0.00
J1949.6+2355	1	0	J195000.70+235439.9	0.73	0.19
			J194953.06+235514.9	0.60	0.19
J1951.8–1102	18	16	J195129.23–110219.9	0.83	0.11
J2028.5+4040c	1	52	J202850.30+403549.8	0.79	0.12
			J202847.69+403349.1	0.51	0.12
			J202839.51+403339.5	0.60	0.26
			J202849.74+403832.5	0.51	0.27
J2037.4+4132c	0	32	J203726.01+413540.7	0.61	0.23
			J203701.28+413119.4	0.51	0.11
			J203707.64+413229.0	0.50	0.06
J2044.0+1035	19	0	J204351.62+103407.2	0.60	0.28
J2046.7–4259	39	0	J204725.08–425321.4	0.57	0.07

Table 7.1: – continued from previous page

3FGL Name	$\|b\|$ / deg	3FGL Flags	WISE Name	BLL Score	FSRQ Score
J2048.8+4436	1	2048	J204859.76+444127.7	0.55	0.28
J2053.9+2922	10	1	J205350.72+292314.5	0.52	0.03
J2221.7+6318	5	1	J222053.70+631206.8	0.53	0.31
J2223.3+0103	45	0	J222329.57+010226.7	0.92	0.04
J2258.2–3645	64	0	J225815.00–364434.3	0.79	0.08
J2321.3+5113	9	0	J232127.44+511115.1	0.68	0.17
J2329.8+6102	0	0	J232938.24+610114.0	0.71	0.03
J2358.5+3827	23	0	J235825.19+382856.5	0.93	0.04
J2358.6–1809	75	0	J235836.83–180717.4	0.83	0.05
FSRQ Candidates					
J0221.1+6059	0	0	J022102.21+610549.2	0.11	0.68
J0221.1+6059	0	0	J022144.66+610554.4	0.07	0.64
J0529.2+3822	2	5	J052923.74+381824.2	0.07	0.78
J0638.4+5704	21	1	J063825.73+570150.6	0.15	0.84
J0647.8+1751	7	1	J064826.73+174422.5	0.08	0.69
J0838.9–4056c	0	40	J083825.90–410302.1	0.11	0.67
			J083903.21–404807.8	0.05	0.80
			J083904.93–410334.5	0.12	0.67
			J083908.49–410036.9	0.15	0.69
			J083913.14–405745.7	0.16	0.78
			J083850.32–405123.0	0.19	0.69
J0847.4–4327	0	4	J084654.45–432917.1	0.19	0.72
			J084655.23–432914.9	0.31	0.64
J0851.4–4244	1	20	J085126.34–424602.8	0.28	0.69
J0854.4–4404	1	17	J085401.16–435918.2	0.21	0.65
J0905.6–4917	1	2049	J090530.80–492011.2	0.10	0.74
J0914.5–4736	1	4	J091450.96–474135.2	0.22	0.66
			J091450.19–473506.5	0.24	0.63
J0935.2+0903	40	0	J093444.73+090352.9	0.16	0.77
J1013.6–5734	1	16	J101317.31–573350.0	0.17	0.72
J1244.3–5515	8	1	J124331.92–551001.3	0.11	0.65
J1259.5–3231	30	0	J125949.83–322328.8	0.06	0.86
J1625.2–2845	14	0	J162516.45–284915.6	0.14	0.76
J1705.5–4128c	0	41	J170522.66–413316.3	0.11	0.71
			J170519.99–412755.1	0.15	0.71
			J170520.45–412432.3	0.13	0.66
J1727.4+0634	22	0	J172644.95+063918.6	0.08	0.79
J1829.9–0203c	4	288	J182959.03–020157.4	0.20	0.63
J1840.5+6116	25	17	J184035.34+611407.0	0.11	0.85
J1855.4+0454	1	8	J185518.56+045539.5	0.22	0.65
J1902.3–3702c	18	32	J190214.65–370033.1	0.30	0.65
J2024.6+3747	0	17	J202433.45+375131.4	0.26	0.65
			J202432.55+374949.3	0.21	0.66
J2032.5+4032	0	0	J203247.70+403545.3	0.18	0.76
J2034.4+3833c	1	2100	J203427.12+383633.2	0.12	0.65
J2058.0+4347	1	1	J205805.94+435230.9	0.25	0.68
			J205729.82+434721.5	0.26	0.63
J2135.6+5742	4	12	J213531.98+574839.4	0.12	0.72
J2147.2+4730c	5	56	J214725.45+473158.7	0.18	0.64
			J214759.91+473546.5	0.14	0.76
			J214755.68+473711.3	0.14	0.64
			J214756.39+473709.8	0.11	0.65
J2233.1+6542	7	0	J223348.53+653907.5	0.20	0.64
J2254.1+6229	3	4	J225336.50+623334.3	0.07	0.70
			J225344.04+623059.6	0.26	0.69

Table 7.2: List of 167 high-confidence BLL and 63 FSRQ candidates and their corresponding ALLWISE counterparts. Out of these, 130 BLL and 42 FSRQ candidates exhibit neither 3FGL analysis flags, nor are located within the Galactic plane.

3FGL Name	$\|b\|$ / deg	3FGL Flags	WISE Name	BLL Score	FSRQ Score
BLL Candidates					
J0003.8−1151	71	0	J000404.91−114858.3	0.55	0.28
J0014.6+6119	1	0	J001448.79+611743.6	0.65	0.16
J0015.7+5552	7	0	J001540.13+555144.7	0.89	0.05
J0031.3+0724	55	0	J003119.71+072453.4	0.84	0.04
J0039.1+4330	19	0	J003908.15+433014.5	0.80	0.14
J0040.5−2339	86	0	J004024.90−234000.7	0.79	0.07
J0049.4−5401	63	0	J004948.86−540243.5	0.84	0.07
J0051.2−6241	54	0	J005116.64−624204.3	0.98	0.01
J0103.7+1323	49	0	J010345.74+132345.3	0.96	0.02
J0132.5−0802	69	0	J013241.13−080404.9	0.91	0.01
J0134.5+2638	35	0	J013428.19+263843.0	0.99	0.01
J0137.8+5813	4	0	J013750.47+581411.3	0.99	0.00
J0139.9+8735	25	0	J013912.38+873757.6	0.90	0.05
J0146.4−6746	49	0	J014554.85−674648.7	0.58	0.12
J0147.0−5204	63	0	J014648.58−520233.5	0.51	0.03
J0148.3+5200	10	0	J014820.33+520204.9	0.96	0.03
J0156.9−4742	66	0	J015646.03−474417.3	0.81	0.08
J0204.2+2420	36	0	J020421.54+241750.7	0.84	0.02
J0211.2−0649	62	0	J021116.95−064419.9	0.85	0.05
J0213.1−2720	72	0	J021255.26−271818.7	0.68	0.19
J0223.3+6820	7	0	J022304.52+682154.8	0.92	0.04
J0224.1−7941	37	0	J022338.74−794015.0	0.70	0.22
J0228.0+2248	35	0	J022744.35+224834.1	0.81	0.15
J0232.9+2606	31	0	J023256.36+260942.9	0.87	0.07
J0241.3+6542	5	0	J024121.75+654311.7	0.92	0.01
J0255.8+0532	46	0	J025549.51+053355.0	0.73	0.23
J0301.4−1652	58	0	J030116.62−165245.0	0.87	0.08
J0318.7+2134	30	0	J031845.66+213436.6	0.84	0.06
J0331.3−6155	46	0	J033118.48−615528.7	0.97	0.02
J0332.0+6308	6	0	J033153.90+630814.1	0.78	0.19
J0333.4+7853	18	0	J033344.58+785028.5	0.94	0.00
J0333.4+4003	13	0	J033346.81+400638.4	0.74	0.15
J0338.5+1303	33	0	J033829.27+130215.5	0.90	0.03
J0339.2−1738	50	0	J033913.70−173600.8	0.76	0.03
J0343.3−6443	44	0	J034320.52−644256.0	0.99	0.01
J0352.9+5655	2	0	J035309.54+565430.7	0.93	0.03
J0431.6+7403	17	0	J043145.06+740326.6	0.95	0.01
J0434.6+0921	25	0	J043440.98+092348.6	0.70	0.15
J0453.2+6321	12	0	J045312.46+632117.9	0.69	0.16
J0505.5−1558	30	0	J050541.57−155838.1	0.95	0.02
J0506.9−5435	37	0	J050657.81−543503.7	0.56	0.10
J0512.9+4038	1	0	J051252.53+404143.7	0.57	0.18
J0521.7+0103	19	0	J052140.82+010255.5	0.92	0.05
J0525.6−6013	34	0	J052542.42−601340.2	0.62	0.08
J0532.0−4827	33	0	J053158.61−482735.9	0.72	0.24
J0553.5−2036	22	0	J055333.12−203418.9	0.94	0.04
J0602.2+5314	15	0	J060200.44+531600.1	0.71	0.04
J0620.4+2644	6	0	J062040.05+264331.9	0.94	0.03
J0622.9+3326	9	0	J062252.21+332610.2	0.58	0.32
J0623.3+3043	8	0	J062316.03+304457.7	0.84	0.12
J0640.0−1252	8	4	J064007.18−125315.0	0.90	0.01
J0644.3−6713	26	0	J064428.03−671257.3	0.73	0.27
J0647.0−5134	22	0	J064710.04−513547.7	0.82	0.03
J0651.3+4014	17	0	J065105.41+401338.1	0.93	0.05
J0653.6+2817	13	0	J065344.26+281547.5	0.97	0.01
J0700.2+1304	8	0	J070014.31+130424.4	0.98	0.02
J0700.3−6310	23	0	J065958.68−631238.6	0.85	0.08
J0703.4−3914	15	0	J070312.65−391418.8	0.64	0.31
J0708.9+2239	14	0	J070858.28+224135.4	0.94	0.04
J0723.2−0728	4	0	J072259.68−073134.7	0.74	0.15
J0725.8−0054	7	0	J072550.63−005456.5	0.75	0.23
J0728.0+4828	26	0	J072759.84+482720.3	0.67	0.24
J0730.5−0537	6	0	J073028.42−053546.4	0.89	0.09
J0733.5+5153	27	0	J073326.79+515355.9	0.61	0.03
J0742.4−8133c	25	32	J074221.14−813136.2	0.86	0.02
J0746.6−4756	11	0	J074642.31−475455.0	0.90	0.06
J0748.8+4929	29	0	J074837.76+493041.0	0.60	0.19
J0749.4+1059	18	0	J074927.38+105733.1	0.59	0.36
J0804.0−3629	3	4	J080405.32−362917.4	0.99	0.00
J0827.2−0711	18	0	J082706.16−070845.9	0.93	0.01
J0829.6−1137	16	0	J082939.17−114103.6	0.88	0.07
J0841.3−3554	4	0	J084121.63−355505.9	0.95	0.04
J0842.0−6055	11	0	J084226.56−605350.4	0.52	0.45
J0849.5−2912	9	0	J084922.10−291150.4	0.57	0.07
J0904.8−5734	7	0	J090453.18−573505.8	0.58	0.31
J0917.3−0344	30	0	J091714.61−034314.2	0.96	0.00
J0947.1−2542	21	0	J094709.52−254059.9	0.99	0.00
J0953.1−7657c	18	32	J095304.35−765802.0	0.52	0.02

Table 7.2: – continued from previous page

3FGL Name	$\|b\|$ / deg	3FGL Flags	WISE Name	BLL Score	FSRQ Score
J1007.8+0026	43	0	J100811.44+002959.9	0.62	0.13
J1009.0–3137	20	16	J100850.54–313905.5	0.62	0.15
J1021.8+8023	35	0	J102202.09+802350.0	0.59	0.36
J1040.4+0615	53	0	J104031.62+061721.7	0.83	0.14
J1052.8–3741	19	0	J105258.09–374318.6	0.94	0.04
J1125.0–2101	38	0	J112508.62–210105.9	0.95	0.03
J1154.0–3243	29	0	J115406.16–324243.0	0.82	0.10
J1155.4–3417	27	0	J115520.52–341719.7	0.97	0.02
J1156.7–2250	38	0	J115633.22–225004.2	0.57	0.05
J1159.6–0723	53	0	J115931.87–072359.4	0.92	0.04
J1200.8+1228	71	0	J120040.03+123103.2	0.67	0.12
J1208.2–7810c	15	32	J120818.33–780948.5	0.80	0.14
J1218.8–4827	14	0	J121902.26–482627.9	0.77	0.15
J1225.7–7314	10	1	J122535.34–731340.1	0.60	0.07
J1229.8–5305	10	0	J122939.88–530332.1	0.80	0.09
J1238.2–1958	43	0	J123824.39–195913.8	0.72	0.16
J1238.3–4543	17	0	J123806.03–454129.6	0.93	0.04
J1244.3–4955	13	8	J124422.98–495422.3	0.92	0.03
J1254.1–2203	41	0	J125422.46–220413.7	0.79	0.13
J1256.1–5919	4	8	J125604.90–591943.9	0.80	0.14
J1256.3–1146	51	0	J125615.95–114637.3	0.84	0.04
J1259.0–2310	40	0	J125908.45–231038.6	0.98	0.01
J1259.8–3749	25	0	J125949.80–374858.1	0.89	0.04
J1304.2–2411	39	0	J130416.69–241216.6	1.00	0.00
J1307.6–4300	20	0	J130737.98–425938.9	0.99	0.01
J1319.6+7759	39	0	J131921.26+775822.2	0.99	0.01
J1322.6–1619	46	0	J132243.65–161724.7	0.76	0.18
J1340.6–0408	57	0	J134042.01–041006.8	0.93	0.04
J1342.7+0945	69	0	J134240.02+094752.4	0.95	0.01
J1344.5–3655	25	0	J134423.77–365628.7	0.91	0.03
J1346.9–2958	31	0	J134706.88–295842.4	0.99	0.01
J1356.3–4029	21	0	J135625.49–402821.8	0.95	0.03
J1406.0–2508	35	0	J140609.60–250809.2	0.99	0.01
J1418.9+7731	39	0	J141900.31+773229.3	0.81	0.06
J1427.8–3215	26	0	J142750.21–321517.0	0.77	0.07
J1434.6+6640	47	0	J143441.46+664026.5	0.92	0.01
J1511.8–0513	43	0	J151148.56–051346.9	0.94	0.01
J1518.0–2732	25	0	J151803.59–273131.1	0.97	0.01
J1539.8–1128	34	0	J153941.19–112835.3	0.91	0.02
J1543.5+0451	43	0	J154333.92+045219.3	0.55	0.03
J1559.7+8512	30	0	J160031.75+850949.2	0.56	0.14
J1607.9–2040	23	0	J160756.90–203943.5	0.70	0.15
J1612.4–3100	15	0	J161219.99–305938.6	0.98	0.01
J1626.4–7640	19	0	J162638.15–763855.5	0.88	0.11
J1628.2+7703	34	0	J162833.06+770650.0	0.54	0.31
J1636.7+2624	40	0	J163651.46+262656.7	0.91	0.05
J1640.9+1142	34	0	J164058.89+114404.2	0.57	0.07
J1647.4+4950	40	0	J164734.91+495000.5	0.64	0.35
J1716.7–8112	23	0	J171711.50–811509.6	0.65	0.06
J1736.0+2033	25	0	J173605.25+203301.1	0.94	0.02
J1740.4+5347	32	0	J174036.52+534623.8	0.90	0.07
J1807.8+6427	29	0	J180732.18+642926.3	0.83	0.08
J1819.1+2134	16	0	J181905.22+213233.8	0.79	0.09
J1841.2+2910	15	0	J184121.73+290940.8	0.97	0.01
J1844.3+1547	9	0	J184425.36+154645.8	0.99	0.00
J1848.1–4230	17	0	J184806.17–423026.4	0.96	0.01
J1855.1–6008	24	0	J185451.68–600923.5	0.80	0.09
J1913.5–3631	20	1	J191320.90–363019.3	0.73	0.20
J1939.6–4925	28	0	J193946.08–492538.1	0.91	0.04
J1942.7+1033	6	0	J194247.48+103327.0	0.93	0.04
J1949.0+1312	6	0	J194855.22+131439.9	0.81	0.13
J1954.9–5640	31	0	J195502.86–564028.8	0.98	0.00
J1959.8–4725	31	0	J195945.66–472519.3	0.99	0.00
J2018.5+3851	2	0	J201831.03+385119.5	0.71	0.24
J2023.2+3154	3	0	J202319.01+315302.2	0.60	0.36
J2024.4–0848	25	0	J202429.37–084804.6	0.91	0.02
J2031.0+1937	12	0	J203057.12+193612.9	0.58	0.04
J2040.2–7115	34	0	J204008.27–711459.9	0.87	0.00
J2041.9–7318	34	0	J204201.93–731913.5	0.81	0.04
J2109.1–6638	38	0	J210851.80–663722.7	0.93	0.04
J2112.7+0819	26	0	J211243.00+081835.2	0.93	0.03
J2133.3+2533	19	0	J213314.36+252859.0	0.79	0.08
J2133.8+6648	11	4	J213349.20+664704.3	0.96	0.03
J2142.2–2546	48	0	J214215.92–255126.5	0.55	0.27
J2144.2+3132	16	0	J214415.22+313339.2	0.88	0.08
J2156.0+1818	28	0	J215601.64+181837.1	0.96	0.01
J2200.9–2412	52	1	J220036.67–241427.7	0.56	0.35
J2212.3–7039	41	0	J221156.23–703914.8	0.56	0.35
J2212.6+2801	23	0	J221239.11+275938.4	0.78	0.10
J2213.6–4755	53	0	J221330.35–475425.2	0.84	0.04
J2250.3–4206	61	0	J225022.22–420613.4	0.99	0.01
J2250.7–2806	63	0	J225044.48–280639.1	0.60	0.36
J2251.5–4928	58	0	J225128.70–492910.8	0.74	0.05

Table 7.2: – continued from previous page

3FGL Name	$\|b\|$ / deg	3FGL Flags	WISE Name	BLL Score	FSRQ Score
J2258.1–8248	33	0	J225759.08–824652.5	0.93	0.02
J2316.8–5209	59	0	J231701.72–521001.4	0.90	0.02
J2336.5+2356	36	0	J233642.10+235529.1	0.68	0.01
J2336.5–7620	40	0	J233627.55–762037.9	0.84	0.15
J2346.7+0705	52	0	J234639.93+070506.8	0.99	0.01
J2348.4–5100	63	0	J234853.10–510314.0	0.74	0.23
FSRQ Candidates					
J0045.2–3704	80	0	J004512.06–370548.5	0.13	0.84
J0050.0–4458	72	0	J004916.62–445711.2	0.19	0.71
J0059.1–5701	60	4	J005846.59–565911.4	0.24	0.75
J0156.3+3913	22	0	J015631.41+391430.8	0.07	0.86
J0205.0+1510	44	0	J020450.42+151410.9	0.08	0.86
J0210.7–5101	62	0	J021046.20–510101.8	0.13	0.87
J0216.6–1019	64	0	J021638.87–101702.9	0.18	0.81
J0218.9+3642	23	0	J021850.04+364042.3	0.07	0.88
J0225.2–2602	69	0	J022510.88–260318.7	0.07	0.82
J0301.8–7157	42	0	J030138.47–715634.5	0.09	0.75
J0356.3–6948	40	0	J035730.10–694844.8	0.03	0.88
J0358.7+0633	34	0	J035827.14+062919.3	0.07	0.83
J0426.3+6827	13	0	J042650.06+682552.9	0.30	0.66
J0427.3–3900	44	0	J042721.64–390100.1	0.02	0.98
J0456.3+2702	10	0	J045617.36+270221.1	0.03	0.95
J0519.5+0852	16	1	J051910.80+084856.6	0.28	0.67
J0529.8–7242	32	0	J052930.02–724528.6	0.12	0.86
J0542.5–0907c	19	41	J054255.87–091330.8	0.09	0.88
J0641.8–0319	4	4	J064151.12–032048.4	0.06	0.86
J0643.4–5358	23	0	J064320.12–535846.1	0.10	0.90
J0647.6–6058	24	0	J064740.86–605805.1	0.25	0.72
J0700.0+1709	10	4	J070001.49+170921.9	0.08	0.91
J0734.3–7709	24	0	J073443.41–771113.3	0.15	0.75
J0744.1–3804	7	8	J074344.81–380356.4	0.08	0.91
J0747.2–3311	4	0	J074719.67–331047.0	0.27	0.73
J0748.5+7910	29	0	J075043.29+790917.1	0.21	0.70
J0816.7–2421	6	4	J081640.41–242106.5	0.08	0.88
J0849.9–3540	5	0	J084945.62–354101.1	0.25	0.74
J0852.6–5756	9	0	J085238.73–575529.2	0.22	0.77
J0858.1–1951	17	0	J085805.36–195036.8	0.02	0.97
J0904.8–3516	8	0	J090442.38–351424.3	0.06	0.81
J0939.2–1732	25	0	J093919.19–173135.6	0.04	0.88
J0956.7–6441	8	1	J095612.14–643929.1	0.17	0.76
J1016.0–0635	40	0	J101626.97–063625.2	0.22	0.68
J1016.1+5555	50	0	J101544.43+555100.6	0.19	0.71
J1026.5+7423	39	16	J102724.13+742826.0	0.17	0.80
J1038.9–5311	5	0	J103840.66–531142.9	0.01	0.82
J1109.4–4815	11	0	J110918.86–481519.2	0.22	0.76
J1138.2+4905	64	0	J113802.06+485857.5	0.06	0.81
J1159.3–2226	39	0	J115911.26–222836.8	0.15	0.80
J1251.0–0203	61	0	J125118.37–020207.6	0.11	0.75
J1304.3–5535	7	0	J130349.21–554031.6	0.26	0.69
J1328.9–5607	6	0	J132901.15–560802.5	0.09	0.91
J1338.6–2403	38	0	J133901.74–240113.9	0.24	0.73
J1351.7–2913	32	0	J135146.84–291217.5	0.11	0.89
J1419.5–0836	48	0	J141922.55–083832.0	0.26	0.69
J1424.6–6807	7	0	J142455.53–680758.1	0.04	0.88
J1508.7–4956	7	16	J150838.83–495302.1	0.08	0.82
J1512.3+8005	35	0	J151032.73+800005.2	0.22	0.75
J1514.8–3623	18	0	J151440.84–361705.0	0.23	0.72
J1604.4–4442	6	0	J160431.00–444132.1	0.19	0.81
J1650.2–5044	4	0	J165016.62–504448.3	0.24	0.69
J1704.0+7646	32	0	J170357.88+764610.1	0.05	0.88
J1723.7–7713	22	0	J172350.80–771350.5	0.07	0.93
J1825.2–5230	17	0	J182513.78–523058.2	0.27	0.73
J1830.0–4439	15	0	J183000.86–444111.4	0.02	0.97
J1858.4–2509	13	0	J185819.06–251050.5	0.06	0.83
J2000.1+4212	6	0	J195958.76+421346.5	0.21	0.78
J2007.7–7728	31	0	J200706.46–773042.2	0.09	0.82
J2103.9–3546	41	0	J210352.95–354620.2	0.08	0.92
J2106.1+2505	15	0	J210639.76+250057.4	0.02	0.91
J2149.6+1915	26	0	J214947.24+192046.4	0.04	0.85
J2328.4–4034	68	0	J232819.26–403509.8	0.20	0.79

7.2 Blazar Candidates and X-Ray Counterparts

In this section, X-ray counterpart candidates from the 1SXPS catalog (cf. subsection 5.1.2) have been considered. The classification model has been generated with a training sample comprising 399 BLLs, 313 FSRQs, and 1233 non-blazars. This model has been applied to 766 counterpart combinations, located in the uncertainty regions of 166 3FGL sources of the UGS sample, and to 163 sources of the BCU sample.

7.2.1 Feature Selection

The feature selection, presented in this section, has been conducted with the similar procedure as performed in subsection 7.1.1. The mean and the standard deviation of the AUC score, obtained in the cross validation, are visualized in Figure 7.10. The score rises, if more features are included in the set, and reaches a saturation value at a particular number of features. Since no clear maximum has been formed, a number of 20 features has been chosen. Out of these, 8 features originate from the 3FGL catalog, 7 from the 1SXPS catalog, and 5 from both catalogs. The AUC scores of this feature set suggest a successful classification. The respective features are listed in Figure 7.11. According to the AUC, the classification of BLLs performs better than the one of FSRQs. This has already been discussed in the previous section.

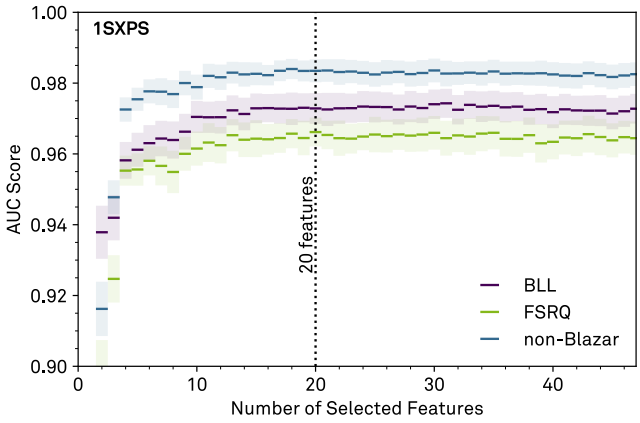

Figure 7.10: Performance of different numbers of selected features for each of the three classes, determined with a recursive backward elimination for the 1SXPS catalog in a 10-fold cross validation. The performance is evaluated with the *one-vs-rest* strategy, and is quantified by the mean of the AUC score, and is saturating above a specific number of features. A number of 20 features has been chosen for the further analysis.

7.2.2 Optimization and Performance of the Classification Model

The optimization and performance evaluation in this section have been performed analogously to the ones in subsection 7.1.2. The optimization of the classification model based on the 1SXPS catalog has lead to the following hyperparameters:

Number of trees 200
Split selection criteria Information gain
Number of features considered at each node 3
Minimum number of samples for split 8

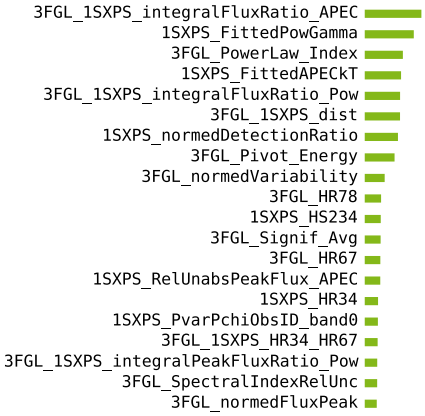

Figure 7.11: List of selected features and their relative feature importance. Out of these, 7 features are based on the 1SXPS catalog, 8 on the 3FGL catalog, and 5 on both catalogs.

With these optimized hyperparameters, the final Random Forest model has been created. The corresponding feature importance is illustrated in Figure 7.11. No feature appears to be significantly more important than the others. The three most important features are the ratio between the integral X-ray and gamma-ray fluxes, and the power law index of the X-ray spectrum and the one of the gamma-ray spectrum, which are the same as in subsection 6.2.2. The corresponding two-dimensional distributions are depicted in Figure 7.12. Although the fit of the spectral slope of the X-ray spectrum mostly fails for non-blazars, it is included in the classification, since this is nevertheless a valuable information. The spectral indices of both the X-ray and the gamma-ray regime are perfectly suited to discriminate between BLLs and FSRQs. Consistent with the previous chapter, the following physical conclusions can be drawn: In the X-ray regime, the connection between properties in the X-ray and gamma-ray regime are important, and the flux ratio of the two energy regimes in connection with the spectral slope in the gamma-ray regime creates discriminable BLL, FSRQ and non-blazar populations. BLLs exhibit harder spectra than FSRQs in the gamma-ray regime, while FSRQs feature steeper spectra in the X-ray regime.

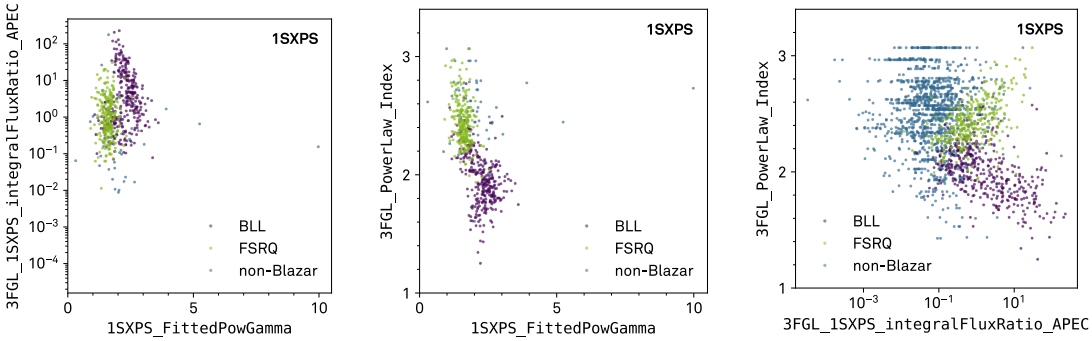

Figure 7.12: Two-dimensional distributions between the three features with the highest feature importance for BLLs, FSRQs and non-blazars. BLLs, FSRQs and non-blazars populate distinctive regions in every distribution.

The BLL and FSRQ score distributions for the three classes in the training sample are shown in Figure 7.13. Consistently, the BLL score distribution features a sharp maximum at 1 for BLLs, while the FSRQ score distribution has a wide maximum below 1 for FSRQs. The remaining classes exhibit always small scores.

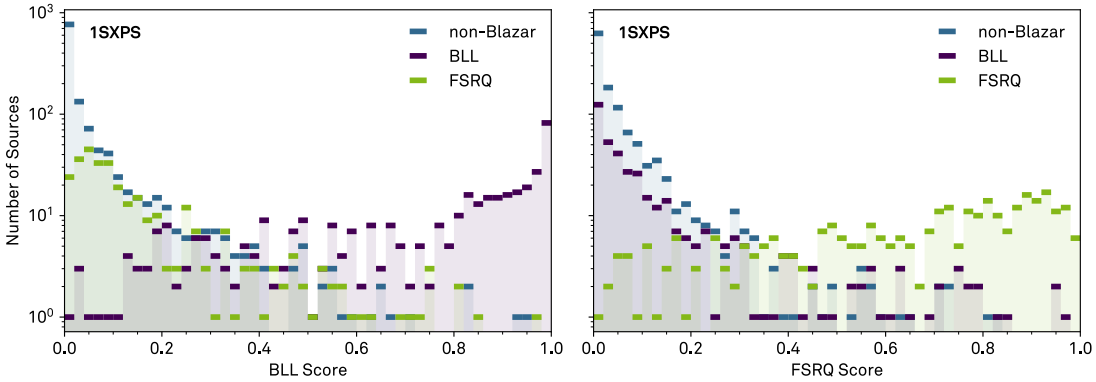

Figure 7.13: Distribution of the BLL (left) and FSRQ (right) scores of the training sample, obtained in a 10-fold cross validation. The scores have been derived from the Random Forest classification model, created with the 3FGL and the 1SXPS catalogs. While the BLL (FSRQ) class features a population at large BLL (FSRQ) scores, the respective remaining classes exhibit small scores.

The performance has been quantified with the ROC curve, which is shown in Figure 7.14. The AUC score, i. e. the area under this curve, amounts to 0.973 ± 0.014 for BLLs, and 0.965 ± 0.013 for FSRQs. The performance measures precision and recall are illustrated in Figure 7.15 dependent on the thresholds of the BLL and FSRQ score, respectively. The AUC, precision and recall scores, as well BLL and FSRQ score distributions, indicate a better classification of BLLs.

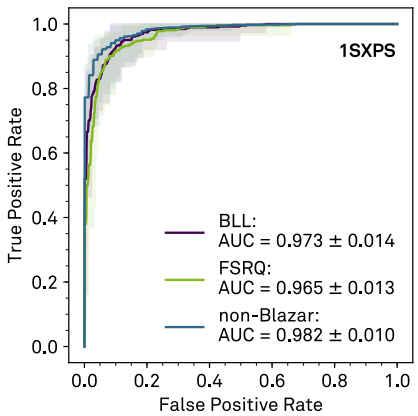

Figure 7.14: ROC curves of the Random Forest classification model for the 1SXPS catalog. The curve has been averaged over 10 curves from the 10-fold cross validation. The shaded area illustrates the minimum and maximum rates. The corresponding AUC and its uncertainty in terms of the standard deviation amounts to 0.973 ± 0.014 for BLLs and 0.965 ± 0.013 for FSRQs.

The classification model has also been applied to the BCU sample and to the counterparts within the 3FGL uncertainty region of the BCUs. The corresponding test sample contains 367 possible 1SXPS counterparts for 163 3FGL sources. Figure 7.16 confirms the capability of the model by large non-blazar scores for the unassociated BCUs, and small scores for the associated ones.

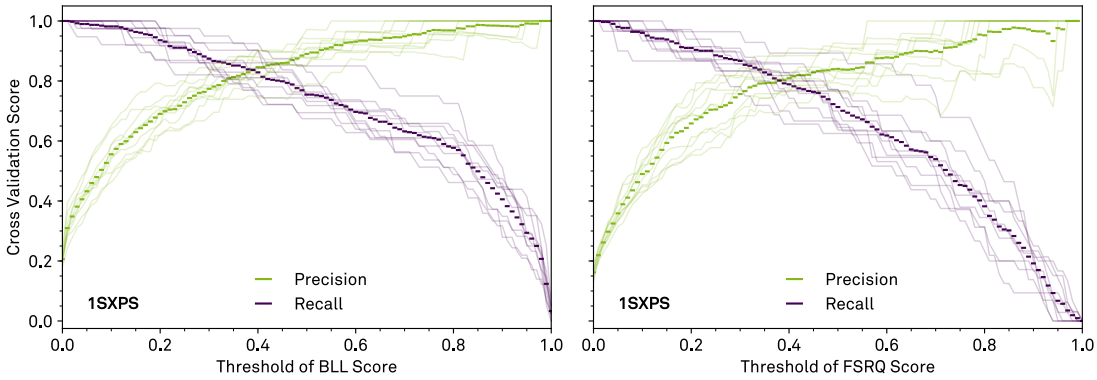

Figure 7.15: Precision and recall for different thresholds of the BLL (left) and FSRQ (right) scores of the Random Forest classification model for the 1SXPS catalog. The averaged precision and recall scores have been determined in a 10-fold cross validation. The individual scores of each cross validation subsample are shown as well.

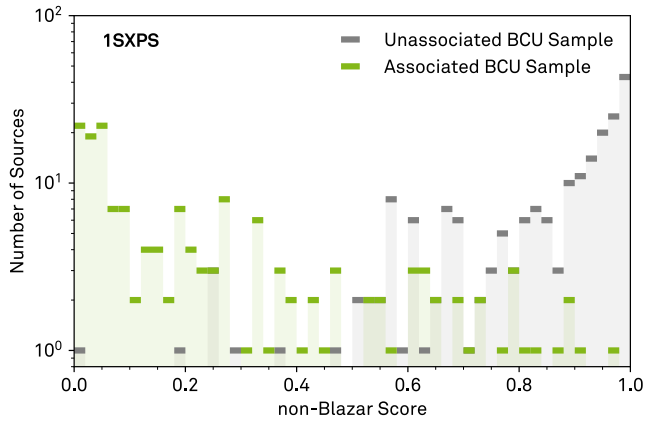

Figure 7.16: Distribution of the non-blazar scores for 3FGL sources associated with BCUs (associated BCU sample), and the counterpart candidates in the corresponding 3FGL uncertainty region (unassociated BCU sample). The shape of the distribution proves the capability of the method.

7.2.3 Application of the Classification Model to the Unassociated Sample and the Uncertain Blazar Sample

Applying the final classification model to the UGS and the BCU sample, the BLL and FSRQ distributions shown in Figure 7.17 have been achieved. The score distributions of the UGS sample have a rather sharp maximum at 0, but no maxima at higher scores. Still, multiple sources exhibit large BLL and FSRQ scores. In case of the BCU sample, the BLL score distributions features a maximum at scores of 1, while the FSRQ distribution has a maximum at 0. This indicates many BLLs and only few FSRQs in the BCU sample. Compared to the training sample, the distributions are quite different for the same reasons explained previously.

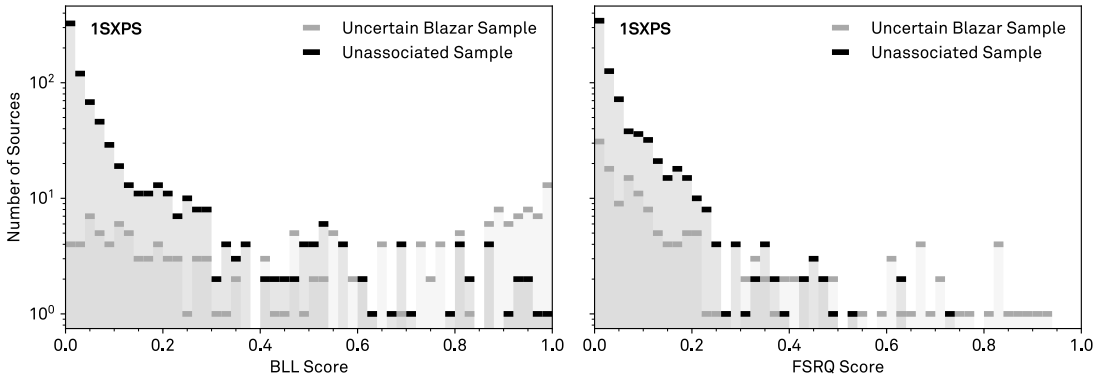

Figure 7.17: Distribution of the BLL (left) and FSRQ (right) scores of the unassociated (UGS) and the uncertain blazar (BCU) sample. The scores have been determined with the Random Forest classification model, created with the 3FGL and the 1SXPS catalog. Multiple sources from both the UGS and the BCU sample feature large BLL and FSRQ scores, respectively.

A threshold of 0.5, and a threshold corresponding to a precision of 0.9 have been applied to generate the medium- and the high-confidence samples, respectively. This results in precisions of 0.89 for BLLs, and 0.84 for FSRQs for the medium-confidence sample, and in a BLL threshold of 0.53 and a FSRQ threshold of 0.66 for the high-confidence 1SXPS sample.
The number of counterpart candidates per 3FGL source is visualized in Figure 7.18 for the UGS sample. Only for the BLLs of the medium-confidence sample some associations are ambiguous. The high-confidence sample contains 32 unambiguous BLL and 1 unambiguous FSRQ candidates. The medium-confidence sample comprises 41 BLL counterpart candidates for 39 3FGL sources, and 4 unambiguous FSRQ candidates. Table 7.3 lists the high-confidence BLL and FSRQ candidates and the names of their corresponding counterparts for the UGS sample, and gives additional information such as the Galactic latitude and the analysis flag.
Applying the thresholds to the BCU sample, the high-confidence 1SXPS sample contains 88 BLL and 19 FSRQ candidates, and the medium-confidence sample 90 BLL and 27 FSRQ candidates. The high-confidence BLL and FSRQ candidates of the BCU sample are listed in Table 7.4.

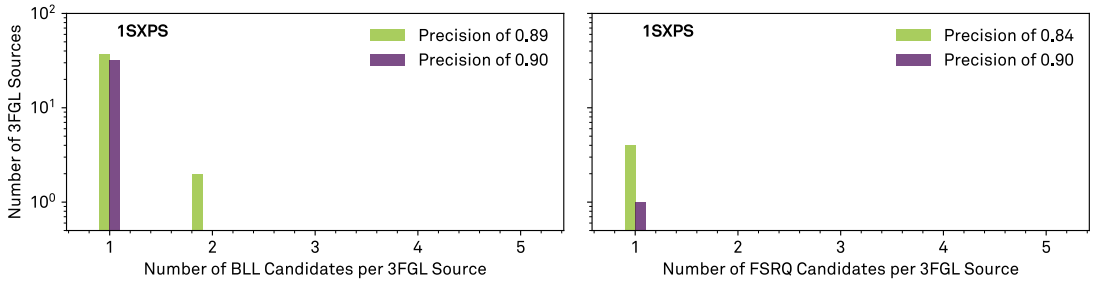

Figure 7.18: Number of 3FGL sources with specific numbers of BLL (left) and FSRQ (right) candidates per 3FGL source, depending on the BLL and FSRQ score, respectively. For a small BLL score threshold, the model associates sometimes 2 counterparts to individual 3FGL sources, making the association ambiguous. For the remaining applied thresholds, the association is definite.

Table 7.3: List of 32 high-confidence BLL and 1 FSRQ candidates and their corresponding 1SXPS counterparts. Out of these, 22 BLL and 1 FSRQ candidates exhibit neither 3FGL analysis flags, nor are located within the Galactic plane.

3FGL Name	$\lvert b \rvert$ / deg	3FGL Flags	1SXPS Name	BLL Score	FSRQ Score
BLL Candidates					
J0156.5–2423	75	0	J015624.4–242008	0.80	0.11
J0200.3–4108	70	0	J020020.9–410936	0.86	0.03
J0212.1+5320	8	0	J021210.6+532136	0.57	0.34
J0322.5–3721	57	0	J032231.8–372023	0.68	0.09
J0420.4–6013	42	0	J042010.9–601502	0.70	0.07
J0704.3–4828	18	0	J070421.6–482645	0.82	0.07
J0721.5–0221	6	256	J072113.8–022053	0.82	0.08
J0725.7–0550	5	0	J072547.7–054828	0.91	0.04
J0737.8–8245	25	0	J073706.2–824837	0.88	0.04
J0746.4–0225	11	0	J074627.1–022550	0.96	0.01
J0802.3–0941	11	0	J080216.1–094207	0.60	0.17
J0826.3–6400	15	0	J082628.0–640414	0.54	0.24
J0919.5–2200	19	0	J091926.2–220043	0.69	0.16
J1016.6–4244	12	0	J101620.6–424720	0.79	0.13
J1033.4–5035	7	0	J103332.1–503529	0.98	0.01
J1117.2–5338	7	4	J111714.9–533817	0.84	0.03
J1132.0–4736	13	1	J113209.5–473857	0.87	0.08
J1240.3–7149	9	0	J124021.3–714857	0.95	0.03
J1249.1–2808	35	0	J124919.2–280832	0.95	0.03
J1251.0–4943	13	0	J125058.9–494450	0.61	0.03
J1315.7–0732	55	0	J131553.0–073301	0.93	0.02
J1424.3–1753	40	0	J142412.5–175010	0.82	0.05
J1704.4–0528	21	0	J170433.7–052841	0.69	0.13
J1808.4–3703	8	0	J180827.6–365843	0.67	0.14
J1817.3–3033	7	8	J181721.0–303259	0.56	0.35
J2034.6–4202	36	0	J203450.9–420036	0.92	0.04
J2133.0–6433	41	0	J213311.3–643822	0.63	0.11
J2144.6–5640	46	0	J214429.3–563847	0.58	0.23
J2300.0+4053	17	0	J230012.4+405225	0.81	0.03
J2309.0+5428	5	0	J230848.5+542613	0.58	0.13
J2351.9–7601	41	0	J235115.3–760018	0.86	0.06
J2358.6–1809	75	0	J235837.1–180714	0.69	0.08
FSRQ Candidates					
J1544.6–1125	33	0	J154439.3–112804	0.21	0.74

Table 7.4: List of 88 high-confidence BLL and 19 FSRQ candidates and their corresponding 1SXPS counterparts. Out of these, 69 BLL and 12 FSRQ candidates exhibit neither 3FGL analysis flags, nor are located within the Galactic plane.

3FGL Name	$\|b\|$ / deg	3FGL Flags	1SXPS Name	BLL Score	FSRQ Score
BLL Candidates					
J0003.2−5246	63	0	J000319.6−524727	0.94	0.02
J0009.6−3211	80	1	J000935.7−321637	0.59	0.19
J0015.7+5552	7	0	J001540.0+555142	0.87	0.08
J0031.3+0724	55	0	J003119.9+072452	0.72	0.06
J0039.0−2218	84	0	J003908.1−222000	0.82	0.16
J0040.3+4049	22	0	J004013.6+405002	0.80	0.14
J0051.2−6241	54	0	J005116.5−624203	0.99	0.00
J0116.2−2744	84	0	J011555.6−274433	0.88	0.04
J0127.2+0325	58	0	J012713.9+032258	0.94	0.02
J0137.8+5813	4	0	J013750.4+581409	0.95	0.03
J0147.0−5204	63	0	J014648.5−520231	0.66	0.17
J0226.5−4442	64	0	J022638.6−444120	0.94	0.02
J0249.1+8438	22	0	J024948.2+843553	0.59	0.08
J0250.6+5630	3	0	J025047.4+562933	0.78	0.19
J0256.3+0335	47	0	J025627.9+033331	0.65	0.09
J0310.4−5015	55	0	J031034.3−501629	0.94	0.03
J0338.5+1303	33	0	J033829.0+130213	0.96	0.02
J0339.2−1738	50	0	J033913.6−173559	0.90	0.06
J0352.9+5655	2	0	J035309.5+565429	0.91	0.04
J0431.6+7403	17	0	J043145.1+740323	0.90	0.01
J0433.1+3228	10	0	J043307.5+322842	0.92	0.04
J0506.9−5435	37	0	J050657.9−543507	0.99	0.00
J0509.7−6418	35	0	J050957.5−641740	0.94	0.06
J0512.9+4038	1	0	J051252.2+404143	0.56	0.23
J0538.4−3909	30	0	J053810.4−390842	0.89	0.07
J0553.5−2036	22	0	J055333.1−203418	0.65	0.07
J0602.8−4016	26	0	J060251.0−401844	0.80	0.05
J0640.0−1252	8	4	J064007.1−125313	0.87	0.07
J0647.0−5134	22	0	J064709.8−513545	0.82	0.15
J0649.6−3138	14	0	J064933.7−313921	0.98	0.00
J0652.0−4808	20	0	J065200.5−480857	0.61	0.11
J0700.2+1304	8	0	J070014.2+130423	0.73	0.07
J0708.9+2239	14	0	J070858.0+224135	0.91	0.06
J0733.5+5153	27	0	J073326.8+515355	0.95	0.02
J0746.6−4756	11	0	J074642.2−475451	0.89	0.07
J0746.9+8511	28	0	J074715.2+851204	0.98	0.00
J0813.3+6509	33	0	J081240.6+650908	0.77	0.15
J0841.3−3554	4	0	J084121.6−355508	0.98	0.00
J0947.1−2542	21	0	J094709.7−254103	0.99	0.00
J0953.1−7657c	18	32	J095303.6−765802	0.92	0.02
J1031.0+7440	39	0	J103121.9+744155	0.80	0.17
J1040.4+0615	53	0	J104031.6+061722	0.53	0.35
J1141.6−1406	45	0	J114141.6−140755	0.86	0.12
J1154.0−3243	29	0	J115406.0−324240	0.73	0.09
J1223.3−3028	32	0	J122336.8−303248	0.74	0.21
J1256.1−5919	4	8	J125604.3−591942	0.54	0.08
J1256.3−1146	51	0	J125615.7−114632	0.94	0.00
J1259.8−3749	25	0	J125950.0−374857	0.68	0.08
J1307.6−4300	20	0	J130737.8−425938	1.00	0.00
J1314.7−4237	20	0	J131503.2−423649	0.77	0.10
J1315.4+1130	73	0	J131532.6+113333	0.76	0.18
J1340.6−0408	57	0	J134042.2−041006	0.88	0.03
J1342.7+0945	69	0	J134240.2+094752	0.64	0.04
J1346.9−2958	31	0	J134707.0−295844	0.89	0.07
J1406.0−2508	35	0	J140609.9−250806	0.93	0.02
J1418.9+7731	39	0	J141900.7+773230	0.98	0.01
J1511.8−0513	43	0	J151148.7−051347	1.00	0.00
J1512.2−2255	29	0	J151212.6−225506	0.89	0.06
J1537.8−8000	19	0	J153739.6−795802	0.55	0.18
J1543.5+0451	43	0	J154333.6+045219	0.63	0.10
J1549.7−0658	35	0	J154952.1−065908	0.98	0.02
J1626.4−7640	19	0	J162638.7−763853	0.77	0.05
J1640.9+1142	34	0	J164058.7+114406	0.72	0.17
J1736.0+2033	25	0	J173605.1+203305	1.00	0.00
J1816.9−4944	15	4	J181655.6−494344	0.55	0.19
J1820.3+3625	22	0	J182020.7+362345	0.86	0.08
J1823.6−3453	10	0	J182338.7−345413	1.00	0.00
J1842.3−5841	22	0	J184229.8−584159	0.99	0.01
J1844.3+1547	9	0	J184425.4+154646	0.97	0.00
J1933.4+0727	6	0	J193320.3+072621	0.91	0.03
J1942.7+1033	6	0	J194247.5+103327	0.98	0.01
J1954.9−5640	31	0	J195502.6−564027	0.80	0.05
J1959.8−4725	31	0	J195945.6−472518	0.94	0.03
J2002.7+6303	16	0	J200245.3+630228	0.56	0.11
J2014.5+0648	15	0	J201431.0+064851	0.94	0.02
J2031.0+1937	12	0	J203057.4+193615	0.98	0.01
J2036.6−3325	35	0	J203649.3−332829	0.89	0.06
J2040.2−7115	34	0	J204007.1−711457	0.95	0.04

Table 7.4: – continued from previous page

3FGL Name	$\|b\|$ / deg	3FGL Flags	1SXPS Name	BLL Score	FSRQ Score
J2056.7+4938	3	0	J205642.6+494009	0.88	0.03
J2109.1–6638	38	0	J210851.5–663722	0.98	0.01
J2133.8+6648	11	4	J213348.9+664704	0.95	0.03
J2232.9–2021	58	0	J223248.9–202221	0.91	0.05
J2246.7–5205	56	256	J224642.0–520641	0.97	0.01
J2250.3–4206	61	0	J225022.3–420608	0.61	0.01
J2250.7–2806	63	0	J225044.5–280639	0.56	0.38
J2312.9–6923	45	0	J231347.6–692329	0.91	0.05
J2322.9–4917	62	0	J232254.3–491628	0.93	0.03
J2346.7+0705	52	0	J234639.8+070508	0.96	0.03
FSRQ Candidates					
J0045.2–3704	80	0	J004511.7–370546	0.05	0.68
J0156.3+3913	22	0	J015631.0+391429	0.10	0.84
J0205.0+1510	44	0	J020450.3+151411	0.05	0.77
J0210.7–5101	62	0	J021046.3–510059	0.07	0.91
J0301.8–7157	42	0	J030138.1–715631	0.03	0.84
J0522.9–3628	33	0	J052257.9–362731	0.06	0.93
J0532.0–4827	33	0	J053158.5–482734	0.26	0.69
J0622.9+3326	9	0	J062252.2+332610	0.23	0.70
J0747.2–3311	4	0	J074719.8–331047	0.15	0.66
J0858.1–1951	17	0	J085805.4–195038	0.12	0.82
J0904.8–3516	8	0	J090442.3–351422	0.05	0.76
J0923.1+3853	45	0	J092314.4+384940	0.09	0.66
J1038.9–5311	5	0	J103840.6–531143	0.03	0.71
J1328.9–5607	6	0	J132901.0–560759	0.16	0.82
J1508.7–4956	7	16	J150838.9–495259	0.03	0.90
J1604.4–4442	6	0	J160430.9–444129	0.08	0.83
J1647.4+4950	40	0	J164734.6+495000	0.10	0.88
J1941.8+7218	22	0	J194126.4+722137	0.02	0.67
J2328.4–4034	68	0	J232819.2–403504	0.11	0.84

7.3 Blazar Candidates and Radio Counterparts

This section deals with radio counterpart candidates from the FIRST catalog (cf. subsection 5.1.3). The classification model is based on a training sample comprising 277 BLLS, 169 FSRQs, and 2 398 non-blazars. This model has been applied to 964 counterpart combinations, located in the uncertainty regions of 103 3FGL sources of the UGS sample, and to 59 sources of the BCU sample.

7.3.1 Feature Selection

The feature selection for the radio counterparts has been conducted analogously to the procedure in subsection 7.1.1. The resulting mean and the standard deviation of the AUC score is shown in Figure 7.19. At some point, the AUC score saturates for all three classes. This saturation value is different for the three classes. The value for the FSRQs is the highest, and the one for the BLLs the lowest. This result is opposite to the ones obtained in the infrared and X-ray regime. Again, the achieved AUC scores indicate a successful classification. Since no obvious maximum has been formed, a number of 20 features has been chosen. Out of these, 6 features originate from the 3FGL catalog, 6 from the FIRST catalog, and 8 from both catalogs. The respective features are listed in Figure 7.20.

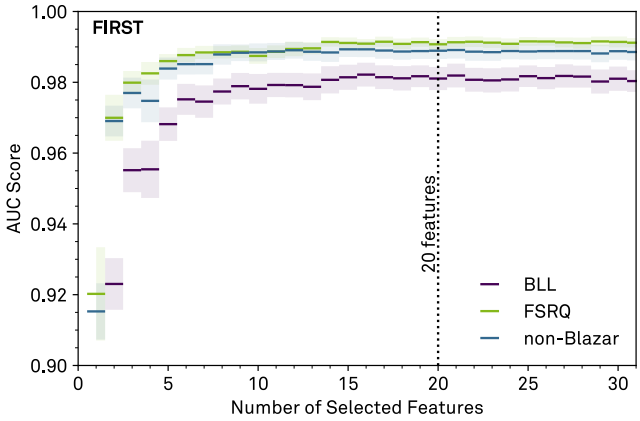

Figure 7.19: Performance of different numbers of selected features for each of the three classes, determined with a recursive backward elimination for the FIRST catalog in a 10-fold cross validation. The performance is evaluated with the *one-vs-rest* strategy, and is quantified by the mean of the AUC score, and is saturating above a specific number of features. A number of 20 features has been chosen for the further analysis.

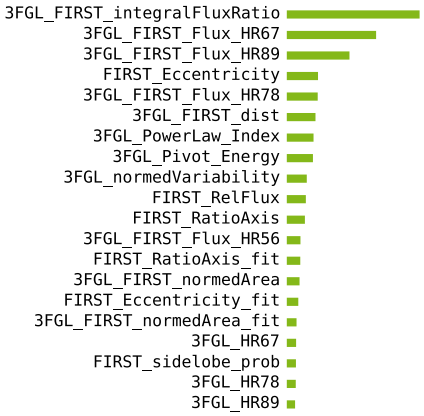

Figure 7.20: List of selected features and their relative feature importance. Out of these, 6 features are based on the FIRST catalog, 6 on the 3FGL catalog, and 8 on both catalogs.

7.3.2 Optimization and Performance of the Classification Model

The classification model has been optimized and its performance has been evaluated, following the same procedure as in subsection 7.1.2. The optimization of the classification model based on the FIRST catalog has resulted in the following hyperparameters:

Number of trees 200
Split selection criteria Information gain
Number of features considered at each node 7
Minimum number of samples for split 9

The final classification model has been generated with the stated optimized hyperparameters. The corresponding feature importance is visualized in Figure 7.20. Three features appear to be more important than the others: The ratio between the integral radio and gamma-ray fluxes, and the ratios between hardness ratios of the 3FGL catalog and the radio flux. The corresponding two-dimensional distributions are illustrated in Figure 7.21. BLLs exhibit both large ratios between radio and gamma-ray fluxes, as well as large gamma-ray hardness ratios compared to non-blazars. FSRQs feature even larger hardness and flux ratios. In all distributions, the diverse classes have formed distinctive populations. Consistent with the previous chapter, the following physical conclusions can be drawn: The larger the ratio between radio and gamma-ray flux of a blazar, the softer the spectrum in the gamma-ray regime.

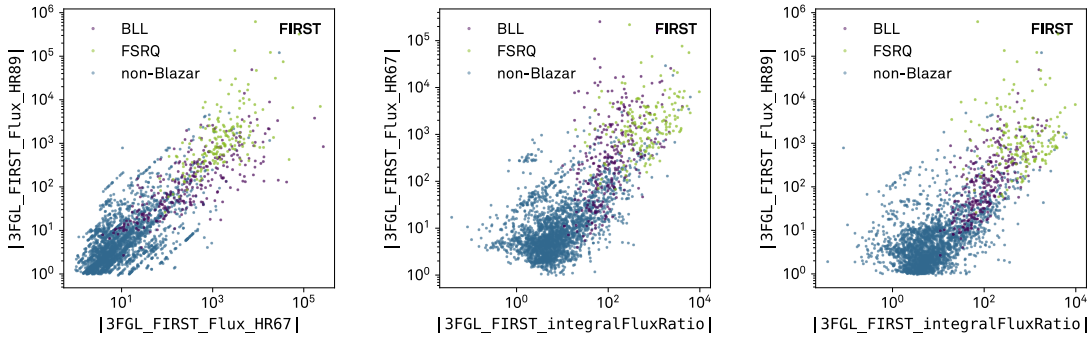

Figure 7.21: Two-dimensional distributions between the three features with the highest feature importance for BLLs, FSRQs and non-blazars. BLLs, FSRQs and non-blazars populate distinctive regions in every distribution.

The BLL and FSRQ score distributions for the three classes in the training sample are shown in Figure 7.22. Both score distributions do not feature a sharp maximum at scores of 1. However, the BLLs feature larger BLL scores than the remaining classes, and the FSRQs larger FSRQ scores than the others.

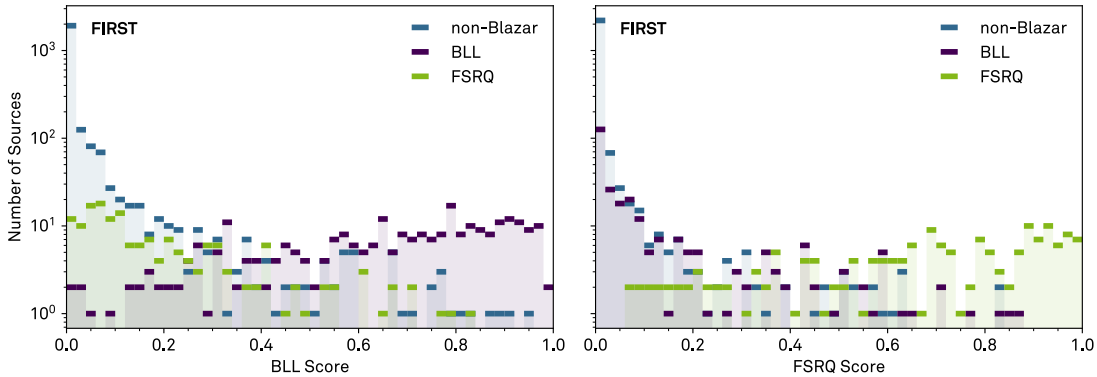

Figure 7.22: Distribution of the BLL (left) and FSRQ (right) scores of the training sample, obtained in a 10-fold cross validation. The scores have been derived from the Random Forest classification model, created with the 3FGL and the FIRST catalogs. While the BLL (FSRQ) class features a population at large BLL (FSRQ) scores, the respective remaining classes exhibit small scores.

The ROC curve, shown in Figure 7.23, quantifies the performance of the classification model. A corresponding AUC score of 0.982 ± 0.009 for BLLs, and 0.991 ± 0.005 for FSRQs has been achieved. The precision and recall has been utilized as performance measures, and their distributions as a function of the BLL and FSRQ score thresholds are presented in Figure 7.24. The AUC, precision and recall scores, as well as the BLL and FSRQ score distributions, indicate a better classification of FSRQs.

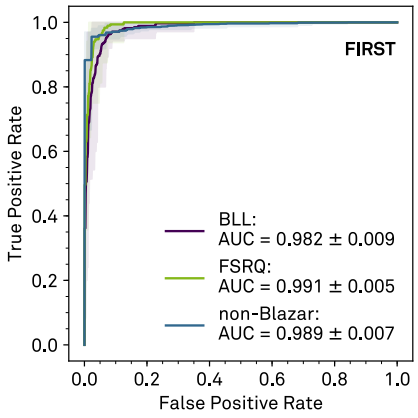

Figure 7.23: ROC curves of the Random Forest classification model for the FIRST catalog. The curve has been averaged over 10 curves from the 10-fold cross validation. The shaded area illustrates the minimum and maximum rates. The corresponding AUC and its uncertainty in terms of the standard deviation amounts to 0.982 ± 0.009 for BLLs and 0.991 ± 0.005 for FSRQs.

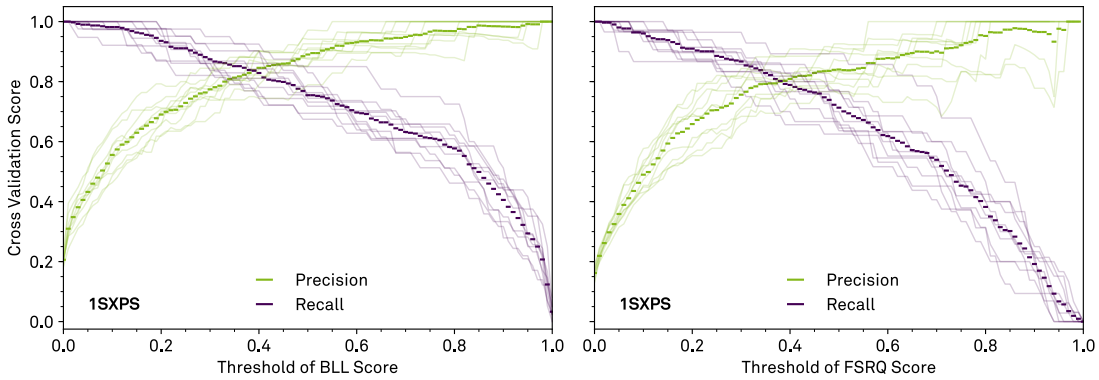

Figure 7.24: Precision and recall for different thresholds of the BLL (left) and FSRQ (right) scores of the Random Forest classification model for the 1SXPS catalog. The averaged precision and recall scores have been determined in a 10-fold cross validation. The individual scores of each cross validation subsample are shown as well.

Moreover, the classification model has been applied to the (associated) BCU sample and to the (unassociated) sample of the counterparts within the 3FGL uncertainty region of the BCUs. The corresponding test sample contains 404 possible FIRST counterparts for 59 3FGL sources. Figure 7.25 confirms the capability of the model by large non-blazar scores for the unassociated BCUs, and rather small scores for the associated ones.

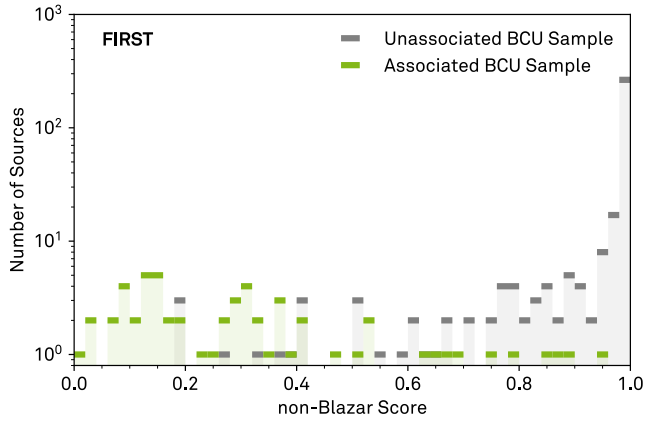

Figure 7.25: Distribution of the non-blazar scores for 3FGL sources associated with BCUs (associated BCU sample), and the counterpart candidates in the corresponding 3FGL uncertainty region (unassociated BCU sample). The shape of the distribution proves the capability of the method.

7.3.3 Application of the Classification Model to the Unassociated Sample and the Uncertain Blazar Sample

The application of the final classification model to the UGS and the BCU sample results in the BLL and FSRQ distributions shown in Figure 7.26. The score distributions of the UGS sample exhibit a rather sharp maximum at 0, but no maxima at higher scores have been formed. However, quite some sources with large BLL scores exist, and also a few with moderate FSRQ scores. In case of the BCU sample, the score distributions are rather flat. The shape of the distributions are differing from the one of the training sample, as previously already explained.

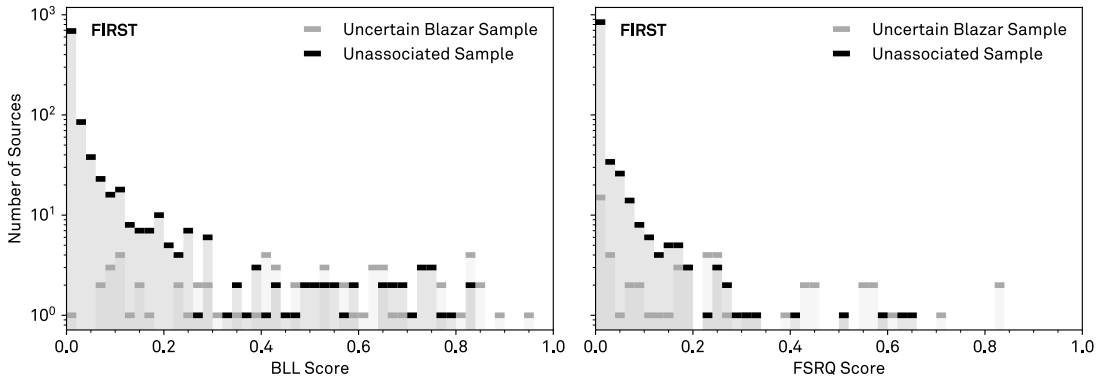

Figure 7.26: Distribution of the BLL (left) and FSRQ (right) scores of the unassociated (UGS) and the uncertain blazar (BCU) sample. The scores have been determined with the Random Forest classification model, created with the 3FGL and the FIRST catalog. Several sources from both the UGS and the BCU sample feature moderate to large BLL and FSRQ scores, respectively.

A threshold of 0.5, and a threshold corresponding to a precision of 0.9 have been applied to generate the medium-, and the high-confidence samples, respectively. This leads to precisions of 0.81 for BLLs, and 0.78 for FSRQs for the medium-confidence sample, and to a BLL threshold of 0.68 and a FSRQ threshold of 0.63 for the high-confidence FIRST sample.

Figure 7.27 illustrates the number of counterpart candidates per 3FGL source for the UGS sample. Some BLLs of the medium-confidence sample exhibit ambiguous associations, while the remaining associations are definite. The high-confidence sample contains 13 BLL and 1 FSRQ candidates. The medium-confidence sample comprises 26 BLL counterpart candidates for 24 3FGL sources, and 4 FSRQ candidates. Table 7.6 lists the high-confidence BLL and FSRQ candidates and the names of their corresponding counterparts for the UGS sample, and gives additional information such as the Galactic latitude and the analysis flag.
Applying the thresholds to the BCU sample, the high-confidence 1SXPS sample contains 12 BLL and 3 FSRQ candidates, and the medium-confidence sample 26 BLL and 10 FSRQ candidates. The high-confidence BLL and FSRQ candidates of the BCU sample are listed in Table 7.5.

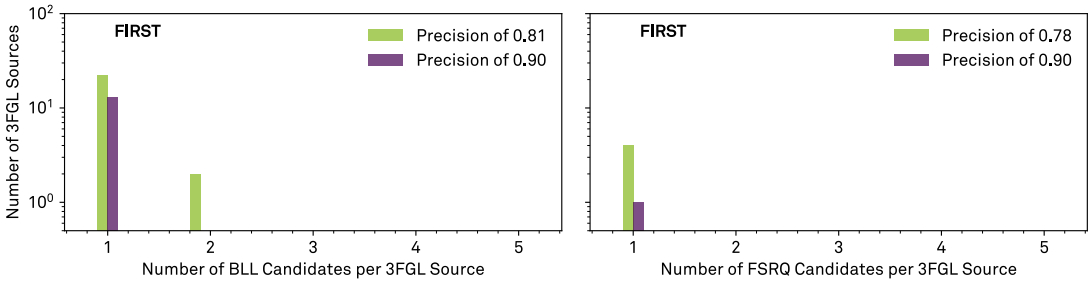

Figure 7.27: Number of 3FGL sources with specific numbers of BLL (left) and FSRQ (right) candidates per 3FGL source, depending on the BLL and FSRQ score, respectively. Despite for BLLs and a precision of 0.81, the association is definite.

Table 7.5: List of 12 high-confidence BLL and 3 FSRQ candidates and their corresponding FIRST counterparts. Out of these, all BLL and FSRQ candidates exhibit neither 3FGL analysis flags, nor are located within the Galactic plane.

3FGL Name	$\|b\|$ / deg	3FGL Flags	FIRST Name	BLL Score	FSRQ Score
BLL Candidates					
J0043.5−0444	68	0	J004334.1−044300	0.81	0.00
J0733.5+5153	27	0	J073326.7+515355	0.89	0.00
J0917.3−0344	30	0	J091714.5−034314	0.76	0.01
J1159.6−0723	53	0	J115931.8−072359	0.82	0.00
J1330.9+5201	64	0	J133042.6+520215	0.77	0.14
J1340.6−0408	57	0	J134042.0−041006	0.83	0.02
J1342.7+0945	69	0	J134240.0+094752	0.86	0.01
J1549.5+1709	48	0	J154929.2+170828	0.69	0.18
J1636.7+2624	40	0	J163651.4+262656	0.84	0.01
J1740.4+5347	32	0	J174036.4+534624	0.95	0.01
J2112.7+0819	26	0	J211243.0+081835	0.82	0.00
J2346.7+0705	52	0	J234639.9+070506	0.86	0.08
FSRQ Candidates					
J0836.3+2143	32	0	J083616.2+213903	0.09	0.83
J0923.1+3853	45	0	J092314.4+384939	0.08	0.84
J2049.7+1002	21	0	J204945.8+100314	0.16	0.71

Table 7.6: List of 13 high-confidence BLL and 1 FSRQ candidates and their corresponding FIRST counterparts. Out of these, 12 BLL and 1 FSRQ candidates exhibit neither 3FGL analysis flags, nor are located within the Galactic plane.

3FGL Name	$\|b\|$ / deg	3FGL Flags	FIRST Name	BLL Score	FSRQ Score
BLL Candidates					
J0008.3+1456	47	0	J000825.3+145635	0.74	0.07
J0749.5+1320	19	0	J074935.9+132156	0.72	0.07
J0952.8+0711	43	0	J095249.5+071330	0.82	0.00
J1049.7+1548	60	0	J104939.3+154837	0.76	0.08
J1052.0+0816	56	0	J105209.4+081605	0.69	0.01
J1151.5+0957	68	0	J115117.3+095826	0.69	0.06
J1200.4+0202	62	0	J120012.3+020208	0.74	0.01
J1232.3+1701	79	0	J123224.7+170428	0.83	0.01
J1234.7–0437	58	512	J123444.2–043622	0.72	0.02
J1411.4–0724	50	0	J141133.3–072253	0.72	0.04
J1517.0+2637	58	0	J151702.6+263859	0.74	0.01
J1541.6+1414	48	0	J154150.0+141437	0.80	0.01
J2043.6+0001	25	0	J204342.1+000119	0.76	0.07
FSRQ Candidates					
J1051.0+5332	56	0	J105106.7+534754	0.15	0.65

7.4 High-Confidence Blazar Candidates and their Multi-Wavelength Counterparts

For each counterpart catalog, samples of medium-confidence counterpart candidates have been provided in the previous sections. The applied BLL and FSRQ score thresholds of 0.5 lead to precisions between 0.81 and 0.94 for the prediction of BLLs, and to precisions between 0.78 and 0.85 for the prediction of FSRQs. To increase the precision, and to achieve more high-confidence candidates as in the sections before, the medium-confidence samples have been joined. In this joining procedure, a positional coincidence between the counterparts is demanded, defined by a specific maximum distance between the counterparts. These maximum distances have been set to the sum of the association radii of the corresponding catalog (cf. section 5.3). The maximum distance between 1SXPS and ALLWISE counterparts amounts to 11.7 arcseconds, between FIRST and ALLWISE counterparts to 6.0 arcseconds, and between FIRST and 1SXPS counterparts to 10.7 arcseconds.
The performance has been calculated based on the training sample with the same applied score thresholds. A bootstrap method with 1000 samples has been utilized to estimate the uncertainty of the obtained precision. The results for different conjunctions of the counterpart catalogs are listed in Table 7.7.
The resulting precisions of more than 0.9 point out the successful creation of a high-confidence sample from medium-confidence samples by demanding a positional coincidence between the counterparts.

Table 7.7: Precisions for different combinations of counterpart catalogs. The medium-confidence candidates of the individual counterpart catalogs have been joined, and high-confidence candidates have been obtained by requiring a specific positional distance between the candidates of different catalogs. The uncertainty has been determined with 1000 bootstrap samples.

Counterpart Catalogs	Precision (BLL)	Precision (FSRQ)
1SXPS / ALLWISE	0.97 ± 0.01	0.90 ± 0.02
FIRST / ALLWISE	0.97 ± 0.01	0.90 ± 0.03
FIRST / 1SXPS	0.93 ± 0.02	0.93 ± 0.03
FIRST / 1SXPS / ALLWISE	0.98 ± 0.02	0.94 ± 0.03

For different conjunctions of the counterpart catalogs, the resulting high-confidence BLL and FSRQ candidates and their corresponding counterparts are listed in Table 7.8, Table 7.10, Table 7.12, and Table 7.14 for the UGS sample. The high-confidence BLL and FSRQ candidates for the BCU sample are listed in Table 7.9, Table 7.11, Table 7.13, and Table 7.15.

Table 7.8: List of 8 high-confidence BLL and 1 FSRQ candidates from the UGS sample and their corresponding 1SXPS and ALLWISE counterpart candidates.

3FGL Name	$\lvert b\rvert$ / deg	3FGL Flags	1SXPS Name	WISE Name	Dist. / ″	1SXPS Score BLL	1SXPS Score FSRQ	WISE Score BLL	WISE Score FSRQ
BLL Candidates									
J0746.4–0225	11	0	J074627.1–022550	J074627.02–022549.3	1.3	0.96	0.01	0.86	0.09
J0802.3–0941	11	0	J080216.1–094207	J080215.90–094210.9	4.7	0.60	0.17	0.63	0.09
J1132.0–4736	13	1	J113209.5–473857	J113209.26–473853.3	5.0	0.87	0.08	0.74	0.03
J1240.3–7149	9	0	J124021.3–714857	J124021.21–714857.7	1.1	0.95	0.03	0.91	0.01
J1249.1–2808	35	0	J124919.2–280832	J124919.31–280834.4	1.8	0.95	0.03	0.62	0.00
J1315.7–0732	55	0	J131553.0–073301	J131552.97–073302.0	0.3	0.93	0.02	1.00	0.00
J2258.2–3645	64	0	J225814.7–364428	J225815.00–364434.3	8.7	0.52	0.03	0.79	0.08
J2358.6–1809	75	0	J235837.1–180714	J235836.83–180717.4	7.4	0.69	0.08	0.83	0.05
FSRQ Candidates									
J1038.0–2425	29	0	J103754.9–242542	J103754.91–242544.5	1.4	0.21	0.52	0.03	0.56

Table 7.9: List of 49 high-confidence BLL and 19 FSRQ candidates from the BCU sample and their corresponding 1SXPS and ALLWISE counterpart candidates.

3FGL Name	$\lvert b\rvert$ / deg	3FGL Flags	1SXPS Name	WISE Name	Dist. / ″	1SXPS Score BLL	1SXPS Score FSRQ	WISE Score BLL	WISE Score FSRQ
BLL Candidates									
J0015.7+5552	7	0	J001540.0+555142	J001540.13+555144.7	2.4	0.87	0.08	0.89	0.05
J0031.3+0724	55	0	J003119.9+072452	J003119.71+072453.4	3.2	0.72	0.06	0.84	0.04
J0051.2–6241	54	0	J005116.5–624203	J005116.64–624204.3	1.1	0.99	0.00	0.98	0.01
J0137.8+5813	4	0	J013750.4+581409	J013750.47+581411.3	1.8	0.95	0.03	0.99	0.00
J0147.0–5204	63	0	J014648.5–520231	J014648.58–520233.5	2.2	0.66	0.17	0.51	0.03
J0333.4+7853	18	0	J033344.8+785026	J033344.58+785028.5	1.9	0.51	0.07	0.94	0.00
J0338.5+1303	33	0	J033829.0+130213	J033829.27+130215.5	3.8	0.96	0.02	0.90	0.03
J0339.2–1738	50	0	J033913.6–173559	J033913.70–173600.8	0.8	0.90	0.06	0.76	0.03
J0352.9+5655	2	0	J035309.5+565429	J035309.54+565430.7	1.5	0.91	0.04	0.93	0.03
J0431.6+7403	17	0	J043145.1+740323	J043145.06+740326.6	2.9	0.90	0.01	0.95	0.01
J0506.9–5435	37	0	J050657.9–543507	J050657.81–543503.7	3.9	0.99	0.00	0.56	0.10
J0512.9+4038	1	0	J051252.2+404143	J051252.53+404143.7	3.6	0.56	0.23	0.57	0.18
J0553.5–2036	22	0	J055333.1–203418	J055333.12–203418.9	1.2	0.65	0.07	0.94	0.04
J0640.0–1252	8	4	J064007.1–125313	J064007.18–125315.0	1.1	0.87	0.07	0.90	0.01
J0647.0–5134	22	0	J064709.8–513545	J064710.04–513547.7	3.4	0.82	0.15	0.82	0.03
J0700.2+1304	8	0	J070014.2+130423	J070014.31+130424.4	3.5	0.73	0.07	0.98	0.02
J0708.9+2239	14	0	J070858.0+224135	J070858.28+224135.4	3.3	0.91	0.06	0.94	0.04
J0733.5+5153	27	0	J073326.8+515355	J073326.79+515355.9	0.0	0.95	0.02	0.61	0.03
J0746.6–4756	11	0	J074642.2–475451	J074642.31–475455.0	3.2	0.89	0.07	0.90	0.06

Table 7.9: – continued from previous page

3FGL Name	$\|b\|$ / deg	3FGL Flags	1SXPS Name	WISE Name	Dist. / ″	1SXPS Score BLL	1SXPS Score FSRQ	WISE Score BLL	WISE Score FSRQ
J0841.3–3554	4	0	J084121.6–355508	J084121.63–355505.9	3.2	0.98	0.00	0.95	0.04
J0947.1–2542	21	0	J094709.7–254103	J094709.52–254059.9	3.2	0.99	0.00	0.99	0.00
J0953.1–7657c	18	32	J095303.6–765802	J095304.35–765802.0	2.5	0.92	0.02	0.52	0.02
J1040.4+0615	53	0	J104031.6+061722	J104031.62+061721.7	0.8	0.53	0.35	0.83	0.14
J1154.0–3243	29	0	J115406.0–324240	J115406.16–324243.0	4.4	0.73	0.09	0.82	0.10
J1256.1–5919	4	8	J125604.3–591942	J125604.90–591943.9	4.1	0.54	0.08	0.80	0.14
J1256.3–1146	51	0	J125615.7–114632	J125615.95–114637.3	5.0	0.94	0.00	0.84	0.04
J1259.8–3749	25	0	J125950.0–374857	J125949.80–374858.1	1.1	0.68	0.08	0.89	0.04
J1307.6–4300	20	0	J130737.8–425938	J130737.98–425938.9	1.1	1.00	0.00	0.99	0.01
J1340.6–0408	57	0	J134042.2–041006	J134042.01–041006.8	3.6	0.88	0.03	0.93	0.04
J1342.7+0945	69	0	J134240.2+094752	J134240.02+094752.4	3.6	0.64	0.04	0.95	0.01
J1346.9–2958	31	0	J134707.0–295844	J134706.88–295842.4	1.4	0.89	0.07	0.99	0.01
J1406.0–2508	35	0	J140609.9–250806	J140609.60–250809.2	7.0	0.93	0.02	0.99	0.01
J1418.9+7731	39	0	J141900.7+773230	J141900.31+773229.3	2.1	0.98	0.01	0.81	0.06
J1511.8–0513	43	0	J151148.7–051347	J151148.56–051346.9	3.7	1.00	0.00	0.94	0.01
J1543.5+0451	43	0	J154333.6+045219	J154333.92+045219.3	3.6	0.63	0.10	0.55	0.03
J1626.4–7640	19	0	J162638.7–763853	J162638.15–763855.5	3.3	0.77	0.05	0.88	0.11
J1640.9+1142	34	0	J164058.7+114406	J164058.89+114404.2	2.5	0.72	0.17	0.57	0.07
J1736.0+2033	25	0	J173605.1+203305	J173605.25+203301.1	4.0	1.00	0.00	0.94	0.02
J1844.3+1547	9	0	J184425.4+154646	J184425.36+154645.8	0.4	0.97	0.00	0.99	0.00
J1942.7+1033	6	0	J194247.5+103327	J194247.48+103327.0	0.7	0.98	0.01	0.93	0.04
J1954.9–5640	31	0	J195502.6–564027	J195502.86–564028.8	2.3	0.80	0.05	0.98	0.00
J1959.8–4725	31	0	J195945.6–472518	J195945.66–472519.3	1.1	0.94	0.03	0.99	0.00
J2031.0+1937	12	0	J203057.4+193615	J203057.12+193612.9	4.4	0.98	0.01	0.58	0.04
J2040.2–7115	34	0	J204007.1–711457	J204008.27–711459.9	5.5	0.95	0.04	0.87	0.00
J2109.1–6638	38	0	J210851.5–663722	J210851.80–663722.7	1.5	0.98	0.01	0.93	0.04
J2133.8+6648	11	4	J213348.9+664704	J213349.20+664704.3	1.5	0.95	0.03	0.96	0.03
J2250.3–4206	61	0	J225022.3–420608	J225022.22–420613.4	4.7	0.61	0.01	0.99	0.01
J2250.7–2806	63	0	J225044.5–280639	J225044.48–280639.1	3.2	0.56	0.38	0.60	0.36
J2346.7+0705	52	0	J234639.8+070508	J234639.93+070506.8	1.8	0.96	0.03	0.99	0.01
FSRQ Candidates									
J0045.2–3704	80	0	J004511.7–370546	J004512.06–370548.5	4.7	0.05	0.68	0.13	0.84
J0059.1–5701	60	4	J005846.3–565910	J005846.59–565911.4	2.4	0.22	0.61	0.24	0.75
J0156.3+3913	22	0	J015631.0+391429	J015631.41+391430.8	4.3	0.10	0.84	0.07	0.86
J0205.0+1510	44	0	J020450.3+151411	J020450.42+151410.9	2.0	0.05	0.77	0.08	0.86
J0210.7–5101	62	0	J021046.3–510059	J021046.20–510101.8	2.7	0.07	0.91	0.13	0.87
J0301.8–7157	42	0	J030138.1–715631	J030138.47–715634.5	3.5	0.03	0.84	0.09	0.75
J0522.9–3628	33	0	J052257.9–362731	J052257.98–362730.8	0.7	0.06	0.93	0.35	0.61
J0647.6–6058	24	0	J064740.8–605804	J064740.86–605805.1	1.1	0.19	0.65	0.25	0.72
J0747.2–3311	4	0	J074719.8–331047	J074719.67–331047.0	3.0	0.15	0.66	0.27	0.73
J0858.1–1951	17	0	J085805.4–195038	J085805.36–195036.8	3.7	0.12	0.82	0.02	0.97
J0904.8–3516	8	0	J090442.3–351422	J090442.38–351424.3	1.4	0.05	0.76	0.06	0.81
J1038.9–5311	5	0	J103840.6–531143	J103840.66–531142.9	0.0	0.03	0.71	0.01	0.82
J1159.3–2226	39	0	J115911.2–222837	J115911.26–222836.8	0.4	0.29	0.60	0.15	0.80
J1322.3+0839	70	0	J132210.2+084232	J132210.17+084232.9	3.6	0.34	0.52	0.33	0.61
J1328.9–5607	6	0	J132901.0–560759	J132901.15–560802.5	3.2	0.16	0.82	0.09	0.91
J1508.7–4956	7	16	J150838.9–495259	J150838.83–495302.1	2.2	0.03	0.90	0.08	0.82
J1604.4–4442	6	0	J160430.9–444129	J160431.00–444132.1	2.9	0.08	0.83	0.19	0.81
J1723.7–7713	22	0	J172349.8–771346	J172350.80–771350.5	4.8	0.12	0.62	0.07	0.93
J2328.4–4034	68	0	J232819.2–403504	J232819.26–403509.8	5.0	0.11	0.84	0.20	0.79

Table 7.10: List of 3 high-confidence BLL candidates from the UGS sample and their corresponding FIRST and ALLWISE counterpart candidates.

3FGL Name	$\|b\|$ / deg	3FGL Flags	FIRST Name	WISE Name	Dist. / ″	FIRST Score BLL	FIRST Score FSRQ	WISE Score BLL	WISE Score FSRQ
BLL Candidates									
J0216.0+0300	54	0	J021600.4+030011	J021600.45+030011.9	0.0	0.52	0.03	0.81	0.13
J1315.7–0732	55	0	J131552.9–073301	J131552.97–073302.0	0.1	0.55	0.01	1.00	0.00
J1548.4+1455	47	0	J154824.3+145702	J154824.38+145702.8	0.0	0.52	0.02	0.74	0.07

Table 7.11: List of 15 high-confidence BLL and 4 FSRQ candidates from the BCU sample and their corresponding FIRST and ALLWISE counterpart candidates.

3FGL Name	$\|b\|$ / deg	3FGL Flags	FIRST Name	WISE Name	Dist. / ″	FIRST Score BLL	FIRST Score FSRQ	WISE Score BLL	WISE Score FSRQ
BLL Candidates									
J0132.5–0802	69	0	J013241.1–080404	J013241.13–080404.9	0.0	0.63	0.24	0.91	0.01
J0211.2–0649	62	0	J021116.9–064419	J021116.95–064419.9	0.5	0.57	0.08	0.85	0.05
J0255.8+0532	46	0	J025549.5+053355	J025549.51+053355.0	0.1	0.66	0.26	0.73	0.23
J0728.0+4828	26	0	J072759.8+482720	J072759.84+482720.3	0.0	0.53	0.22	0.67	0.24
J0733.5+5153	27	0	J073326.7+515355	J073326.79+515355.9	0.4	0.89	0.00	0.61	0.03
J0917.3–0344	30	0	J091714.5–034314	J091714.61–034314.2	0.1	0.76	0.01	0.96	0.00
J1159.6–0723	53	0	J115931.8–072359	J115931.87–072359.4	0.1	0.82	0.00	0.92	0.04
J1200.8+1228	71	0	J120040.0+123103	J120040.03+123103.2	0.0	0.66	0.04	0.67	0.12
J1340.6–0408	57	0	J134042.0–041006	J134042.01–041006.8	0.0	0.83	0.02	0.93	0.04
J1511.8–0513	43	0	J151148.5–051346	J151148.56–051346.9	0.0	0.59	0.00	0.94	0.01
J1543.5+0451	43	0	J154333.9+045219	J154333.92+045219.3	0.4	0.66	0.23	0.55	0.03
J1636.7+2624	40	0	J163651.4+262656	J163651.46+262656.7	0.0	0.84	0.01	0.91	0.05
J1740.4+5347	32	0	J174036.4+534624	J174036.52+534623.8	0.4	0.95	0.01	0.90	0.07
J2112.7+0819	26	0	J211243.0+081835	J211243.00+081835.2	0.1	0.82	0.00	0.93	0.03
J2346.7+0705	52	0	J234639.9+070506	J234639.93+070506.8	0.0	0.86	0.08	0.99	0.01
FSRQ Candidates									
J0216.6–1019	64	0	J021638.8–101703	J021638.87–101702.9	0.0	0.28	0.57	0.18	0.81
J0836.3+2143	32	0	J083616.2+213903	J083616.20+213903.7	0.0	0.09	0.83	0.02	0.61
J1014.2+4115	55	0	J101417.8+411217	J101417.87+411217.7	2.7	0.36	0.50	0.09	0.52
J1239.4+0727	70	0	J123924.5+073017	J123924.58+073017.2	0.1	0.43	0.55	0.37	0.62

Table 7.12: List of 1 high-confidence BLL candidate from the UGS sample and its corresponding FIRST and 1SXPS counterpart candidates.

3FGL Name	$\|b\|$ / deg	3FGL Flags	FIRST Name	1SXPS Name	Dist. / ″	FIRST Score BLL	FIRST Score FSRQ	1SXPS Score BLL	1SXPS Score FSRQ
BLL Candidates									
J1315.7–0732	55	0	J131552.9–073301	J131553.0–073301	0.1	0.55	0.01	0.93	0.02

Table 7.13: List of 8 high-confidence BLL and 2 FSRQ candidates from the BCU sample and their corresponding FIRST and 1SXPS counterpart candidates.

3FGL Name	$\|b\|$ / deg	3FGL Flags	FIRST Name	1SXPS Name	Dist. / ″	FIRST Score BLL	FIRST Score FSRQ	1SXPS Score BLL	1SXPS Score FSRQ
BLL Candidates									
J0127.2+0325	58	0	J012713.9+032300	J012713.9+032258	2.0	0.62	0.01	0.94	0.02
J0733.5+5153	27	0	J073326.7+515355	J073326.8+515355	0.4	0.89	0.00	0.95	0.02
J1315.4+1130	73	0	J131532.6+113331	J131532.6+113333	1.4	0.57	0.15	0.76	0.18
J1340.6–0408	57	0	J134042.0–041006	J134042.2–041006	3.6	0.83	0.02	0.88	0.03
J1342.7+0945	69	0	J134240.0+094752	J134240.2+094752	3.6	0.86	0.01	0.64	0.04
J1511.8–0513	43	0	J151148.5–051346	J151148.7–051347	3.7	0.59	0.00	1.00	0.00
J1543.5+0451	43	0	J154333.9+045219	J154333.6+045219	3.7	0.66	0.23	0.63	0.10
J2346.7+0705	52	0	J234639.9+070506	J234639.8+070508	1.7	0.86	0.08	0.96	0.03
FSRQ Candidates									
J0923.1+3853	45	0	J092314.4+384939	J092314.4+384940	0.7	0.08	0.84	0.09	0.66
J1647.4+4950	40	0	J164734.9+495000	J164734.6+495000	0.4	0.30	0.55	0.10	0.88

Table 7.14: List of 1 high-confidence BLL candidate from the UGS sample and its corresponding FIRST, 1SXPS and ALLWISE counterpart candidates.

3FGL Name	FIRST Name	1SXPS Name	WISE Name	FIRST Score BLL	FIRST Score FSRQ	1SXPS Score BLL	1SXPS Score FSRQ	WISE Score BLL	WISE Score FSRQ
J1315.7–0732	J131552.9–073301	J131553.0–073301	J131552.97–073302.0	0.55	0.01	0.93	0.02	1.00	0.00

Table 7.15: List of 5 high-confidence BLL candidates from the BCU sample and their corresponding FIRST, 1SXPS and ALLWISE counterpart candidates.

3FGL Name	FIRST Name	1SXPS Name	WISE Name	FIRST Score BLL	FIRST Score FSRQ	1SXPS Score BLL	1SXPS Score FSRQ	WISE Score BLL	WISE Score FSRQ
J0733.5+5153	J073326.7+515355	J073326.8+515355	J073326.79+515355.9	0.89	0.00	0.95	0.02	0.61	0.03
J1340.6–0408	J134042.0–041006	J134042.2–041006	J134042.01–041006.8	0.83	0.02	0.88	0.03	0.93	0.04
J1511.8–0513	J151148.5–051346	J151148.7–051347	J151148.56–051346.9	0.59	0.00	1.00	0.00	0.94	0.01
J1543.5+0451	J154333.9+045219	J154333.6+045219	J154333.92+045219.3	0.66	0.23	0.63	0.10	0.55	0.03
J2346.7+0705	J234639.9+070506	J234639.8+070508	J234639.93+070506.8	0.86	0.08	0.96	0.03	0.99	0.01

The conjunctions of the counterpart catalogs have lead to a reduction of the number of partly ambiguous candidates from the medium-confidence samples to a high-confidence sample of unambiguous candidates. Compared to the number of AGN candidates in section 6.4, the number of blazar candidates is slightly smaller. This is quite reasonable, since the classification task has been extended from a 2-class to a 3-class problem, implying i. a. smaller class representations.

The majority of the candidates obtained with the joint model are BLL candidates. Especially when considering the infrared, radio, and X-ray bands in combination, only BLL candidates have been predicted. One source from the UGS sample has been predicted as a BLL – the same source, which has been predicted as an AGN in section 6.4, and has been studied and discussed in more detail, confirming the statements made there. Due to this success, the 5 BLL candidates from the BCU sample have also been compared to the BLL and FSRQ sequences, presented in Figure 7.28. Similarities between the measurements and the BLL sequence are visible, and rough estimates of the redshift have been performed. Nevertheless, it has to be stressed that these redshift estimates have to be considered with extreme caution, since the observations have not been conducted simultaneously. The BLL candidate 3FGL J2346.7+0705 features the largest sum of BLL scores, and is studied in more detail in chapter 8.

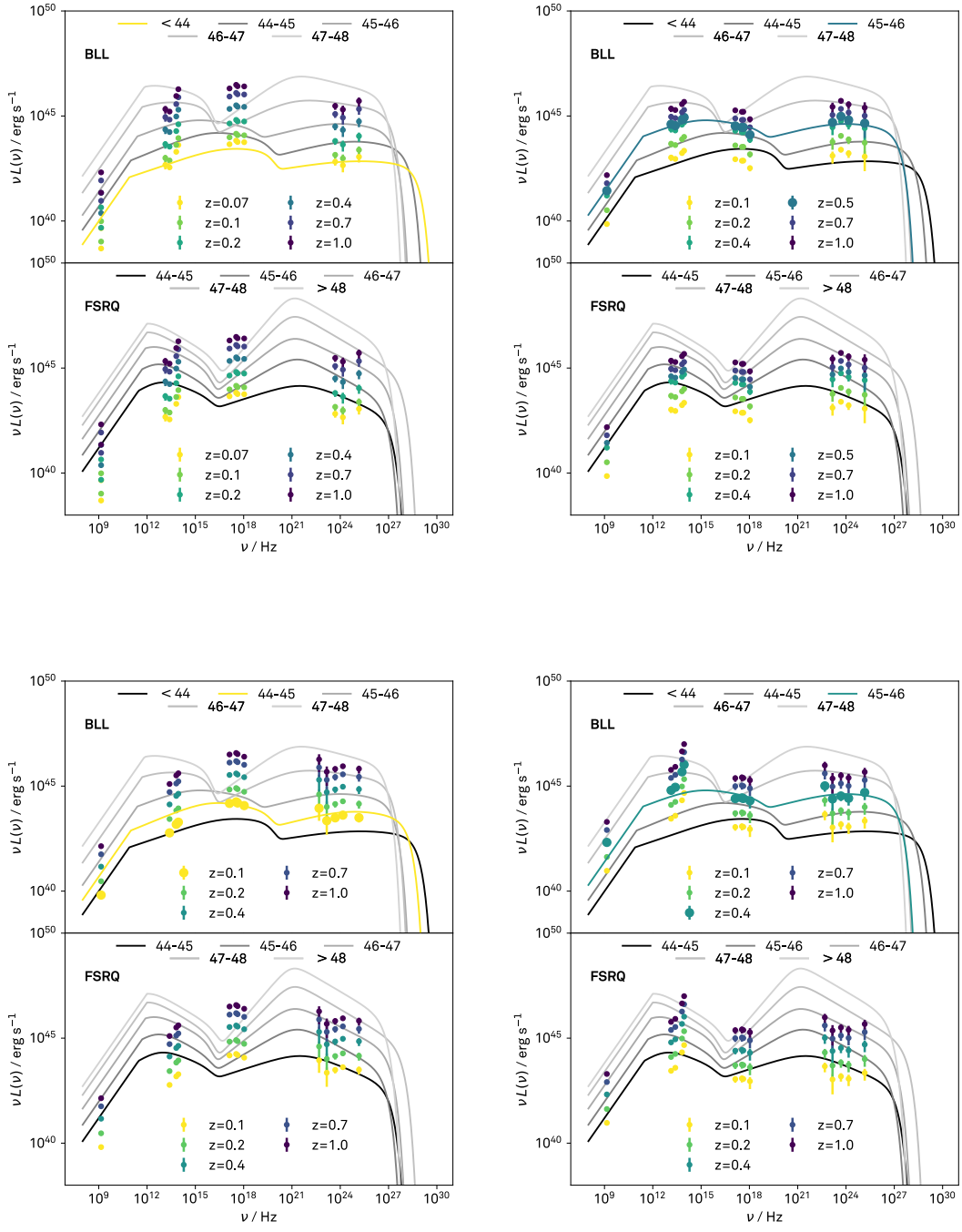

Figure 7.28: Luminosities of the BLL candidates 3FGL J0733.5+5153 (top left), 3FGL J1340.6–0408 (top right), 3FGL J1511.8–0513 (bottom left), and 3FGL J1543.5+0451 (bottom right), compared to the BLL (top) and FSRQ (bottom) sequences. The similarity between the high-lighted luminosities and color-coded spectral energy distribution underline the predicted BLL nature and estimates redshifts of roughly $z = 0.07$, $z = 0.5$, $z = 0.1$, and $z = 0.4$, respectively.

7.5 Conclusion

In addition to the conclusions drawn in the previous chapter, the capability to discriminate also BLLs, FSRQs, and non-blazars has been proven. For this purpose, the developed machine learning procedure has been adapted to this 3-class problem, e. g. regarding the performance evaluation. The generated classification models for the infrared, X-ray and radio wavebands have not only been applied to the UGS sample, they have also been utilized to specify the blazar type of the uncertain blazars in the 3FGL catalog.
The obtained results in terms of the precision and the number of candidates are summarized in Table 7.16. Comparing these results with those of the AGN classification, it turns out that the blazar models based on a single counterpart catalog provide more high-confidence candidates, although some of them are ambiguous. Considering multiple counterpart catalogs, the AGN classification procedure provides in total more high-confidence candidates than the one for blazars. This is to be expected, since the classification task to discriminate between BLLs, FSRQs and non-blazars is more extended and thus, more difficult than to differentiate AGNs and non-AGNs. Precisions above 0.9 have been achieved for both the single and the multiple counterpart catalog approach. However, the transfer of the precisions to the UGS and the BCU sample is not directly possible, since the training, the UGS and BCU samples exhibit different properties. The training sample features properties that enabled the association and source type affiliation, which was not possible for the UGS. Thus, the precision obtained from the associated sample underestimates the precision of the UGS sample. The BCUs have already been associated, i. e. they are more similar to the training sample, but the properties have not been distinctive enough for a source type affiliation. Since the BCU sample only contains a single counterpart per 3FGL source, simplifying the classification task, the precision is presumably overestimated. In the infrared and X-ray regime, the classification of BLLs performs better, while in the radio-regime the classification of FSRQs performs best.

Table 7.16: Overview of the obtained results. Depending on the considered counterpart catalog, different numbers of BLL and FSRQ candidates are achieved with different precisions, depending on the BLL and FSRQ score thresholds. Both with solely one counterpart catalog, as well as with multiple catalogs precisions of more than 0.9 can be obtained.

Counterpart Waveband	BLL Candidates Unassoc.	Uncert.	Precision (BLL)	FSRQ Candidates Unassoc.	Uncert.	Precision (FSRQ)
Infrared	147	167	0.94 ± 0.04	77	91	0.85 ± 0.03
X-ray	39	90	0.89 ± 0.05	4	27	0.84 ± 0.06
Radio	24	26	0.81 ± 0.09	4	10	0.78 ± 0.12
Infrared	147	167	0.94 ± 0.04	29	63	0.90 ± 0.04
X-ray	32	88	0.90 ± 0.05	1	19	0.90 ± 0.08
Radio	13	12	0.90 ± 0.09	1	3	0.91 ± 0.10
X-ray / Infrared	8	49	0.97 ± 0.01	1	19	0.90 ± 0.02
Radio / Infrared	3	15	0.97 ± 0.01	0	4	0.90 ± 0.03
Radio / X-ray	1	8	0.93 ± 0.02	0	2	0.93 ± 0.03
Radio / X-ray / Infrared	1	5	0.98 ± 0.02	0	0	0.94 ± 0.03

Five high-confidence BLL candidates with corresponding counterparts in the infrared, X-ray and radio have been found in the BCU sample. The source 3FGL J2346.7+0705 exhibits the highest sum of all three BLL scores, and is analyzed further in the next chapter. Although this source is included in the BCU sample, it has recently been associated with a BLL (Álvarez Crespo et al., 2016) by its optical spectrum (see chapter 8 for details).

8 Analysis of the Active Galactic Nucleus 3FGL J2346.7+0705

There is a moment
in every dawn when light floats,
there is the possibility of magic.

— Douglas Adams, *The Hitchhiker's Guide to the Galaxy*

The 3FGL catalog (cf. subsection 5.1.1) has a few things to say on the subject of the source 3FGL J2346.7+0705. Its position at $(RA, Dec) = (23^{h}46^{m}42.8^{s}, 07°05'10'')$ has been obtained in an analysis of four years of observations with the *Fermi*-LAT, which has also provided results of spectral fitting. The spectral shape has been determined to follow a power law of $\mathrm{d}F/\mathrm{d}E = F_0 \cdot \left(E/E_0\right)^{-\gamma}$ with $F_0 = 6.12 \times 10^{-14}\,\mathrm{cm}^{-2}\,\mathrm{MeV}^{-1}\,\mathrm{s}^{-1}$, $E_0 = 4090\,\mathrm{MeV}$ and $\gamma = 1.78$. Neither a curvature of the spectrum, nor a variability of the flux over time have been detected significantly.
Although no sources from the preceding 1FGL (Abdo et al., 2010) and 2FGL (Nolan et al., 2012) catalogs are linked to this source in the 3FGL catalog, the sources 1FGL J2347.3+0710 and 2FGL J2347.2+0707 are located at a very small distance. Comparing the physical properties of these sources, such as the spectral shape, the spectral index and the integral flux, and taking into account the uncertainties of these features, they appear very similar. It is concluded that these sources are associated with the same astrophysical object, and the missing link in the 3FGL catalog is stemming from the automatic association procedure. While the source could not have been associated to any source class or counterpart in another wavelength in the 1FGL and 2FGL catalogs, it has been associated with a blazar of uncertain type and the radio source TXS 2344+068 in the 3FGL catalog.
Based on the 3FGL catalog, the Third Catalog of Active Galactic Nuclei detected by the *Fermi*-LAT (3LAC, Ackermann et al., 2015b) was created, containing additional information e. g. on the associated counterpart, the fluxes in other wavelengths and the synchrotron peak position. According to this catalog, a specific blazar type could not have been established due to the lack of an optical spectrum. Nevertheless, the synchrotron peak position ν^{S}_{peak} has been identified by a reliable evaluation of the Spectral Energy Distribution. The source's peak position of $\nu^{S}_{\mathrm{peak}} = 10^{14.76}\,\mathrm{Hz}$ (not corrected for redshift) specifies it as an intermediate-synchrotron-peaked blazar.
Álvarez Crespo et al., 2016 have investigated optical spectra from the latest release (Alam et al., 2015) of the Sloan Digital Sky Survey regarding the redshift and type of blazars. The spectrum of 3FGL J2346.7+0705 shows the absorption feature Ca H & K (H and K lines of calcium) with an equivalent width of $EW = 0.3 - 0.9$ Å. Since blazars with equivalent widths below 5 Å are classified as BL Lacs, the blazar type of 3FGL J2346.7+0705 is specified to be an intermediate-peaked BL Lac. The specific absorption feature allows also an estimation of the redshift of $z = 0.17$. The corresponding source SDSS J234639.93+070506.8 is located at $(RA, Dec) = (23^{h}46^{m}39.93^{s}, 07°05'06.85'')$.

The source 3FGL J2346.7+0705 has been of interest also for researches with the MAGIC telescopes (see section 2.3). At the time of the 1FGL catalog, the object has been proposed as a dark matter clump candidate within an application for observation time in 2010 (Nieto et al., 2010) due to the lack of an obvious counterpart at other wavelengths and its steady and hard spectrum. After approximately 8 h of observations of good quality, no significant detection has been obtained, neither for the 1FGL position, nor for the 3FGL position in a re-analysis in subsequent years. In 2015, a continuation of the by then associated blazar has been proposed in another application of observing time (Paiano et al., 2015). The corresponding observations have been conducted in late 2016, and their analysis is presented in the subsequent section. Finally, the measurements with the MAGIC telescopes are set into context with measurements at other wavelengths.

8.1 MAGIC Standard Data Analysis

The analysis of 3FGL J2346.7+0705 in the very high-energy regime has been accomplished with the MAGIC standard software *MARS V2-17-2* in combination with the scientific software framework *ROOT V5-34-36*. The standard data analysis pipeline has been applied to observations of the object in 2016, and to observations of the Crab Nebula in a similar time span to verify the configurations and to evaluate the performance thereof.

8.1.1 Data Selection

Observations entering this analysis have to fulfill specific criteria, which are summarized in Table 8.1. Since all observations of the source 3FGL J2346.7+0705 in the year 2016 have been obtained in wobble mode (cf. section 2.3) under small zenith angles (below 35°), the Crab Nebula observations for the sanity check, the observations necessary for the creation of the gamma/hadron separation model (cf. section 9.1), and the Monte Carlo simulations are selected accordingly. In addition, the standard analysis requires particular light conditions, quantified by the maximum median of the dark currents (`DCMax`) in the camera of the MAGIC-II telescope. Further criteria concern the weather conditions, e. g. the presence of clouds, which is related to the sky temperature (`Cloudiness`) and the atmospheric transmission measured by the LIDAR. The trigger rate gives some additional indication of weather conditions, but also of hardware problems. The basic selection criteria can be easily applied with the *MARS* executable *quate*, employed on data files on *star* or *superstar* level. The nightly runbooks, daily checks, and so-called superplots have been inspected further for e. g. hardware problems or other abnormalities.
This inspection has revealed that in the considered time range of observations several problems with the LIDAR occurred. During one night in August, the transmission value dropped significantly for no apparent reason, while the trigger rate and the `Cloudiness` parameter behaved perfectly stable. By mid of November, the LIDAR system has been upgraded, i. a. the laser has been replaced by a much stronger one, requiring a new calibration. From time to time, it is not possible to operate the LIDAR or the movement of the LIDAR is faulty. If the latter is not spotted directly by the on-site shifters, the shooting direction of the LIDAR is outdated, and does not coincide with the observed source, and its result is mostly useless. Since these issues do not allow a homogeneous selection based on the LIDAR results, no automatic selection with *quate* has been performed, but the LIDAR measurements have been investigated manually.

Table 8.1: Criteria for the data selection, applied to all observations used in the presented analysis. The criteria concern weather, light, telescope and observation conditions.

Parameter	Condition
Observation mode	Wobble
Zenith angle	$< 35^{\circ}$
`DCMax`	< 2000 nA
`Cloudiness`	< 35
Transmission (manual inspection)	> 0.8 (@9 km)
Trigger rate	stable
Run duration	> 10 s
Hardware issues	none

Table 8.2: Summary of the observations in the year 2016 used for the analysis of the source 3FGL J2346.7+0705. Observations of the Crab Nebula have been employed for the sanity check of the analysis, while the data samples of the remaining sources have been incorporated in the creation of the model for the gamma/hadron separation. The dates refer to the end of the night. The exact subrun and run numbers are listed in Appendix B.1.

Source	Month	Days
3FGL J2346.7+0705	August	4., 7., 8., 9., 11., 13., 14., 16.
	October	21., 22., 29.
	November	27.
	December	5.
Crab Nebula	September	11., 13., 27., 29., 30.
	October	12., 30.
	November	01., 28.
	December	24., 27.
1ES 0229+200	October	7.
	November	1., 20.
	December	6.
M15	August	4., 13.
	September	23.
Triangulum II	August	29.
	September	9., 27.
	October	11., 23.
S3 0218+035	September	2., 14., 30.

The final observation time of 3FGL J2346.7+0705 amounts to 15.85 h. Apposite to the source's properties regarding the conditions outlined above and distributions thereof, observations of the Crab Nebula and sources with a low gamma-ray flux have been selected in the same time range. The dates of all used observations are listed in Table 8.2. The corresponding subrun and run numbers are itemized in Appendix B.1.

8.1.2 Low-Level Reconstruction

The gamma/hadron separation, and the direction and energy reconstruction requires the creation of corresponding models, which have been obtained with the executable *coach* with default settings, and with 50% of the available Monte Carlo simulations. The generated models have been applied to the observations of 3FGL J2346.7+0705 and Crab Nebula, and to the remaining 50% of the simulations with the executable *melibea*, as well with default settings.

By using the reconstructed simulations, the energy threshold of the conducted analysis, defined as the maximum of the energy distribution of the survived events, can be derived. For comparability, it is common practice to perform the extraction of the threshold on simulations generated with a spectral index of 2.6 (similar to the index of the Crab Nebula). Since the used simulations feature a spectral index of 1.6, they have been re-weighted with the inverse energy of the simulations. The energy threshold is dependent on the applied cuts, and has been determined for different ones. In *MARS*, specific cuts, optimized for particular energy ranges, are typically used, for instance during the significance detection.

The so-called low-energy (LE) cuts imply:

```
MHillas_{1,2}.fSize > 60
MHadronness.fHadronness < 0.28
```

The so-called full range (FR) cuts imply:

```
MHillas_{1,2}.fSize > 300
MHadronness.fHadronness < 0.16
```

The so-called high-energy (HE) cuts imply:

```
MHillas_{1,2}.fSize > 400
MHadronness.fHadronness < 0.10
MStereoParDisp.fEnergy > 1000
```

The corresponding energy thresholds have been derived to be 90 GeV, 265 GeV and 1180 GeV, respectively.

Since the energy threshold increases with the zenith angle, and the source 3FGL J2346.7+0705 features zenith angles between 21° and 35°, the threshold has been also evaluated for these conditions, and a threshold of 95 GeV has been determined for the cuts optimized to the low-energy regime.

8.1.3 Crab Nebula Sanity Check

The Crab Nebula features a very strong gamma-ray flux in the very high-energy regime, and is entitled the standard candle of high-energy astrophysics. Due to that, it is perfectly suited for determination of specific settings, and for a sanity check of this analysis.
The knowledge of the point spread function of the system and the analysis is important, i. a. for the choice of the size of the signal region, and for the generation of the sky maps. A common procedure to estimate the point spread function is to fit a double two-dimensional Gaussian to the θ^2 distribution of Crab Nebula observations with the executable *odie* with the command `Odie.doExtensionFit: 2`. The result of the fit has been used for all subsequent executions of *odie*.
odie provides also a suggestion for the choice of the size of the signal region, determined by an analytical approach, based on the point spread function. The suggested size has been chosen for the determination of the significance, and lead to a significance according to Li and Ma, 1983 of 120.05 σ for LE cuts, and 91.82 σ for FR cuts in 11.40 h of observation. The corresponding θ^2 distributions, together with additional information, are shown in Figure 8.1.

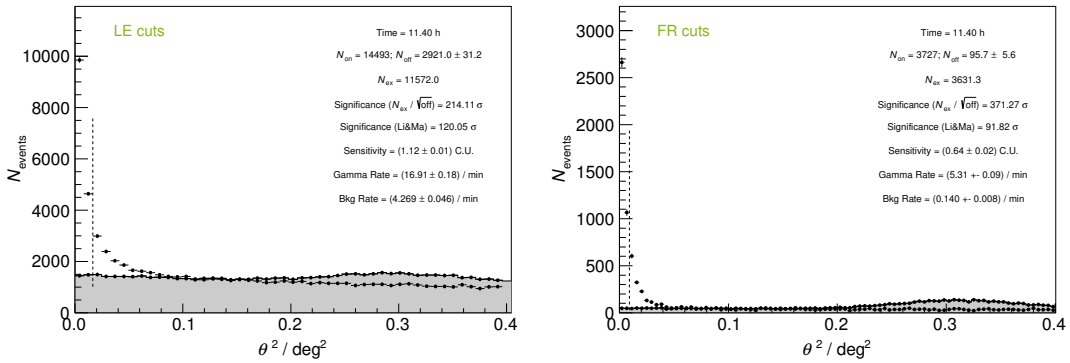

Figure 8.1: θ^2 distributions of the signal (on) and the background (off) source position for observations of the Crab Nebula. Cuts optimized for the low energy (LE) and the full range (FR) have been applied (cf. subsection 8.1.2).

Sensitivity values, stating the flux which can be detected with a significance of 5 σ in 50 h of observations in units of the Crab Nebula flux, have been calculated as well for the different cuts, and are summarized in Table 8.3. These values give some indication of the performance of the telescope system and the adapted analysis.

Table 8.3: Sensitivity values for different cuts (cf. subsection 8.1.2). They state the flux which can be detected with a significance of 5 σ in 50 h of observations in units of the Crab Nebula flux (C. U.).

Cuts	Sensitivity / C. U.
LE	1.12 ± 0.01
FR	0.64 ± 0.02
HE	0.87 ± 0.11

The corresponding spectral energy distribution and light curve have been derived with *flute*, and are shown in Figure 8.2 and Figure 8.3. The results are compatible with former published observations, as e. g. by J. Albert et al., 2008b or Aleksić et al., 2015, and the daily fluxes are stable. Different values for the efficiencies of the `Hadronness` and θ^2 cuts have been tested, and have been rated as robust. This is why the default values have been chosen for all subsequent executions of *flute*.

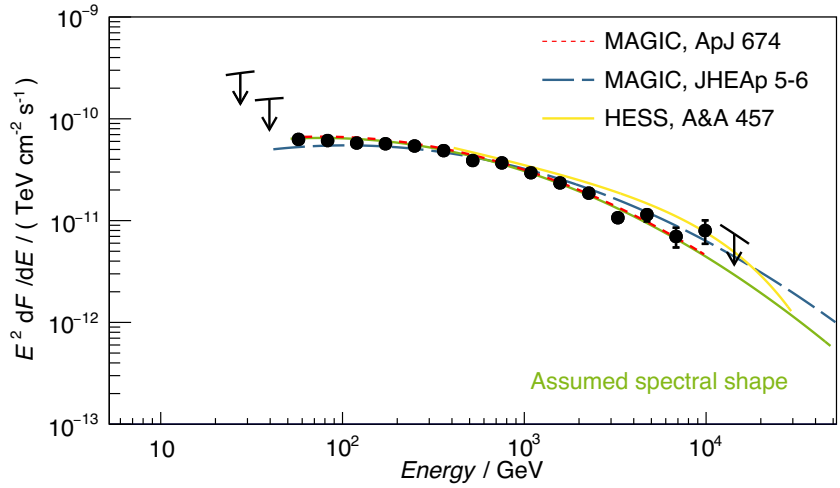

Figure 8.2: Spectral energy distributions of the Crab Nebula. The results of the observation (black) are compatible with former published observations (J. Albert et al., 2008b, Aleksić et al., 2015 and Aharonian et al., 2006), and indicate a reasonable adjusted analysis.

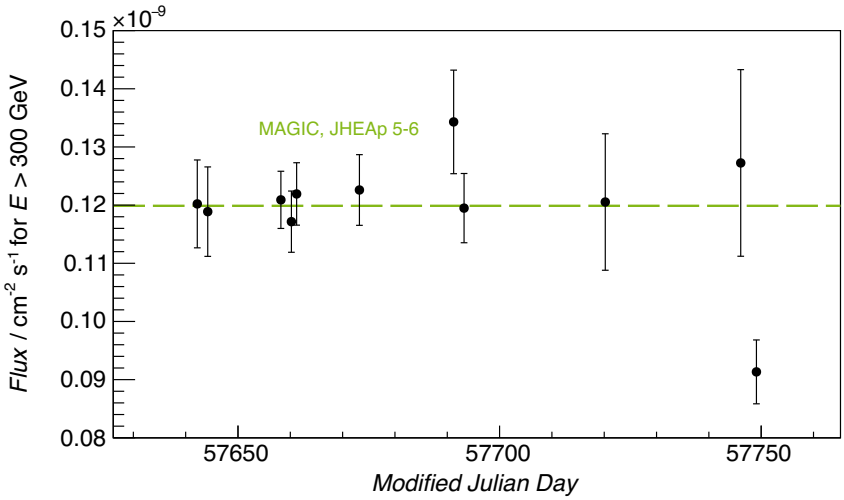

Figure 8.3: Light curve of the Crab Nebula in the year 2016. The daily fluxes are stable and compatible with former observations (green) of Aleksić et al., 2015.

8.1.4 Source Detection

The total and daily significances of the source 3FGL J2346.7+0705 have been determined with *odie* for the position of the associated source SDSS J234639.93+070506.8, obtained from optical observations, since its position is much more precise as the one of the *Fermi*-LAT. Table 8.4 lists these values. The corresponding total θ^2 distributions in Figure 8.4 do not reveal a significant excess, which is also supported by the respective significance values. Accordingly, neither a detection ($5\,\sigma$) nor a hint of detection ($3\,\sigma$) of the source can be claimed at this point.

Table 8.4: Daily observation times and significances according to Li and Ma, 1983 for the position of SDSS J234639.93+070506.8.

Date	t_{effObs} / h	Signif. LE / σ	Signif. FR / σ
2016/08/04	0.55	0.11	−0.71
2016/08/07	1.81	−0.60	−1.82
2016/08/08	1.63	0.38	0.40
2016/08/09	1.60	0.29	0.40
2016/08/11	1.30	0.85	0.91
2016/08/13	2.10	0.28	0.40
2016/08/14	0.92	−0.09	1.73
2016/08/16	0.56	−0.04	−1.00
2016/10/21	1.30	2.43	0.75
2016/10/22	0.92	−1.57	1.72
2016/10/29	1.15	0.82	−0.01
2016/11/27	0.43	0.58	1.51
2016/12/05	1.58	−0.17	2.20
Total	15.85	1.19	2.06

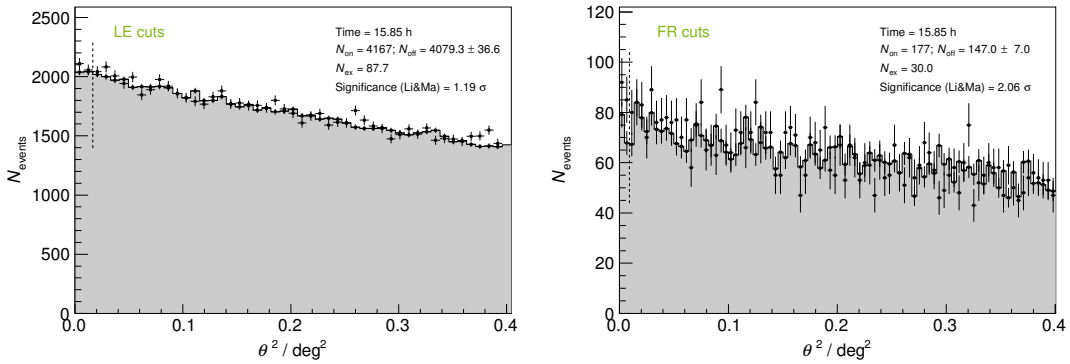

Figure 8.4: θ^2 distributions of the signal (on) and the background (off) source position for observations of 3FGL J2346.7+0705. Cuts optimized for the low energy (LE) and the full range (FR) have been applied (cf. subsection 8.1.2).

The region around the analyzed source is inspected on sky maps of Test Statistic (TS) values compiled with *caspar*. In the applied execution, the background has been estimated from the so-called wobble map, i. e. from regions outside of the signal region. The applied Gaussian smearing is dependent on the point spread function, defined by a single two-dimensional Gaussian. The values are translated from the double Gaussian fit and can be found in the output of *odie*. Figure 8.5 shows the described sky maps. No significant excess can be spotted in the sensitive region, in line with the *odie* results.
The used settings of both *odie* and *caspar* can be found in Appendix B.2.

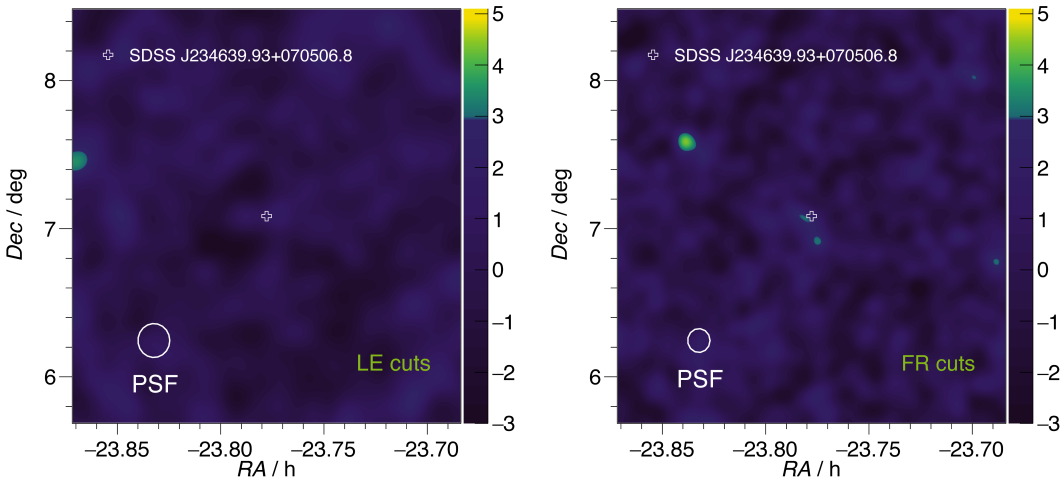

Figure 8.5: Sky maps of the Test Statistic (TS) values in units of σ of the region around 3FGL J2346.7+0705. Cuts optimized for the low energy (LE) and the full range (FR) have been applied (cf. subsection 8.1.2).

8.1.5 Energy Spectrum

Results related to the energy spectrum of 3FGL J2346.7+0705 have been derived with *flute* with the same settings found in subsection 8.1.3. Nevertheless, different efficiencies have been tested to check the robustness. The results turned out to be stable, taking into account fluctuations from low statistics. Due to the latter, and supported by the results obtained with *odie* and *caspar*, no significant spectral energy points could have been derived.
Instead, flux upper limits assuming a power law spectrum with a spectral index of −2 and the position of SDSS J234639.93+070506.8 have been extracted at a confidence level of 95%. These upper limits are illustrated in Figure 8.6, and compared to the flux of the Crab Nebula. In addition, significance values according to Li and Ma, 1983 have been computed for each single energy bin, based on the extracted number of on and off events. Table 8.5 summarizes these results.
No correction for the absorption by the extragalactic background light have been applied. For this topic and the further interpretation of the results, the reader is referred to section 8.2.
The settings to obtain theses results with *flute* can be found in Appendix B.2.

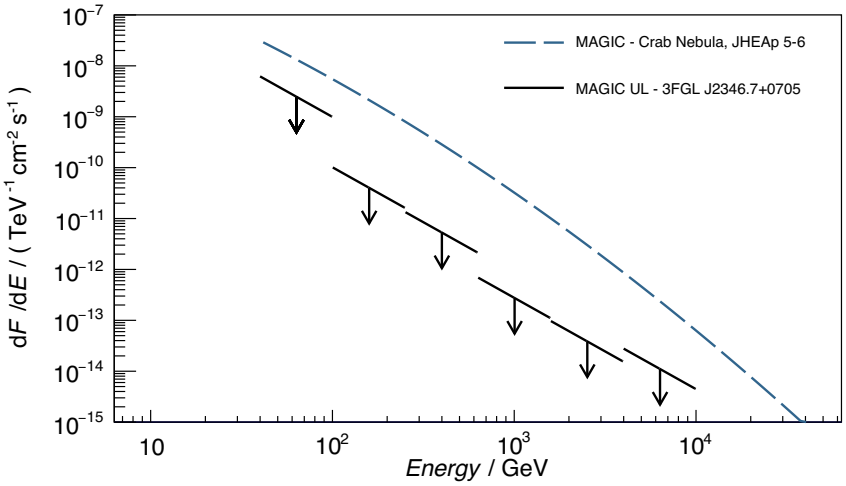

Figure 8.6: Upper limits at a confidence level of 95% for the differential energy spectrum of the source 3FGL J2346.7+0705. A power law spectrum with an index of −2.0 has been assumed for the determination of the limits. For comparison also the spectrum of the Crab Nebula is shown (Aleksić et al., 2015).

Table 8.5: Detailed information of each energy bin of the differential energy spectrum of the source 3FGL J2346.7+0705 concerning the number of on and off events N_{on}, N_{off}, the obtained upper limit of excess events N_{ex}^{UL}, the significance according to Li and Ma, 1983 and the upper limits of the differential flux at a confidence level of 95%, assuming a power law with an index of −2.

Energy / GeV	N_{On}	N_{Off}	N_{ex}^{UL}	Signif. / σ	$(dF/dE)^{UL}$ / (TeV^{-1} cm^{-2} s^{-1})
40 – 100	15409	15 157.0 ± 71.1	799.7	1.77	2.47×10^{-9}
100 – 251	3632	3639.3 ± 34.8	158.9	−0.11	4.01×10^{-11}
251 – 631	491	471.7 ± 12.5	94.4	0.76	5.30×10^{-12}
631 – 1585	48	50.7 ± 4.1	17.0	−0.33	2.74×10^{-13}
1585 – 3981	7	8.0 ± 1.6	7.3	−0.31	3.85×10^{-14}
3981 – 10000	3	3.3 ± 1.1	5.6	−0.16	1.11×10^{-14}

8.2 Multi-Wavelength Context

The results obtained with the MAGIC telescopes can be set into context with measurements at other wavelengths. The adjacent energy regime to the one of MAGIC is the regime of the Large Area Telescope (LAT) on board the *Fermi* satellite. The Third *Fermi*-LAT (3FGL) source catalog (Acero et al., 2015) provides information about the fluxes in various energy bands, the energy spectrum and the variability of the source. These flux measurements are depicted in Figure 8.7, together with the upper limits at a confidence level of 95% of the MAGIC measurements. The *Fermi*-LAT power law spectrum with a spectral index of 1.78 has been extrapolated to the very-high-energy regime of MAGIC. Moreover, the absorption by the extragalactic background light (EBL) has been derived from Domínguez et al., 2011 and has been applied to the extrapolated spectrum. The corresponding 1σ uncertainties of the spectrum are illustrated as well. Based on this, the direct extrapolation of the high-energy regime of *Fermi*-LAT to the very-high-energy regime of MAGIC can be excluded, e. g. due to the absorption of the EBL or intrinsic effects of the source. However, the upper limit in the regime between 100 and 250 GeV indicates an intrinsic cutoff in addition to the absorption of the EBL.
Although the 3FGL catalog does not state a significant variability of the source, it is important to check the flux state during the MAGIC and *Fermi* observations. The MAGIC observations have been conducted from August to December 2016, while the information from the 3FGL catalog is based on observations between 2008 and 2012. Figure 8.8 shows the *Fermi*-LAT light curve for the year 2016. The MAGIC observations have been marked in this light curve. During these observations, the source has not been in an exceptional state, and its flux is compatible to the average flux of the time interval that has been used to construct the 3FGL catalog. Thus, the spectra of MAGIC and *Fermi*-LAT connect.

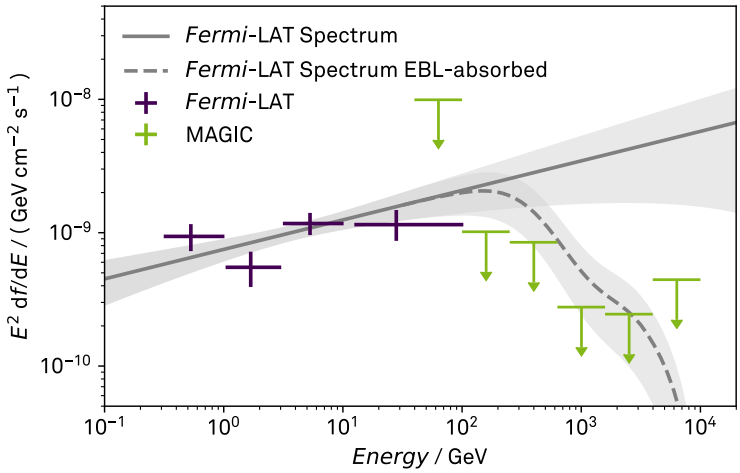

Figure 8.7: Spectral energy distribution in the gamma-ray regime. The MAGIC upper limits at a confidence level of 95% have been determined in the previous section, and the fluxes of *Fermi*-LAT have been extracted from the 3FGL catalog (Acero et al., 2015). The corresponding *Fermi*-LAT spectrum has been extrapolated to the energy regime of MAGIC. In addition, the spectrum is depicted with an absorption of the extragalactic background light (EBL) according to Domínguez et al., 2011. The shaded area illustrates the 1σ uncertainty of the spectrum. The MAGIC measurements exclude a direct extrapolation of the spectrum, and suggest intrinsic source effects in addition to the EBL absorption.

The long-term weekly *Fermi*-LAT light curve of the source is shown in Figure 8.9. No significant variability has been identified, and the weekly fluxes are compatible with the average 3FGL flux. This variability property is common for sources of the type BL Lac.
Considering the results from the previous chapters, information from further wavebands is included in the investigation of the source. In the previous chapter, searches for AGN and blazar candidates and their counterparts have been conducted using machine learning methods. The application of these methods has provided scores, indicating the source type, such as AGN, BLL or FSRQ. In the radio regime, an AGN score of 0.98, and a BLL score of 0.86 has been obtained. In the X-ray regime, the AGN score has amounted to 0.97, and the BLL score to 0.96. In the infrared regime, the method has lead to a AGN score of 0.98 and a BLL score of 1.00. Thus, the derived properties of the source are compatible to the determined BLL type by Álvarez Crespo et al., 2016. The corresponding counterparts are the following: The X-ray counterpart 1SXPS J234639.8+070508, the radio counterpart FIRST J234639.9+070506 and the infrared counterpart WISE J234639.93+070506.8.

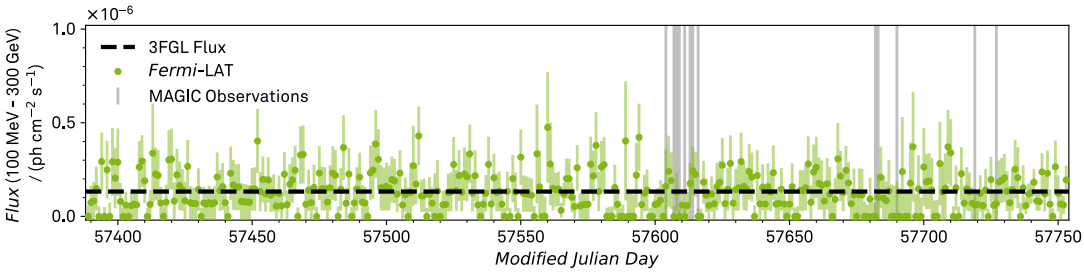

Figure 8.8: Daily light curve of *Fermi*-LAT measurements in an energy regime between 100 MeV and 300 GeV for the year 2016. The average flux between August 2008 and August 2012 corresponds to the measurements, used to construct the 3FGL catalog. In 2016, the daily fluxes do not significantly deviate from this average 3FGL flux.

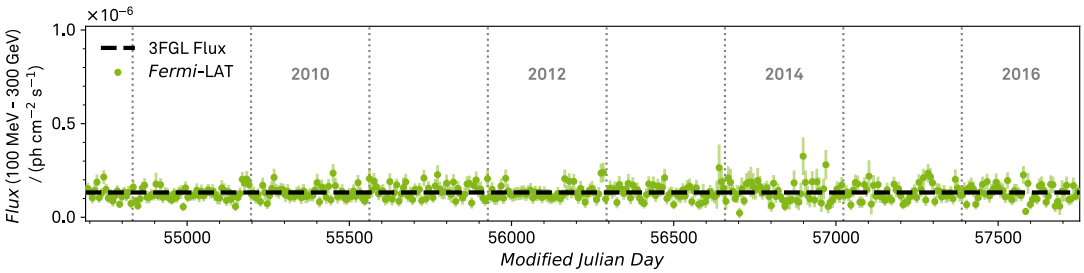

Figure 8.9: Long-term weekly light curve of *Fermi*-LAT measurements in an energy regime between 100 MeV and 300 GeV for the years 2008-2016. The average flux up to August 2012 corresponds to the measurements, used to construct the 3FGL catalog, and is depicted as well. No significant variabilities or different flux states have been detected.

The amount of counterparts has even been extended by querying astronomical databases[1]. All measurements have been summarized to construct the spectral energy distribution, depicted in Figure 8.10. A two-humped structure, typical for blazars, is clearly recognizable.

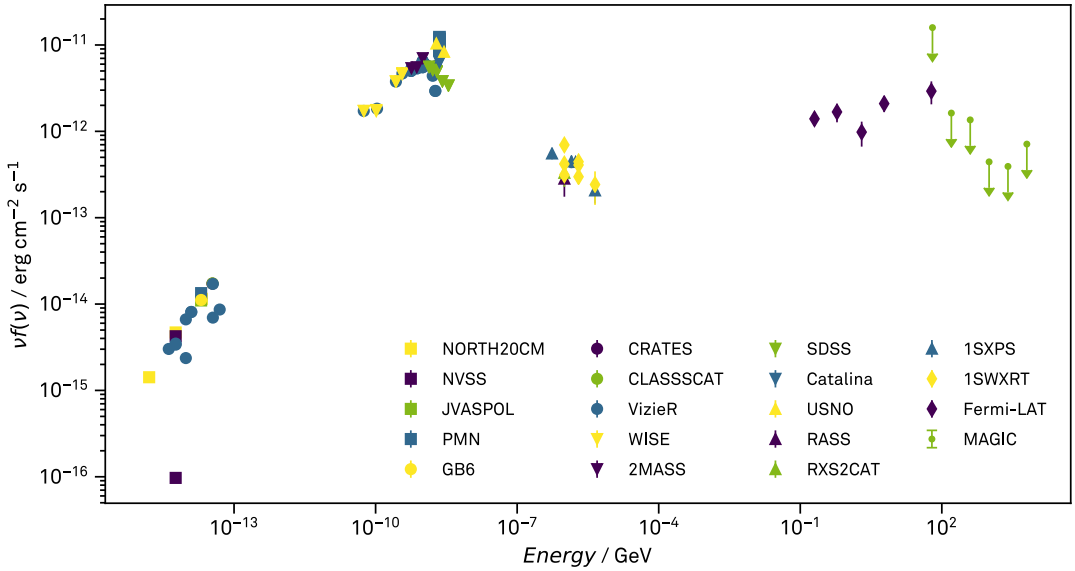

Figure 8.10: Spectral energy distribution of 3FGL J2346.7+0705 for the SDSS position. The typical two-humped distribution of blazars is recognizable. The depicted data originate from Myers et al., 2003, Healey et al., 2007, Jackson et al., 2007, Condon et al., 1998, A. E. Wright et al., 1994, Gregory et al., 1996, White and Becker, 1992, E. L. Wright et al., 2010, D'Elia et al., 2013, Evans et al., 2014, Voges et al., 1999, Boller et al., 2016, Acero et al., 2015. The MAGIC upper limits have been derived in the previous section.

8.3 Conclusion

Based on the 3FGL catalog, the source 3FGL J2346.7+0705 has been rated as an interesting candidate for observations in the very-high-energy regime, due to its steady and hard spectrum. Moreover, it stood out in the candidate list of the previous chapter, since it has achieved the highest BL Lac type score in all three wavebands.

The MAGIC telescopes observed this source for almost 16 hours in 2016. However, no significant detection has been made. Nevertheless, the obtained upper limits of 95% confidence have enabled the exclusion of a direct extrapolation of the high-energy spectrum to the very-high-energy regime. The upper limits have indicated a cutoff in the very-high-energy regime, which could not solely be due to the absorption of the extragalactic background light. Instead, it suggests an intrinsic absorption of the source in addition to the absorption by the extragalactic background light.

The spectral energy distribution over multiple magnitudes in energy, and the light curve in the high-energy regime are common for this type of objects.

[1]Part of this work is based on archival data, software or online services provided by the Space Science Data Center - ASI.

9 Revision of the Energy Reconstruction and Gamma/Hadron Separation for the MAGIC Telescopes

Don't panic.

— Douglas Adams, *The Hitchhiker's Guide to the Galaxy*

For stereoscopic observations with the MAGIC telescopes, the energy reconstruction is by default performed with the lookup table method (see section 2.3). This method makes use of only a very small number of features, and it estimates the energy of an event for every telescope individually, before the final reconstructed energy is obtained by the average of the individual energies weighted with the corresponding uncertainty. The disadvantages are that only a very small fraction of the available information (in terms of multiple features) is used, and that the benefits of stereoscopic observations are not fully exploited. Following this, the use of a Random Forest regression (see section 3.2), estimating the final reconstructed energy directly with additional features from both telescopes, is not far to seek. Independently from the method, the performance of the reconstruction is of interest, e. g. to adjust the bin width of the energy spectrum to the energy resolution, but also for comparisons between different instruments and methods. While in Aleksić et al., 2016b the performance is evaluated on data samples with high-quality cuts applied, this work studies a more realistic sample with quality cuts as typically applied in subsequent analysis steps. A further difference in this study is that no specific distribution for the performance measures is assumed.

The execution of a Random Forest classification in *MARS* for the gamma/hadron separation does not provide any performance measure apart from the out-of-bag estimate. Nevertheless, the knowledge of e. g. the robustness of the built classification model, and the estimated purity or in other words the hadronic contamination of the obtained data sample is extremely important. In *MARS*, the classification model is typically applied to data samples with less strict quality cuts as to the data samples the model is trained on, which is quite uncommon, and consequently subject to this work.

In the following, the implementation of the Random Forest algorithms (both regression and classification) used in *MARS* are reviewed, mainly concerning the performed split selection. A comparison between the results obtained with *MARS* and the machine learning library *scikit-learn* (Pedregosa et al., 2011) will give further indication of the similarity. Since an optimization of the Random Forest settings, a validation of the built models and a performance evaluation is rather complicated to be performed directly with *MARS*, these tasks are performed with *scikit-learn*. In case of similar implementations, the findings concerning optimized settings obtained with *scikit-learn* are transferred to *MARS*, and verified by an analysis of Crab Nebula data, the standard candle in very high-energy astrophysics.

9.1 Random Forest Algorithms in *MARS*

The largest differences between various implementations of Random Forest algorithms concern usually the determination of the best split in a certain feature at a specific node (referred to as *split selection*, see also section 3.2). The best split is usually selected according to the maximum difference of a quality measure Q_{parent} of the parent node and the sum of the quality measures Q_{left} and Q_{right} of the left and right child nodes, scaled to the number of events N at each node:

$$\text{criterion}_{\text{best split}} = Q_{\text{parent}} - \frac{N_{\text{left}}}{N} \cdot Q_{\text{left}} - \frac{N_{\text{right}}}{N} \cdot Q_{\text{right}} \,. \tag{9.1}$$

In *MARS*, the Gini impurity is used as the quality measure of a node in case of classifications (here: classification of gamma and hadron):

$$Q_{\{\text{node}\}} = \sum_{i=0}^{n_{\text{classes}}} p_i(1-p_i) = 1 - \frac{N_\gamma^2 + N_h^2}{N^2} \tag{9.2}$$

with N_γ, N_h the number of gamma and hadron events, $N = N_\gamma + N_h$ and $p_i = N_i/N$. Accordingly, the impurity is decreased during the tree growing.
For regressions (here: regression of energy E), the variance σ^2 of a node is used as the quality measure,

$$Q_{\{\text{node}\}} = \sigma^2 = \frac{\sum E_i^2}{N} - \left(\frac{\sum E_i}{N}\right)^2 , \tag{9.3}$$

in *MARS*, and is reduced during the growth of the tree.
These are typical definitions for the split selection and implemented accordingly in *MARS*, although stated differently in J. Albert et al., 2008a.

Another generic difference between distinct implementations is related to the output of the classifier and regressor. Based on the output of each tree in the forest, the final output can be determined by e. g. a majority vote, the mean or the median over the individual decision trees. In *MARS*, the mean is used for both the classification and the regression.

The model for the energy reconstruction is built with Monte Carlo simulations of gamma-ray events, whose source position in the camera have the same offset to the camera center as the data to be analyzed (usually 0.4°). Typically, the same simulations are also used for the gamma/hadron separation as the gamma-ray training sample, in combination with real observations, serving as the hadron training sample. The simulation of hadrons is rather challenging, because hadron-induced showers contain more interactions and different particles than gamma-induced showers of the same energy, and are much more complicated to model. Since the number of hadronic showers is more than a factor 1000 larger than the number of gamma-ray induced showers, the utilization of real observations is acceptable.

Before the models are created, the events of the data samples to be used for the creation process have to fulfill the following conditions:

```
# Only for the gamma/hadron separation:
MImagePar_{1,2}.fNumIslands <= 1
MHillas_{1,2}.fSize >= 50
MHillas_{1,2}.fSize <= 50000
MNewImagePar_{1,2}.fLeakage1 <= 0.15
MStereoPar.fValid >= 0.5
```

```
# Only for the energy reconstruction:
MStereoPar.fCherenkovRadius >= 4000
MStereoPar.fCherenkovDensity >= 0
MStereoPar.fValid >= 0.5
MStereoPar.fTheta2 <= 0.1
```

For the gamma/hadron separation, the gamma and hadron training samples are equalized to each other, i. a. to suppress artifacts. This equalization is conducted in a way that the number of gamma-ray and hadron events in bins of `log10(Size)`, `cos(Zd)` and `Az` is the same. Commonly, `cos(Zd)` is binned in 30 bins from 0 to 1, and `Az` in 1 bin from 0 to 2π, while `log10(Size)` is always binned in 90 bins from 1 to 5.5.

The basic settings of the Random Forest classification for the gamma/hadron separation are summarized as follows:

Number of features 12

Number of features considered at each node 3

Minimum number of samples for split 5

The features or generation of new features with the help of these features, respectively, are listed below. A detailed list of the generated features can be found in the input card of *coach* in Appendix B.2.

```
MHillas_{1,2}.fSize
MHillas_{1,2}.fWidth
MHillas_{1,2}.fLength
MStereoPar.fM{1,2}Impact
MStereoPar.fMaxHeight
MHillasTimeFit_{1,2}.fP1Grad
MPointingPos_1.fZd
```

9.2 Preparation of Data Samples

For the training of the models, Monte Carlo simulations of gamma-ray events and real observations are used. Additional tests are performed with observations of the Crab Nebula. These data samples are in general the same as in section 8.1 and fulfill the same criteria, and are listed in Appendix B.1.
To allow a direct comparison on single events, the study is performed with *melibea* files, i. e. the data samples have passed the gamma/hadron separation and energy reconstruction in *MARS* and contain the results thereof. Since the models must not be applied to the same data they have been trained on, and to perform the study on the same data (sub)samples which are used for the training in section 8.1, the following approach has been employed:
Equivalent to the standard analysis, the Monte Carlo simulations are divided in two halves. The models for the gamma/hadron separation, the energy and direction reconstruction are created with one half of the complete hadron sample, and applied to the second half. Additional models for the gamma/hadron separation are built analogously, but one file (out of 25) is left out at a time for the training. Like this, the model for the gamma/hadron separation can be applied to this specific file, together with the former models for the energy and direction reconstruction.

9.3 Comparison between Data and Monte Carlo Simulations

When using simulations in machine learning applications, it is always important that they reproduce the real observations, otherwise the obtained models and derived results are effectively useless. To know the level of compatibility of each feature, and to reject features showing an obvious mismatch between real gamma-ray events and Monte Carlo simulations (hereafter referred to as *data* and *MC*, respectively), a comparison is conducted. In this case, the comparison is not straight-forward, since no simulations of hadronic events are used in this analysis, and consequently a direct comparison of the recorded observations with simulations is not possible. For a comparison of the gamma-ray simulations with real gamma-ray events, hadronic events of the sample of real observations (which contains both hadronic and gamma-ray events) need to be suppressed.

A fair comparison requires the same underlying distribution, or rather the same underlying energy spectrum, and accordingly the knowledge of this spectrum. Observations of the Crab Nebula are well-suited, since this source is known as the standard candle in astrophysics, and is commonly used for these kind of test purposes. Aleksić et al., 2016b have fitted its spectrum with a curved power law of $\mathrm{d}F/\mathrm{d}N = F_0 \cdot (E/1\,\mathrm{TeV})^{a+b\,\log_{10}(E/1\,\mathrm{TeV})}$ with $F_0 = 3.39 \times 10^{-11}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{TeV}^{-1}$, $a = -2.51$ and $b = -0.21$ in the very high-energy regime.
The MC simulations have been generated with a simple power law with a spectral index of $\gamma = -1.6$, a maximum impact r_{max} of 350 m, and a number N_{gen} of 2.5 million events entering the atmosphere in an energy range between $E_{\mathrm{min}} = 10\,\mathrm{GeV}$ and $E_{\mathrm{max}} = 30\,\mathrm{TeV}$. Since the energy spectra do not coincide, the MC events with energies E are re-weighted as follows:

$$\mathrm{weight} = E^{-\gamma + a_{\mathrm{Crab}} + b_{\mathrm{Crab}} \cdot \log_{10}(E)}\,. \tag{9.4}$$

Furthermore, the distributions are normalized according to their observation time T. The observation time of the simulations is taking into account the re-weighting and their settings of the production,

$$T_{\mathrm{MC}} = \frac{N_{\mathrm{gen}} \cdot (\gamma + 1)}{F_{\mathrm{Crab}} \cdot r_{\mathrm{max}}^2 \cdot \pi \cdot \left(E_{\mathrm{max}}^{\gamma+1} - E_{\mathrm{min}}^{\gamma+1}\right)}\,. \tag{9.5}$$

This results in an observation time of $T_{\mathrm{MC}} = 203.2\,\mathrm{h}$. The observation time of the Crab Nebula observations amounts to $T_{\mathrm{Crab}} = 11.4\,\mathrm{h}$.

Before the comparison, the events have to fulfill the following conditions:

```
MHadronness.fHadronness < 0.95
MStereoParDisp.fTheta2 < 0.02
MStereoParDisp.fValid > 0
MHillas_{1,2}.fSize >= 50
```

The cut in `Size` concerns mostly events below the energy threshold, and is a standard minimum cut in the *MARS* analysis chain. In the subsequent studies it has always been performed. The cut in `Valid` ensures that only events with a proper reconstruction (at *melibea* level) enter the comparison, and is as well a standard cut in the data analysis. The cut in `Hadronness` suppresses a rather large amount of hadronic events, and is by default the extreme cut in *flute* when reconstructing the spectrum.

The cut in `Theta2` defines the size of the signal and the background region, respectively. It has been chosen smaller than the maximum cut in *flute*, but larger than typical signal cuts in *odie*. Like this, the signal-to-background ratio is still high, but at the same time a reasonable amount of events fulfills the condition, and the comparison is valid for a quite large range of typical `Theta2` values.

To extract the gamma-ray excess event distribution of the Crab Nebula observations, the signal and the background distributions have to be subtracted from each other. At the same distance from the camera center, one background region opposite to the signal region has been used to estimate the background. In order to put the compatibility of the gamma-ray excess with the simulation in concrete terms, the maximum distance between their cumulative distributions is determined.
Figure 9.1 and Figure 9.2 show examples of features without and with an obvious mismatch. The maximum distances of 0.02 and 0.29, respectively, support this assessment.

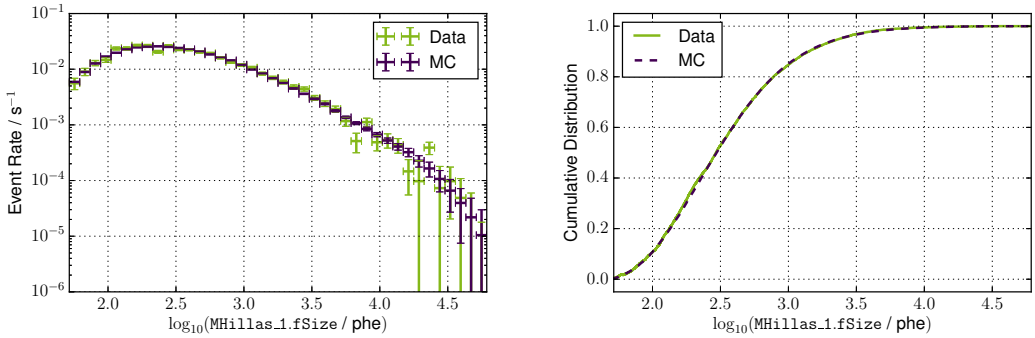

Figure 9.1: Comparison between gamma-ray simulations (MC) and excess events of Crab Nebula observations (Data) of the feature `MHillas_1.fSize`. The maximum distance between the cumulative distributions is 0.02. Both the event rate and the cumulative distributions are compatible.

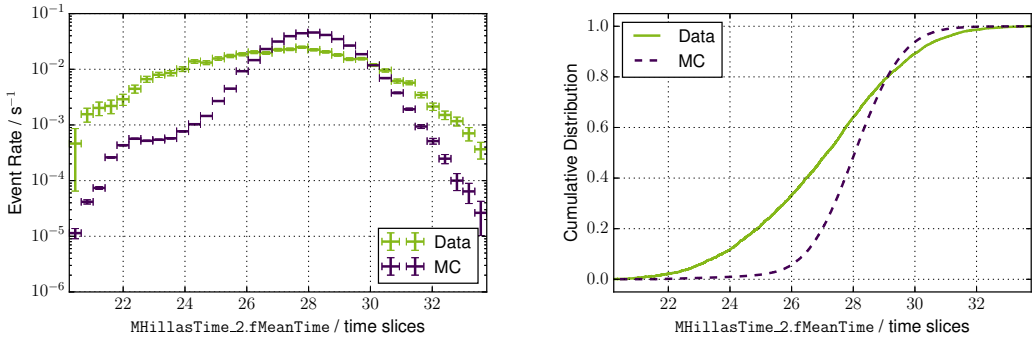

Figure 9.2: Comparison between gamma-ray simulations (MC) and excess events of Crab Nebula observations (Data) of the feature `MHillasTime_2.fMeanTime`. The maximum distance between the cumulative distributions is 0.29. Both the event rate and the cumulative distribution show a clear mismatch.

9.4 Feature Generation and Pre-Selection

For the application of machine learning methods, such as Random Forests, the selection of a set of proper features (regarding i. a. relevance and quality) is very important. The combination and transformation of existing features to new ones can simplify the growing process of the decision trees, e. g. in terms of the number of necessary splits and the number of features in the Random Forest, to perform the classification or regression task. This often matters for the run time of the method or for systematics the method implies.
A feature generation is applied to a base set of *MARS* features. The generation is conducted by applying mathematical operations to features, such as the logarithm, the absolute value, a multiplication, or more complex operations. Correction factors dependent on the zenith angle, the magnetic field, and the fraction of the shower in the camera image, are calculated in analogy to the corrections within the *LUT* method.

The available features from *MARS* and their description are listed in section B.4. Certain kinds of features are removed from the feature set, such as source dependent features, features related to the moon position and specific fits, and features containing irrelevant information for the considered classification and regression task, e. g. on the observation date.
Features with an obvious mismatch between real data and simulations are removed as well. The mismatch is evaluated by the largest distance between the cumulative distributions of a feature, as described in the previous section. Considering the distances of all features, a rather large gap in the values between a distance of 0.06 and 0.11 is identified. Following this, and supported by a visual inspection of the comparisons, a maximum distance of 0.10 is chosen to rank a feature as a feature without an obvious mismatch. The subsequent features have a larger maximum distance of 0.10:

```
MHillasTimeFit_{1,2}.fP{1,2}Const
MHillasExt_{1,2}.fTmeanCore
MHillasTime_{1,2}.fMeanTimeW
MHillasTime_{1,2}.fMeanTime
MHillasTime_1.fRMSTime
MStereoPar.fCherenkovDensity
MStereoPar.fCherenkovRadius
```

In addition, the feature `MHillasTime_2.fRMSTime` is removed to include only pairs of features. The resulting feature set comprises 64 features extracted directly from *MARS*, and 44 features generated thereof.

9.5 Optimization of the Energy Reconstruction

The optimization of the energy reconstruction is performed on a sample of Monte Carlo simulations with a spectral index of -1.6, which have been prepared as described in section 9.2. This sample comprises 220 275 events before any cut is applied. The optimization itself is conducted with *scikit-learn*, but the findings are transferred to *MARS* and are validated using observations of the Crab Nebula.
Various sets of selected features are studied, as well as different settings of the regressor. The performance of the energy reconstruction is evaluated by different approaches with regard to the method and the applied quality cuts. The results are also compared to those obtained with *MARS* with the *LUT* method.

9.5.1 Evaluation of the Energy Reconstruction

The performance of the energy reconstruction is evaluated on Monte Carlo simulations by inspecting the location and the width of the distribution of the relative differences between the reconstructed energy E_{rec} and the simulated energy E_{MC}. To study the energy dependency of the performance, the distribution $(E_{\mathrm{rec}} - E_{\mathrm{MC}})/E_{\mathrm{MC}}$ is typically investigated in bins of E_{MC}. The location and the width of the distribution quantify the bias and the resolution of the energy reconstruction, respectively, and can be defined by different methods. The resolution can be determined by calculating the standard deviation, or the half of the interquantile range between the 16% and the 84% quantile of the distribution. The mean or the 50% quantile (also referred to as the median) can be used for the determination of the bias. A further possibility is to fit a Gaussian function to the distribution, and take its resulting mean and standard deviation as bias and resolution.

A typical distribution of the relative differences is shown in Figure 9.3. It is a slightly asymmetric and skewed distribution. In the *MAGIC* collaboration it is common to apply the following performance cuts before conducting a Gaussian fit to determine the performance:

```
MImagePar_{1,2}.fNumIslands = 1
MNewImagePar_{1,2}.fLeakage1 < 0.15
MStereoParDisp.fTheta2 < 0.1
MHadronness.fHadronness < 0.5
MStereoParDisp.fCherenkovDensity > 0
MStereoParDisp.fCherenkovRadius > 4000
```

In the central region the distribution is rather similar to a Gaussian function, but at the margins non-Gaussian tails arise. The Gaussian fit is not sensitive to these tails, and the resulting resolution is underestimated, leading to an overestimation of the performance.
Both the standard deviation and the interquantile range are sensitive to these tails, but the interquantile range is more robust.

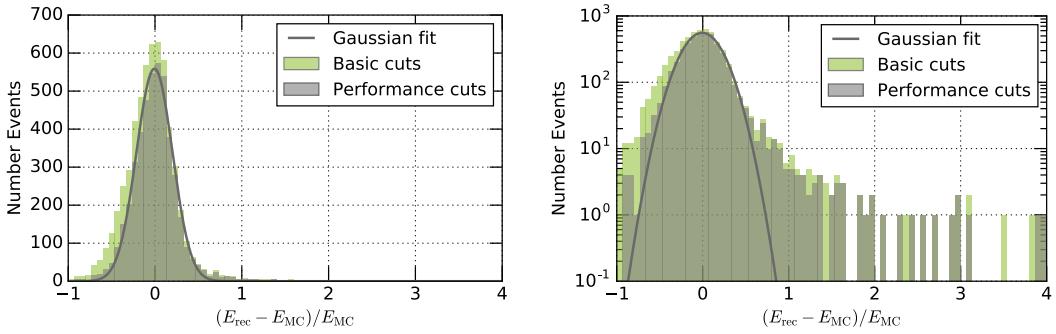

Figure 9.3: Example of the distribution of relative distances between E_{rec} (obtained with the lookup table method) and E_{MC} in an energy range of E_{MC} between 3 TeV and 4 TeV. Apart from some basic cuts, performance cuts as typically used by the *MAGIC* collaboration for the performance evaluation of the energy reconstruction have been applied to demonstrate their influence. A Gaussian fit has been performed on the distribution with applied performance cuts.

9.5.2 Feature Selection

For the following evaluation of the performance of different sets of selected features, the median is used to quantify the energy bias and the interquantile range for the energy resolution. Both approaches are very robust, do not require an assumption of a specific distribution, and the interquantile range is also sensitive to the non-Gaussian tails. The robustness of the Random Forest is measured with a 10-fold cross-validation (see section 3.3), and provides the opportunity to check for a possible overfitting. The consequent uncertainties of the performance values are visualized in the corresponding figures as vertical bars. The evaluation is performed in 5 bins per decade of the simulated energy E_{MC} from about 50 GeV to 20 TeV. These bins are visualized as horizontal bars.
For the investigations in this subsection, some particular settings need to be chosen initially: The output of each decision tree in the Forest is combined by computing the median over the trees. The number of features considered for the split selection depends on the number of features entering the Forest and is set to the square root thereof. Each split requires a minimum sample size of 5 events in the node, and each Random Forest comprises 200 decision trees.
For the creation of the Random Forest and the evaluation of the reconstruction, some very basic cuts are applied to the data sample:

```
MHillas_{1,2}.fSize >= 50
MStereoPar.fValid > 0
MStereoPar.fCherenkovDensity > 0
```

For the evaluation of the reconstruction an additional cut is applied, ensuring the validity of the direction reconstruction:

```
MStereoParDisp.fValid > 0
```

They are necessary to ensure e. g. the generation of specific features (to avoid a division by zero), or they are applied in any case in subsequent steps of the data analysis. A number of 206 984 events remains after the application of the basic cuts. Consequently, a number of roughly 186 000 events is used for the creation of one Random Forest (in the 10-fold cross-validation).
Many features are available for each telescope separately. They are always added or removed in pairs from a certain feature set, to allow individual models for each telescope for consistency checks (see subsection 9.5.6).

In a first step, the performance is evaluated for a feature set (*w/o feature generation*) comprising 64 features extracted directly from *MARS*. Another set (*w/ feature generation*) contains the same features, and in addition 44 features generated thereof. For the third set (*correlation removed*), features which have a correlation greater than 0.95 to another feature are removed. From the correlated features, the feature with the smallest feature importance obtained in the *w/ feature generation* set is removed. In this way, the number of features is reduced from 108 to 69. The results of the performance obtained with these feature sets are shown in Figure 9.4. In general, the feature sets comprising also generated features perform very similar and better than the set without generated features. As expected, the removal of correlated features does not degrade the performance, since effectively no information is withheld.

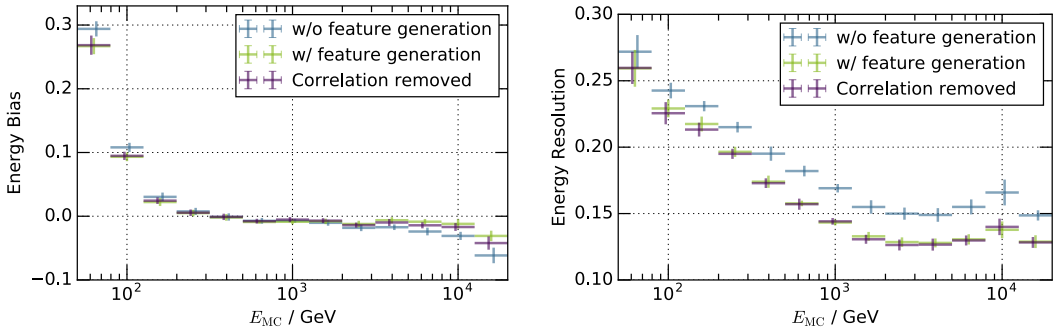

Figure 9.4: Comparison of the performance of the energy reconstruction for different feature sets. The feature sets comprising also generated features perform better. The removal of correlated features does not degrade the performance. The bias and the resolution have been determined with the median and the interquantile range, respectively. Only basic cuts have been applied during the evaluation. The vertical bars represent the uncertainty obtained within the cross-validation, while the horizontal bars indicate the energy range of the performance evaluation. The vertical bars are shifted horizontally for visibility purposes.

The reduction of the number of features is the next step. Based on the *correlation removed* set, two different approaches are pursued for this feature selection:
All features with a feature importance greater than 1/69 (the inverse of the number of features) are selected (*feature importance* set). In addition, the zenith angle is added, since the shower in the camera changes with the orientation of the telescope, especially for medium to large zenith angles ($> 35°$). The total number of features in this set amounts to 19.
In the second approach, a variant form of the backward elimination has been conducted. A small number of features with the minimal feature importances is removed, a new Random Forest is created with this set, and the performance is evaluated. If the performance degrades significantly, the removed features are added again, and the next lowest features are removed. This step is repeated iteratively until a similar number of features as the previous set is reached. Analogously, the zenith angle is added as well. The total number of features in this (*backward elimination*) set amounts to 18. Figure 9.5 shows the resulting performances. The bias is rather similar for the base and the feature selection sets, but the *backward elimination* set performs slightly better than the others, especially at the lowest and the highest energies. The resolution of this set is also significantly better as the *feature importance* set, in particular at energies below 1 TeV.
Correspondingly, the *backward elimination* set is chosen for the subsequent studies.
Figure 9.6 lists the selected features in combination with their feature importance. A large fraction of the features contain information about the light content of the shower. Due to that, they exhibit a large correlation to the energy, and are well suited for the energy reconstruction, which is underlined by the largest importances of these features. The importance of the zenith angle is rather small, since it is not correlated to the energy, but the correlation of other features to the energy is dependent on the zenith angle.

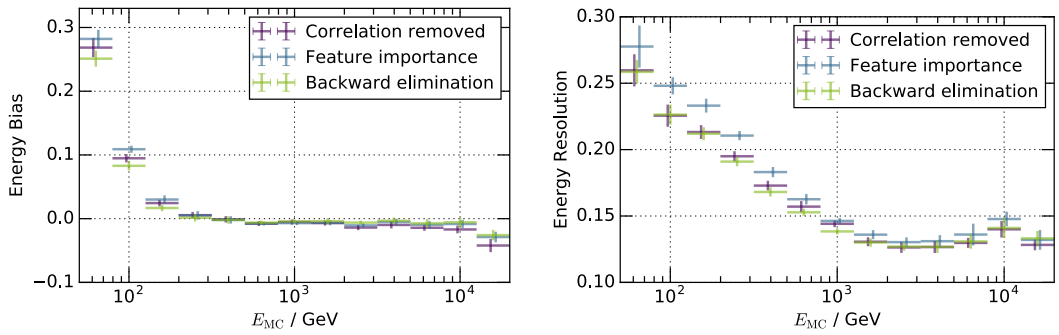

Figure 9.5: Comparison of the performance of the energy reconstruction for different feature sets. The reduction of features using a simple approach based on the feature importance leads to degradation of the performance. A backward elimination reduces the number of features, and results in more robust and better performance. The bias and the resolution have been determined with the median and the interquantile range, respectively. Only basic cuts have been applied during the evaluation. The vertical bars represent the uncertainty obtained within the cross-validation, while the horizontal bars indicate the energy range of the performance evaluation. The vertical bars are shifted horizontally for visibility purposes.

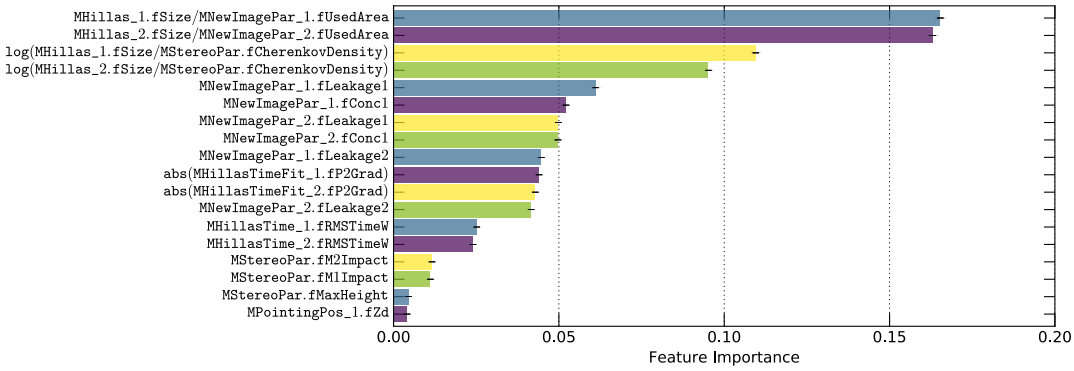

Figure 9.6: Feature importance of the selected features of the *backward elimination* set. Features including information about the light content of the shower have the largest importance and represent a large fraction of the selected features.

9.5.3 Settings of the Random Forest Regression

Several settings of the Random Forest influence the results. For instance, the number of decision trees or the number of features considered for the split selection at each node have a limited impact. Settings regarding the quantity to be estimated or the merging of the individual results of each decision tree to a final result have a larger effect. Figure 9.7 compares the performance between the median and the mean over the particular decision trees, and between the estimation of the energy and the logarithm of the energy. At large, the median performs better than the mean over the trees. If the logarithm of the energy is estimated, the performance improves slightly. Therefore, the logarithm of the energy and the median are used in the following studies.
The feature importances (Figure 9.8) of the estimation of the logarithmic energy changed compared to the linear energy estimation (cf. Figure 9.6).

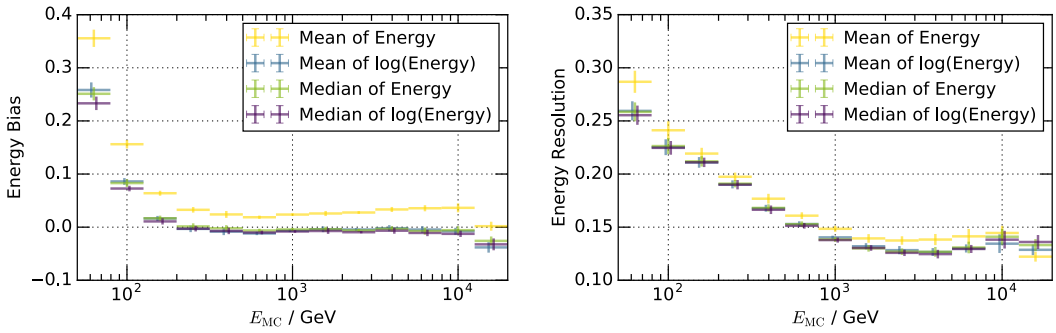

Figure 9.7: Comparison of the performance of the energy reconstruction for settings of the Random Forest regression. In general, the median over the decision trees performs better than the mean. The performance slightly improves, if the logarithm of the energy is estimated by the Random Forest regression. The bias and the resolution have been determined with the median and the interquantile range, respectively. Only basic cuts have been applied during the evaluation. The vertical bars represent the uncertainty obtained within the cross-validation, while the horizontal bars indicate the energy range of the performance evaluation. The vertical bars are shifted horizontally for visibility purposes.

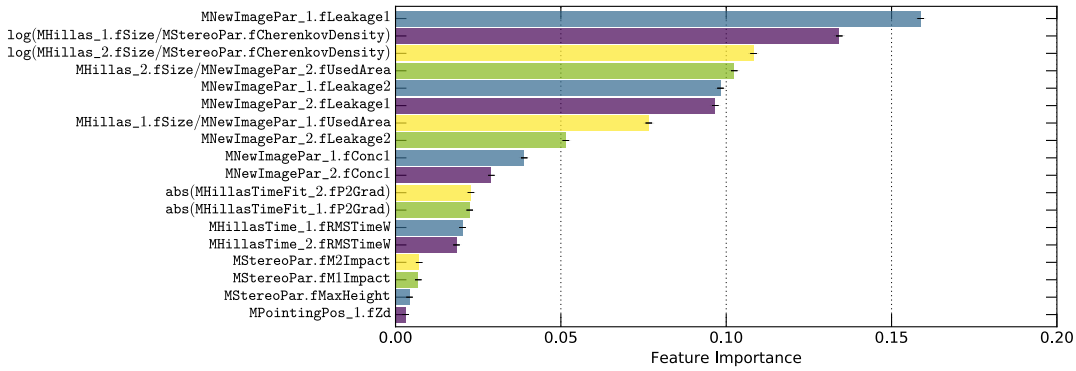

Figure 9.8: Feature importance of the selected features for an estimation of the logarithm of the energy. The importances changed compared to the estimation of the energy in linear scale (cf. Figure 9.6).

9.5.4 Verification with Crab Nebula Observations

In the previous subsections, the energy reconstruction has only been evaluated on Monte Carlo simulations. A verification of the found settings on real data can be performed with observations of the Crab Nebula – the standard candle in very high-energy astrophysics. For this purpose, the software package *MARS* is necessary, but it needs some adaptations to transfer the found settings. The starting point has been *MARS V2-17-2*.

In the executable *coach*, a Random Forest regression to estimate the energy with features from both telescopes in a single model has been implemented. Before, only individual models for each telescope could have been derived. The corresponding input card to steer the executable has been adapted accordingly, and has been extended with options to change the settings of the Random Forest. Additionally, the request of cuts via the input card has been cleaned up and restructured to associate each cut to a specific task (g/h separation, energy and direction reconstruction) without looking into the exact implementation of *coach*. The found settings have been integrated in the new input card, which can be found in Appendix B.3.

In the executable *melibea*, the possibility to use the median instead of the mean over the single decision trees in the Random Forest has been added. If in *coach* the logarithm of the energy has been estimated, this has to be taken into account in *melibea*. Both options can be activated in the input card, which can be found with the used settings in Appendix B.3. Some further basic adaptations for the application of the stereoscopic energy reconstruction model have been conducted by K. Ishio.

In the executable *flute*, the usage of an alternative container the reconstructed energy is stored in was corrupt, and has been rectified. An additional precut (`MStereoParDisp.fValid = 1`) has been added to the routine, determining the `Hadronness` and the θ^2 cuts from given efficiencies. This precut is implicitly conducted if the lookup table method has been used, since it sets the reconstructed energy to -1 for an invalid reconstruction (`MStereoParDisp.fValid = 0`). If the Random Forest regression has been selected, the reconstructed energies are not modified dependent on the validity of the reconstruction, and an explicit cut is necessary.

Different models for the energy reconstruction have been created with *coach* – one model with the lookup table method, the other with the Random Forest regression. Together with the models for the direction reconstruction and the particle type determination, these models are applied to the data in *melibea*.

The `Hadronness` and θ^2 cuts are determined in *flute* dependent on the reconstructed energy. The `Hadronness` cut is chosen in such a way that 90% of the simulated gamma-ray events survive the cut, but only cut values between 0.15 and 0.95 are allowed. Subsequent to the application of the `Hadronness` cut, the θ^2 cut is derived in such a way that 75% of the simulated gamma-ray events survive the cut, but only cut values between 0.01 and 0.20 are allowed. If an energy bin is lacking simulated events, no cut can be determined and all data ending up in this bin in subsequent steps will not survive. The derived cuts dependent on the reconstructed energy for both the lookup table method and the Random Forest regression are compared in Figure 9.9. Although the models for the particle type determination (`Hadronness`) and the direction reconstruction (θ^2) are the same, the cuts derived from efficiency are diverse, since the reconstructed energy changed and the cuts are dependent on it. In the medium-energy regime, the cuts are very similar, while in the low- and the high-energy range the cuts differ significantly. Below a reconstructed energy of $\approx$ 30 GeV and above $\approx$ 35 TeV, no cuts could have been derived for the Random Forest regression due to the lack of events in these energy bins.

This is explained by the nature of a Random Forests regression: The reconstructed energy of a decision tree is a combination of the energies of the events in a node. Thus, the estimate will always be in the range of the input energies, and extrema are reconstructed extremely rarely. In other words, a Random Forest regression always tends to the center of the distribution. This means that generally speaking low energies are overestimated and high energies underestimated.

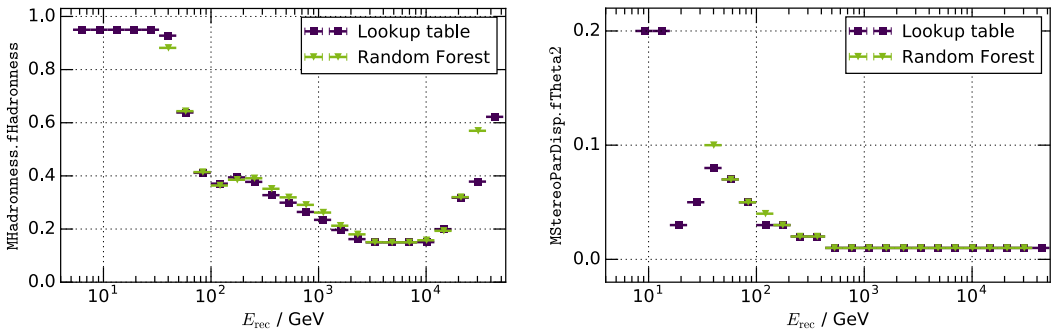

Figure 9.9: Comparison of `Hadronness` and θ^2 cuts obtained with *flute* for the energy reconstruction with the lookup table method and the Random Forest regression. The `Hadronness` cut is chosen in such a way that 90% of the simulated gamma-ray events survive the cut, but only cut values between 0.15 and 0.95 are allowed. The θ^2 cut is chosen in such a way that 75% of the simulated gamma-ray events survive the cut, but only cut values between 0.01 and 0.20 are allowed. If an energy bin is lacking simulated events, no cut can be determined and all data ending up in this bin in subsequent steps will not survive. In the medium-energy regime, the cuts are rather similar, while in the low- and the high-energy range the cuts differ significantly.

The resulting spectral energy distributions in terms of the reconstructed energy are shown in Figure 9.10. Both the spectra obtained with the lookup table method and the Random Forest regression are compatible with former results for Crab Nebula observations with MAGIC (J. Albert et al., 2008b, Aleksić et al., 2015). Anew, the results are very similar in the medium-energy range, while exhibiting differences for the lowest and the highest energies. As stated above, no events below $\approx$ 30 GeV survive the cuts in case of the Random Forest regression, and correspondingly neither a spectral flux point nor a flux upper limit can be derived. In the high-energy regime, the energy spectrum reconstructed with the Random Forest regression performs much more stable than the one with the lookup table method, and a significant flux point instead of only a flux upper limit could has been calculated. These unstable and insignificant results of the lookup table method can be explained with a large bias in the high-energy regime (see subsection 9.5.5 for details).
The conducting of an unfolding leads to spectral energy distributions in terms of a more correct energy. For the unfolding with the macro *CombUnfold*, the Tikhonov regularization has been used. The strength of the regularization has been adjusted in such a way that the noise of the covariance matrix of E_{rec} is the same as the one of E_{MC}. The bins of E_{rec} and E_{MC} need to be chosen by the user, dependent on e. g. the event distribution of E_{rec} and the collection area. The bins derived for the lookup table method are also valid for the Random Forest regression and lead to very similar spectra over the complete energy regime.

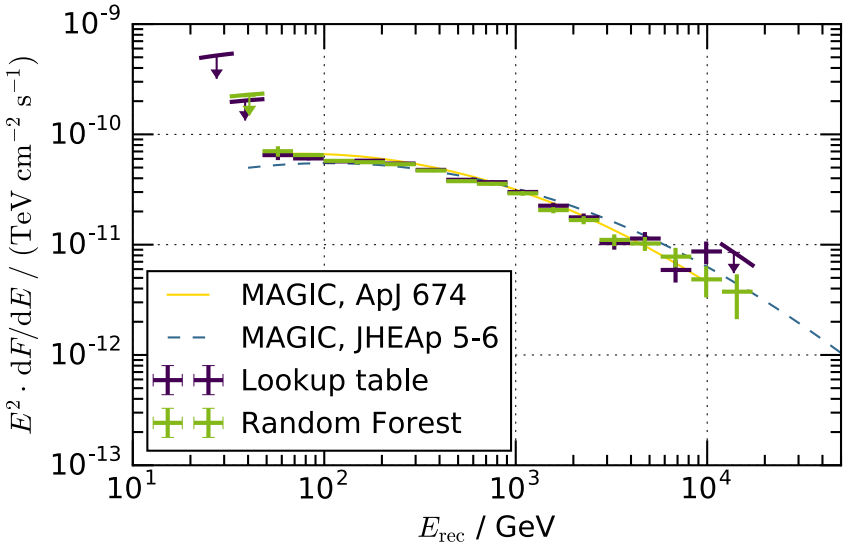

Figure 9.10: Spectral energy distribution obtained with *flute* for the energy reconstruction with the lookup table method and the Random Forest regression. Both spectra are compatible with former results (Aleksić et al., 2015, J. Albert et al., 2008b). The Random Forest regression is more robust and produces more significant flux points in the high-energy regime.

For the Random Forest regression, it is also valid to choose one bin more in each E_{rec} and E_{MC} towards higher energies. The corresponding spectra (see Figure 9.11) are compatible with former results for Crab Nebula observations with MAGIC. In contrast to the spectra in terms of the reconstructed energy (cf. Figure 9.10), the differences between the spectra obtained with the lookup table method and the Random Forest regression are very small. As discussed above, the unstable results in the high-energy regime are due to a bias. Since an unfolding is capable to correct possible biases, the differences between the two methods are much smaller. Nevertheless, the use of the Random Forest regression is able to reconstruct the spectrum to higher energies.
A light curve, representing the time-resolved flux, has been created with *flute*, and is shown in Figure 9.12. The results of both the lookup table method and the Random Forest regression agree within their uncertainties, and are compatible with a former result obtained by MAGIC for Crab Nebula observations. It is noteworthy that the fluxes are systematically lower for the Random Forest regression (except for one day). This is explained with the larger bias of the lookup table method (cf. Figure 9.13): The energy reconstructed with the lookup table method is overestimated. Correspondingly, the flux is integrated starting from lower energies than the flux derived with the Random Forest regression, resulting in this systematic effect. In general, the light curve obtained with the Random Forest regression is more stable (indicated by a lower reduced chi-squared of a constant fit), but features a lower flux level.

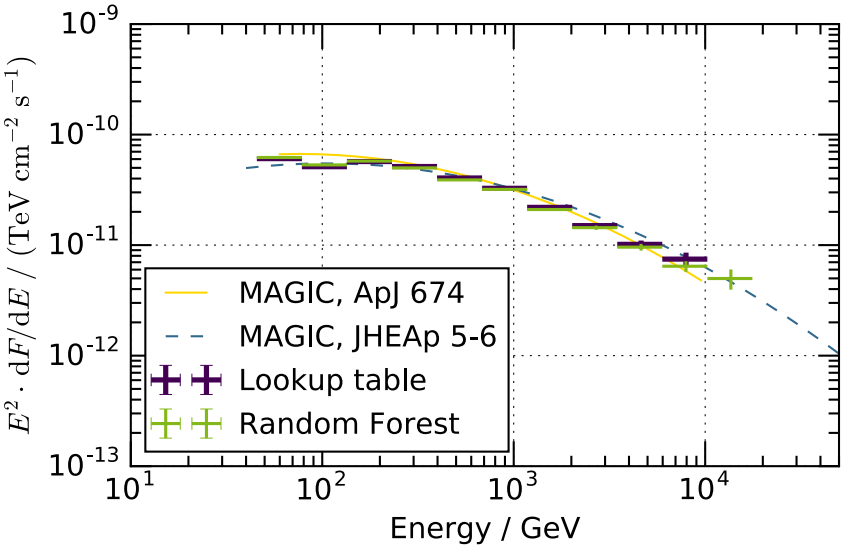

Figure 9.11: Spectral energy distribution, unfolded with *CombUnfold*. The bins in E_{rec} and E_{MC} have been chosen separately. The results are very similar and compatible with former measurements. Using the Random Forest regression, it is possible to reconstruct the spectrum to higher energies.

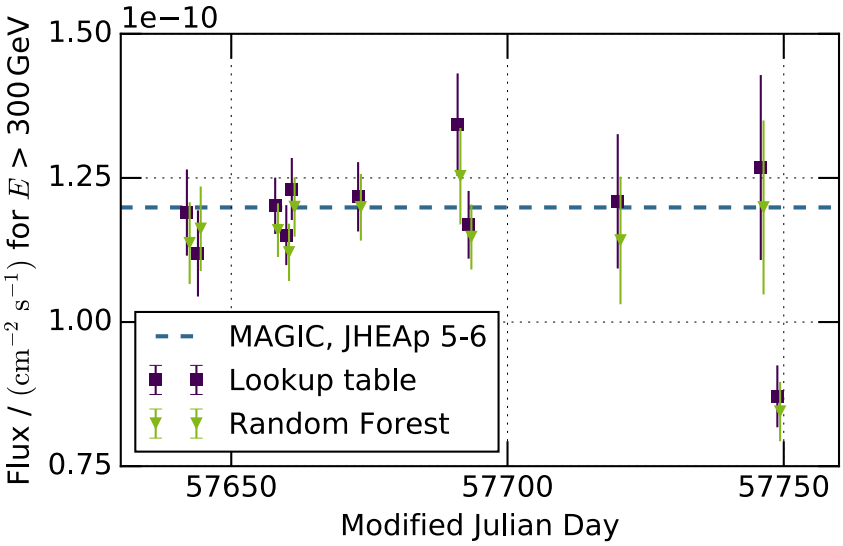

Figure 9.12: Light curve, representing the daily flux, obtained with *flute*. The fluxes derived with the lookup table method and the Random Forest regression are consistent and compatible with former measurements. The fluxes are shifted horizontally for visibility purposes.

9.5.5 Performance of the Energy Reconstruction

The performance of the energy reconstruction is visualized in Figure 9.13 in terms of the bias and resolution. The direct comparison between the two approaches (with applied `Hadronness` and θ^2 cuts according to Figure 9.9) reveals that the Random Forest regression generates better performance values starting from $\approx$ 200 GeV. Especially in the high-energy regime, the regression performs significantly better in both bias and resolution. In the low-energy range, the performance is slightly better for the lookup table method.

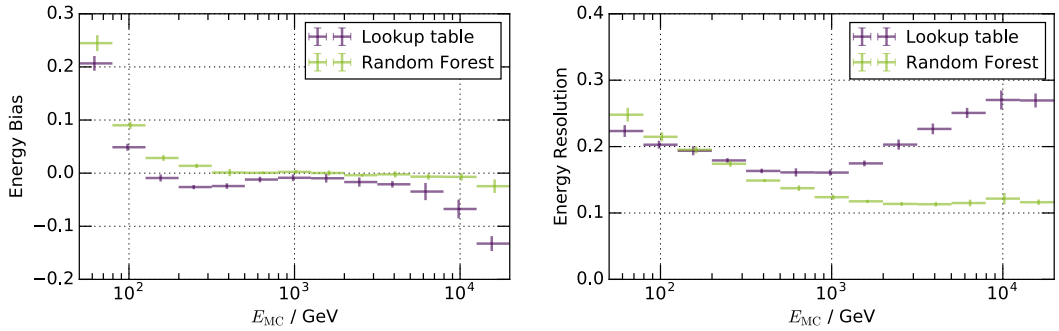

Figure 9.13: Comparison of the performance between the lookup table method and the Random Forest regression. `Hadronness` and θ^2 cuts according to Figure 9.9 have been applied. In the low-energy range the lookup table method performs slightly better, while in the medium- to high-energy range the Random Forest regression performs significantly better. The vertical bars represent the uncertainty obtained with a bootstrap validation (lookup table method) or a cross-validation (Random Forest regression), while the horizontal bars indicate the energy range of the performance evaluation.

Both approaches tend to an overestimation of the reconstructed energy (positive bias) for low energies and an underestimation (negative bias) for high energies. Different matters are converging in this regard: In both approaches the reconstructed energy is the average of a number of events (in a final node of a decision tree or in a bin of the lookup tables). Following this, the reconstructed energy will always be in the energy range of the accepted Monte Carlo simulations if no further correction or extrapolation is applied, and extrema are rather rare. In the high-energy range, the further difficulty of events not fully contained in the camera arises. The features `Leakage1` and `Leakage2` denote the ratio between the light in the (one or two) most outer ring(s) and the total light content, and values larger than zero indicate events only partly contained in the camera. The larger these values, the more of the shower is generally outside of the camera. Nonetheless, these events can be reconstructed – the better the estimate of the fraction of the shower outside of the camera, the better the reconstruction. In the lookup table method, only a correction factor dependent on `Leakage2` is determined and applied, while much more information enters the Random Forest regression.
For instance, by including `Leakage1` in addition to `Leakage2`, information regarding the gradient towards the camera border is derived, which is a better indicator of the fraction of the shower not contained in the camera. It can be assumed that the estimate of the shower fraction outside of the camera is more precise within the Random Forest regression, and correspondingly the compensation of such events is better. A further reason for the large differences in this energy regime between the methods is the condition `Leakage2` < 0.2 when creating the lookup tables. Events with a larger `Leakage2` are not included in the model, and the reconstruction of these events will be difficult.

As outlined above, the bias influences subsequent analysis steps, such as the creation of the light curve or the determined `Hadronness` and θ^2 cuts. The different biases for the lookup table method and the Random Forest regression need to be kept in mind when comparing analysis results.
In Figure 9.13 the performance is evaluated only using Monte Carlo simulations. Their validity is examined following the same approach as in section 9.3 by comparing the simulations with real data. In addition to the cuts stated in the mentioned section, `Hadronness` and θ^2 cuts according to Figure 9.9 have been applied.
Although the performance of the lookup table method in the low-energy range is better than the regression, the comparison of the event rate distributions (see Figure 9.14) reveals that these findings cannot be transferred to real data in their entirety. The distributions of the reconstructed energies obtained with the lookup table method show a significant mismatch below $\approx$ 50 GeV. The maximum distances between the cumulative distributions support the larger consistency of the distributions of the reconstructed energies with the Random Forest regression: The maximum distance is 0.07 for the lookup table method, and 0.04 for the Random Forest regression.

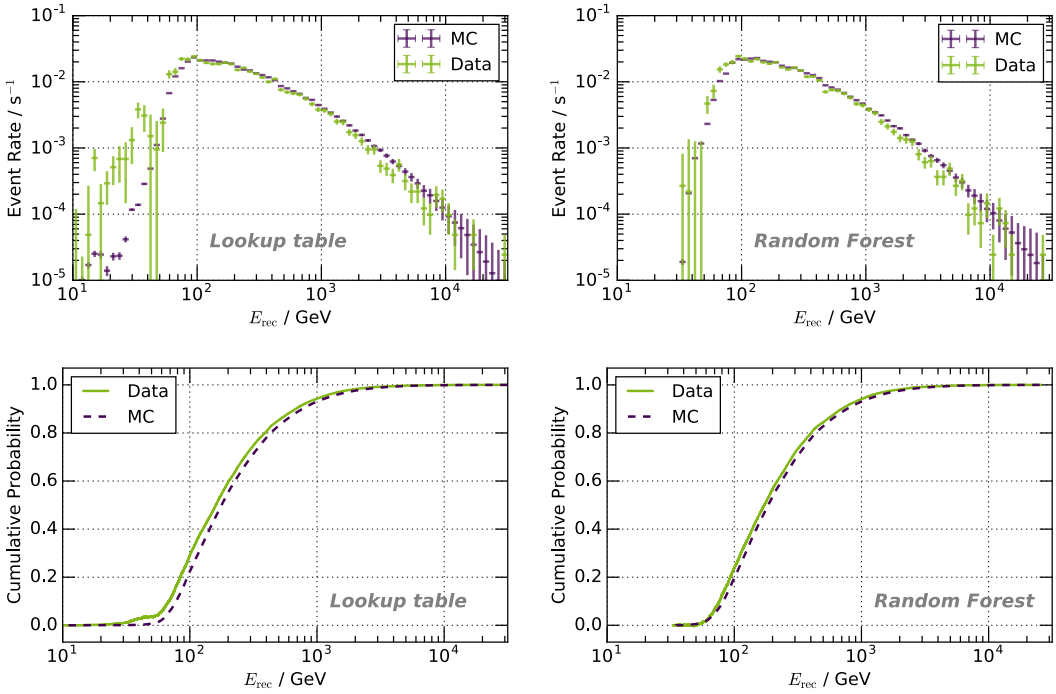

Figure 9.14: Comparison between gamma-ray simulations (MC) and excess events of Crab Nebula observations (Data) of the reconstructed energies. The maximum distance between the cumulative distributions is 0.07 for the lookup table method, and 0.04 for the Random Forest regression. In the low-energy regime a significant mismatch is noticeable. `Hadronness` and θ^2 cuts according to Figure 9.9 have been applied.

Plotting the reconstructed energy E_{rec} against the simulated energy E_{MC} illustrates in which bins of E_{rec} events of a specific energy bin E_{MC} end up, or in other words, where events migrate. This illustration is called migration matrix, which is shown in Figure 9.15 for both reconstruction methods. The migration matrix is utilized in unfolding procedures, where it is also known as the response matrix.

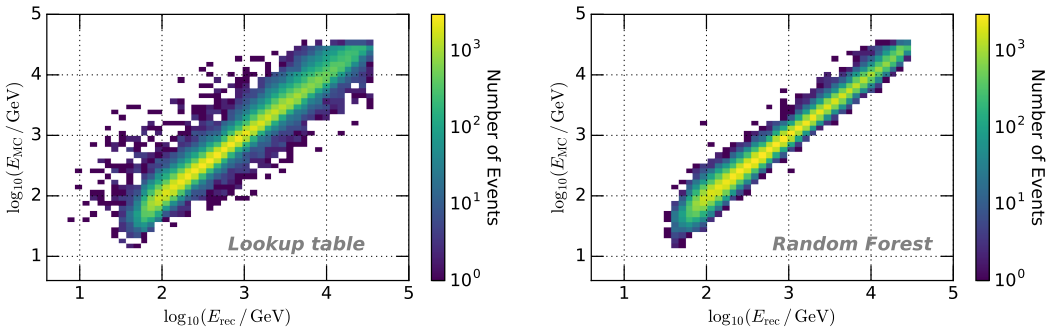

Figure 9.15: Migration matrices derived for the lookup table method and the Random Forest regression. The resolution is better for the Random Forest regression. Due to the applied `Hadronness` and θ^2 cuts according to Figure 9.9, the lookup table method reconstructs more events with low energies.

By means of the width of the two-dimensional distributions and the amount of outliers, the better resolution of the Random Forest regression is clearly identifiable. In principle, this allows the choice of smaller energy bins in a subsequent unfolding. A further difference is that the lookup table method estimates more events with a low reconstructed energy, due to to the application of `Hadronness` and θ^2 cuts according to Figure 9.9. In total, 130 489 events survive the cuts in case of the Random Forest regression, and 128 842 events in case of the lookup table method.
The executable *odie* has been applied to 11.4 h of observations of the Crab Nebula to compare the performance in terms of significance and sensitivity between the two methods. The condition of a reconstructed energy larger than zero has been added to the standard cuts conducted in *odie*, since some events obtained with the lookup table feature a reconstructed energy of −1 and are not removed during the validity check (`MStereoParDisp.fValid>0`). The significance is calculated according to Li and Ma, 1983. The sensitivity denotes the flux, which can be detected with a significance of 5 σ in 50 h of observations. Three background regions have been used to determine the number of background events. The results are summarized in Table 9.1 for different cuts (cf. subsection 8.1.2). By analogy with the migration matrices, more events pass the cuts if reconstructed by the regression. The significance and sensitivity are very similar in general, the Random Forest regression performs only slightly better for *LE* and *FR* cuts. For *HE* cuts, the lookup table method performs slightly better. But the comparison for these particular cuts should be interpreted with caution, since in contrast to the *LE* and *FR* cuts (in `Size` and `Hadronness`), the *HE* cuts comprise an additional cut in the reconstructed energy at 1 TeV. Due to different reconstruction methods and different biases, the *HE* cut is also different for both methods.

Table 9.1: Comparison of the results derived with the executable *odie* for different cuts (see subsection 8.1.2 for details). The significance according to Li and Ma, 1983 and the sensitivity have been calculated with the number of events N_{on} in the signal (on) region and N_{off} in the background (off) region for 11.4 h of observation of the Crab Nebula. Using the Random Forest regression, more events survive the cuts. Nonetheless, the resulting significances and sensitivities are compatible.

	N_{on}	N_{off}	N_{ex}	Signif. / σ	Sens. / C.U.
LE cuts					
Lookup table	14 409	2889.0 ±31.0	11 520.0 ±124.0	119.96	1.11 ±0.01
Random Forest	14 501	2914.3 ±31.2	11 586.7 ±124.4	120.22	1.11 ±0.01
FR cuts					
Lookup table	3717	93.3 ±5.6	3623.7 ±61.2	91.86	0.64 ±0.02
Random Forest	3731	94.0 ±5.6	3637.0 ±61.3	92.01	0.64 ±0.02
HE cuts					
Lookup table	696	5.3 ±1.3	690.7 ±26.4	42.26	0.80 ±0.11
Random Forest	705	7.0 ±1.5	698.0 ±26.6	42.15	0.91 ±0.11

9.5.6 Consistency Check

Another check of the validity of the energy reconstruction is to examine the consistency of the individual results of each telescope. For this purpose, the energy is reconstructed separately for both telescopes, i. e. two Random Forest regressions are performed, each with features from only one single telescope. The beforehand selected features (cf. Figure 9.8) serve as starting point. Apart from few stereoscopic features and information regarding the pointing direction of the telescope, all features in the set are available for each telescope individually. In total, 10 features are used for one Random Forest regression. The resulting reconstructed energies are denoted as $E_{rec,1}$ and $E_{rec,2}$. The term $E_{rec,1+2}$ refers to the reconstructed energy of the previous section and the features stated in Figure 9.8. Analogous to the lookup table method, the weighted arithmetic mean is calculated from the individual reconstructed energies $E_{rec,1}$ and $E_{rec,2}$. The weights are chosen as the inverse of the relative uncertainty of the reconstructed energies. As before, `Hadronness` and θ^2 cuts according to Figure 9.9 have been applied.

The results of the stated Random Forest regressions are compared in Figure 9.16. The results of the individual reconstructions are very similar, both in terms of the bias and the resolution. The mean of these individual reconstructions is alike the combined reconstruction of the previous subsection. In general, the combined reconstruction performs slightly better. Especially in the low-energy regime, the performance is better. This is explained by the fact that low-energy events comprise only few photons. Larger statistics and as much information as possible is important so that Cherenkov photons stick out from fluctuations of the night sky background, which can be accomplished by assembling information from two telescopes.

By examining the two-dimensional distribution of $E_{rec,1}$ and $E_{rec,2}$, Figure 9.17 checks the consistency of the individual reconstructed energies event-wise for both Monte Carlo simulations (top) and excess events of Crab Nebula observations (bottom). This consistency check is also very useful for an inter-telescope calibration.

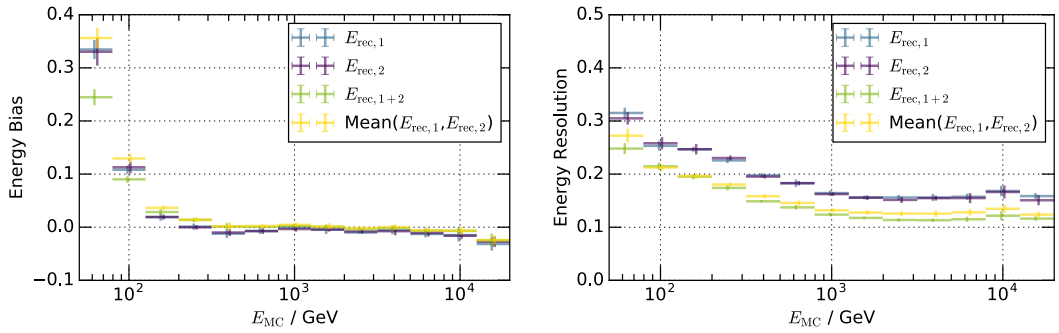

Figure 9.16: Comparison of the performance between stereo and mono regression. For the reconstructed energies $E_{\mathrm{rec,1}}$ and $E_{\mathrm{rec,2}}$ only stereo features and features from the corresponding telescope have been used. From these two estimates, the weighted arithmetic mean has been calculated. $E_{\mathrm{rec,1+2}}$ denotes the stereo reconstructed energy, as used in the previous subsection. The vertical bars represent the uncertainty obtained with a cross-validation, while the horizontal bars indicate the energy range of the performance evaluation. The vertical bars are shifted horizontally for visibility purposes.

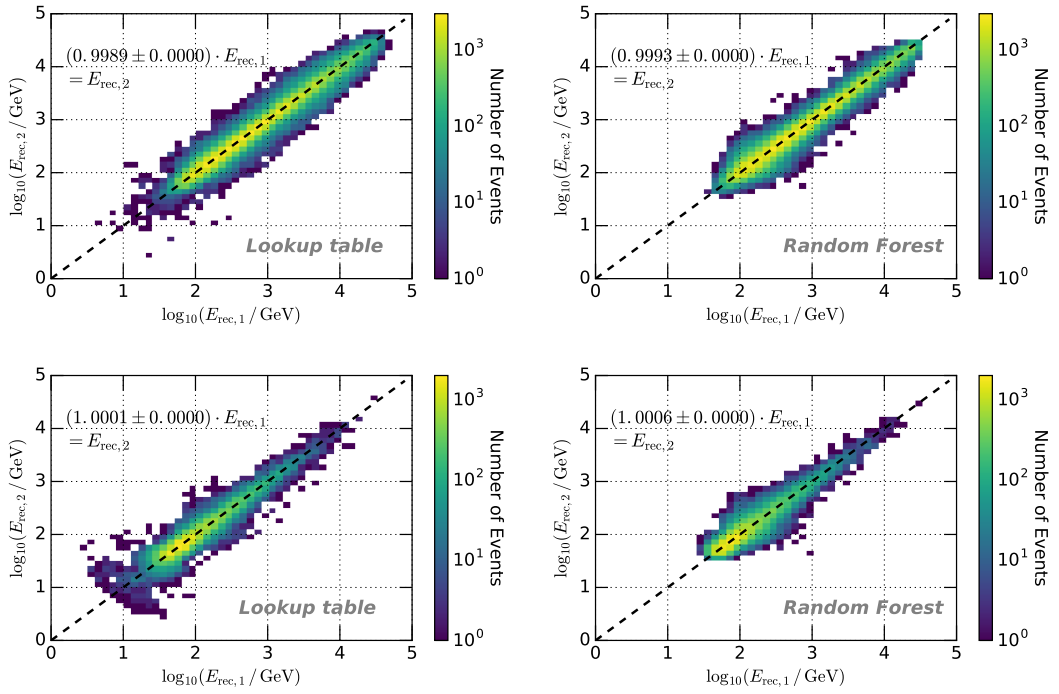

Figure 9.17: Consistency of telescope-wise reconstructed energies for Monte Carlo simulations (top) and excess events of Crab Nebula observations (bottom). `Hadronness` and θ^2 cuts according to Figure 9.9 have been applied. The dependency of the individual reconstructed energies is clearly visible for both simulations and real data.

Subsequent to the application of the Hadronness and θ^2 cuts according to Figure 9.9, 128 842 (lookup table) and 130 489 (Random Forest) of the simulated events, and 37 653 (lookup table) and 34 823 (Random Forest) of the excess events survive. For the determination of the dependencies between the individual energies, a linear fit is conducted. All distributions feature a direct dependency between the individual reconstructed energies, confirmed by the fit results, which are stated in the figures. The slope of the fit of the simulated events is below 1, the one of the excess events greater 1. The differences in the slope and the number of surviving events can be explained by a different underlying energy spectrum and a mismatch between Monte Carlo simulations and real data. Comparing the distributions of the lookup table method and the Random Forest regression, differences in the low-energy regime are noticeable. This is related to the performed Hadronness and θ^2 cuts. In contrast to the lookup table method, no cuts below $\approx$ 30 GeV could have been derived for the Random Forest regression, since not enough events have been reconstructed with smaller energies. This is, of course, also visible in the investigated two-dimensional distributions. Nevertheless, reconstructed energies below 10 GeV, as they occur when reconstructing the energy with the lookup table method, seem unrealistic.

9.5.7 Comparison of Different Evaluation Procedures

In this work, not only the performance of the energy reconstruction is of interest, but also the influence of the evaluation measure and the choice of cuts to derive the data sample on the performance is to be studied. The different evaluation measures are explained in more detail in subsection 9.5.1.
The different measures are compared for both the lookup table method and the Random Forest regression in Figure 9.18. The results of the fit with a Gaussian function and the median or the interquantile range, respectively, are rather similar for the bias and resolution, but also for both reconstruction methods. The mean as a measure of the bias deviates from the other measures, especially in the low-energy regime. Also the standard deviation shows a large discrepancy to the other measures in the complete energy range. For the lookup table method, the behavior of the standard deviation is conspicuous, and very different compared to the one of the Random Forest regression. It deviates extremely from the other methods, and the uncertainties are quite large. This is explained by the fact that the standard deviation is very sensitive to outliers, since it is dependent of the squared distance of an event to the mean. Thus, it can be concluded that the lookup table method produces several significant outliers.
The different cuts to derive the data sample, on which the performance is evaluated, are compared in Figure 9.19. The basic cuts are described in subsection 9.5.2, and the performance cuts are listed in subsection 9.5.1. The flute (Hadronness and θ^2) cuts have been derived dependent on the reconstructed energy, and are depicted in Figure 9.9. In general, the differences between the data samples have a larger impact on the performance of the lookup table method. The performance of the Random Forest regression is only slightly influenced by the choice of the cuts. Especially in the high-energy regime, the impact of the cuts on the performance of the lookup table method is significant. Since the performance cuts (or even more stringent cuts) are typically applied for statements regarding the telescope's performance, the performance values as e. g. stated in Aleksić et al., 2016b should be interpreted carefully. The application of flute cuts is more suited and corresponds to the same cuts energy spectra and light curves have been derived with, and thus, this evaluation presents a more realistic performance of the energy reconstruction.

In conclusion, the median and interquantile range as measures of the bias and the resolution have been the appropriate choice in the evaluation of the energy reconstruction. Both measures are robust and no assumptions regarding a specific distribution have been necessary.

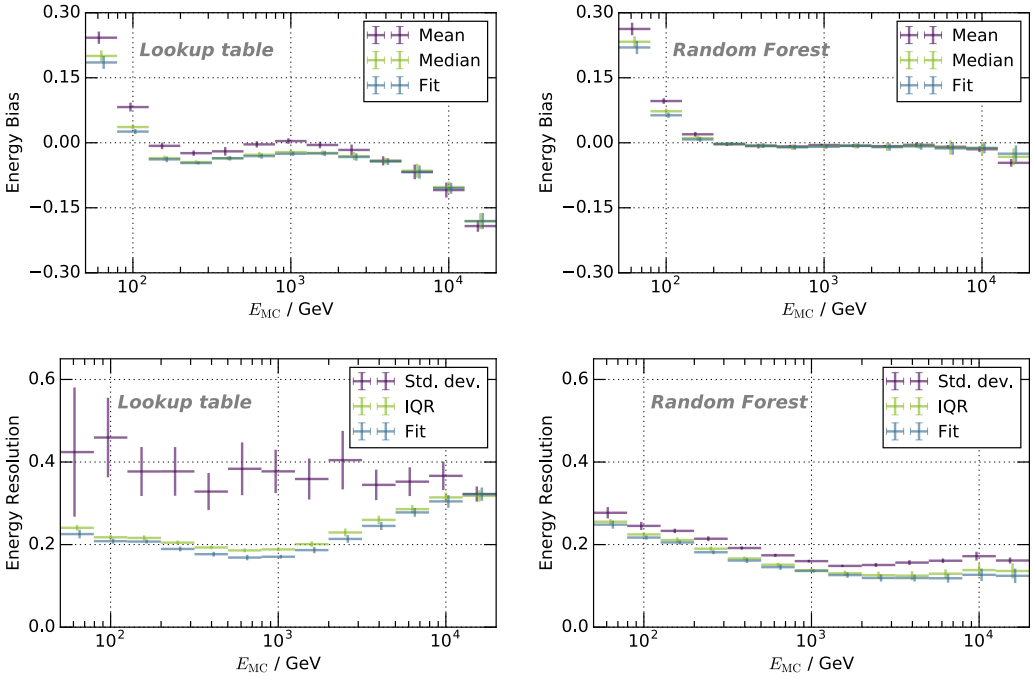

Figure 9.18: Comparison of different evaluation methods (see subsection 9.5.1 for a description of the methods), applied to the results obtained with the lookup table method and the Random Forest regression. The standard deviation performs rather different for both methods. Since the standard deviation is quite sensitive to outliers, it can be concluded that the lookup table method produces more significant outliers. `Hadronness` and θ^2 cuts according to Figure 9.9 have been applied. The vertical bars represent the uncertainty obtained with a bootstrap validation (lookup table method) or a cross-validation (Random Forest regression), while the horizontal bars indicate the energy range of the performance evaluation. The vertical bars are shifted horizontally for visibility purposes.

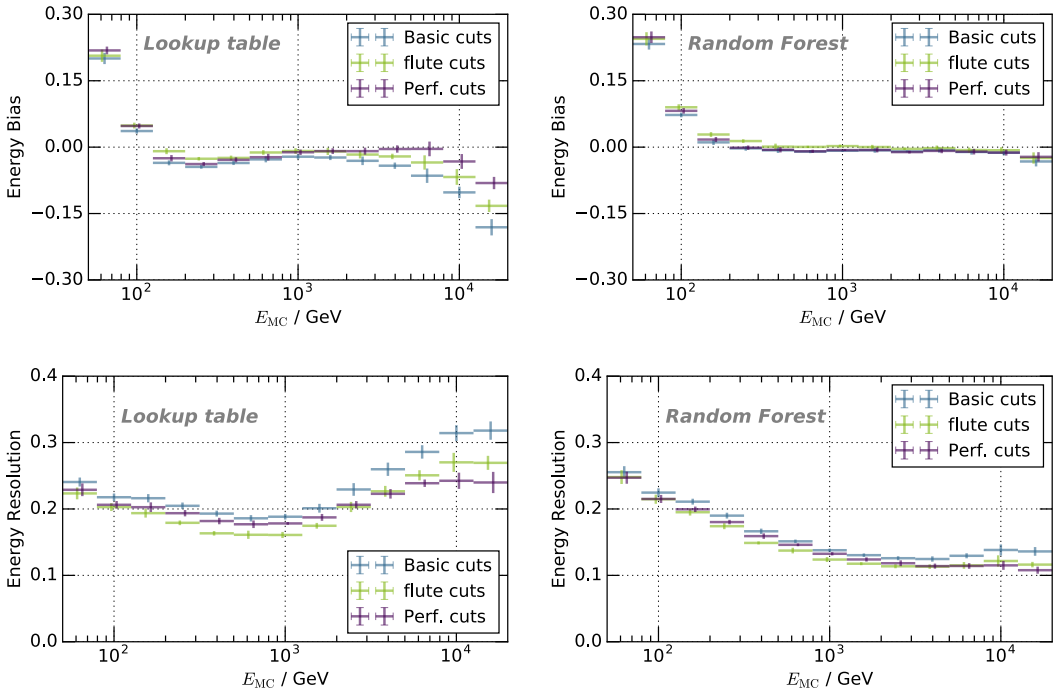

Figure 9.19: Comparison of different quality cuts, applied to the results obtained with the lookup table method and the Random Forest regression. The different methods obtain very similar and stable performance values for the evaluation of the Random Forest regression. In contrast, the evaluation of the lookup table method is quite sensitive to the applied cuts. The performance cuts are listed in subsection 9.5.1. The flute (`Hadronness` and θ^2) cuts have been derived dependent on the energy, and are depicted in Figure 9.9. The vertical bars represent the uncertainty obtained with a bootstrap validation (lookup table method) or a cross-validation (Random Forest regression), while the horizontal bars indicate the energy range of the performance evaluation. The vertical bars are shifted horizontally for visibility purposes.

9.5.8 Combination of the Regression with the Lookup Table Method

A possibility for further improvements is the combination of the Random Forest regression with results obtained with the lookup table method and the direction reconstruction. Subsequent to a successful direction reconstruction, the container, comprising features such as `CherenkovDensity`, `CherenkovRadius`, `MaxHeight` and `Valid`, is updated from `MStereoPar` to `MStereoParDisp`. These updated features, but also the energy E_{LUT} reconstructed with the lookup tables as an additional feature, could improve the Random Forest regression.
For an equitable comparison, the following basic cuts have been applied in addition to the `Hadronness` and θ^2 cuts according to Figure 9.9:

```
MHillas_{1,2}.fSize >= 50
MStereoPar.fValid > 0
MStereoPar.fCherenkovDensity > 0
MStereoParDisp.fValid > 0
MStereoParDisp.fCherenkovDensity > 0
```

Figure 9.20 compares the set with the updated features and the set with the updated features in combination with the energy E_{LUT} to the feature set which has been intensively studied and discussed in the previous subsections (see Figure 9.8 for a feature list). It turns out that the bias improves already for the updated feature set, and improves even more for the updated feature set in combination with the energy from the lookup tables, especially in the low-energy regime. The resolution is slightly better for the updated feature set in combination with E_{LUT} than the others.

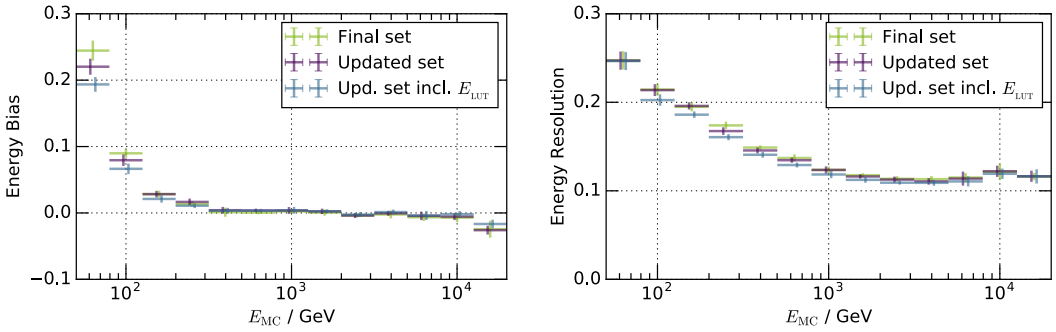

Figure 9.20: Comparison of the performance for different feature sets. `Hadronness` and θ^2 cuts have been applied according to Figure 9.9. In general, the set with the updated features (from the container `MStereoPar` to `MSteroParDisp`) performs better than the one discussed in the previous subsections (final set). Including the reconstructed energy E_{LUT} derived with the lookup tables as an additional feature to the updated ones, the performance improves even more. The vertical bars represent the uncertainty obtained within the cross-validation, while the horizontal bars indicate the energy range of the performance evaluation. The vertical bars are shifted horizontally for visibility purposes.

This set is also compared to the results of the default lookup table method (see Figure 9.21). The resolution is only better for the lookup tables in the lowest energy bin. At high energies, the Random Forest regression is by far better. With the new feature set, the bias improved in the low-energy regime, and is now rather similar to the one of the lookup tables. As before, for high energies the regression performs significantly better.

In *MARS*, the models for the direction and energy reconstruction and the g/h separation are generated independently, while the application of the models is conducted all at once. Thus, the update of the feature set based on the results of the direction reconstruction and the combination with the lookup table results cannot be integrated in a straight-forward manner in *MARS*. This is beyond the scope of this work and is left as an outlook of future improvements and an idea of what to expect from this specific enhancement.

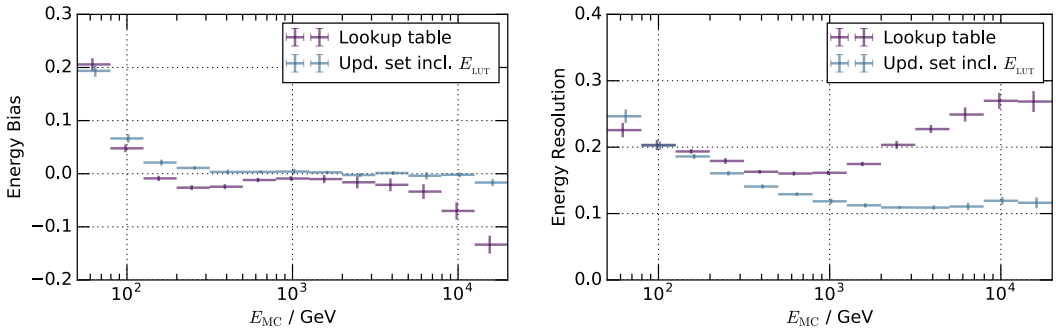

Figure 9.21: Comparison of the performance between the lookup table method and the Random Forest regression in combination with the lookup table method. `Hadronness` and θ^2 cuts have been applied according to Figure 9.9. Still, the lookup table performs slightly better in the low-energy regime, but in the high-energy regime the Random Forest is by far better. The vertical bars represent the uncertainty obtained with a bootstrap validation (lookup table method) or a cross-validation (Random Forest regression), while the horizontal bars indicate the energy range of the performance evaluation. The vertical bars are shifted horizontally for visibility purposes.

9.6 Optimization of the Gamma/Hadron Separation

The optimization of the gamma/hadron separation has been conducted on a sample of Monte Carlo simulations of gamma rays with a spectral index of -1.6, and a sample of recorded data. Observations with small gamma-ray contributions have been chosen, since the recorded data are to be used as hadronic events. The preparation of the samples is described in section 9.2.
The performance evaluation of the classification and the optimization are to be evaluated with *scikit-learn*. To transfer the findings subsequently to *MARS*, it has to be ensured that the Random Forest implementations in *scikit-learn* and *MARS* produce similar results. Therefore, they have been compared in a first step, before the classification has been optimized, and has been validated with observations of the Crab Nebula.

9.6.1 Comparison of *MARS* and *scikit-learn*

For the comparison between *MARS* and *scikit-learn*, the default pre-cuts, Random Forest settings and features, listed in section 9.1, have been utilized in both frameworks. The application of the pre-cuts have lead to approximately 200 000 gamma and hadron events each.
Figure 9.22 compares the distributions of the output of the Random Forest classification – the `Hadronness` – between *MARS* and *scikit-learn* for gamma and hadron events. The distributions are compatible for both event types. The shape of the distributions indicate a successful classification: Many hadron events feature large `Hadronness` values, while a large fraction of gamma events exhibits small values. Also the event-to-event comparison of the `Hadronness` values dependent on the energy indicates consistent results (see Figure 9.23). The red crosses mark the mean and the standard deviation of a particular energy bin. Correspondingly, it is valid to perform the optimization with *sklearn*, and to transfer the findings to *MARS*.

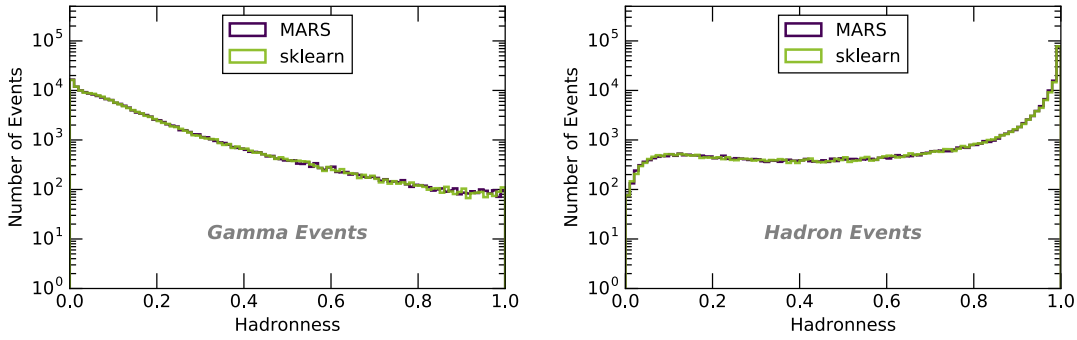

Figure 9.22: Comparison of the `Hadronness` distributions between *MARS* and *sklearn* for gamma and hadron events. For both event types, the distributions are very similar.

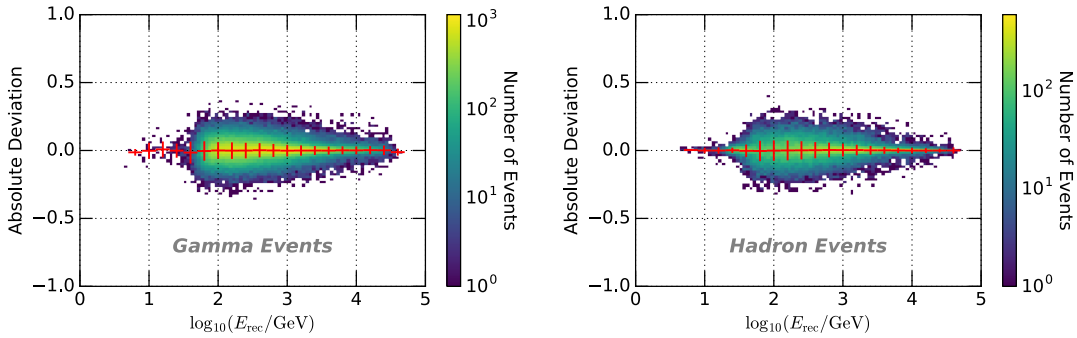

Figure 9.23: Deviation between the `Hadronness` values on event basis, obtained with *MARS* and *sklearn*. The red crosses mark the mean and the standard deviation of the particular energy bin, and indicate consistent results.

9.6.2 Feature Selection

The feature selection has been conducted analogously to the one in the previous section. First, some very basic cuts have been applied to the data samples (cf. subsection 9.5.2) before the equalization has been performed (cf. section 9.2). Moreover, a feature generation and a pre-selection has been conducted (cf. section 9.4). Subsequent to a removal of correlated features, a recursive backward elimination has been utilized for the reduction of features, since this procedure turned out to be performant in the previous section. For this purpose, the following initial settings have been chosen:

Number of trees 100
Number of features considered at each node $\lfloor \sqrt{\text{Number of available features}} \rfloor$
Minimum number of samples for split 5

The selected features and their importances are listed in Figure 9.24.

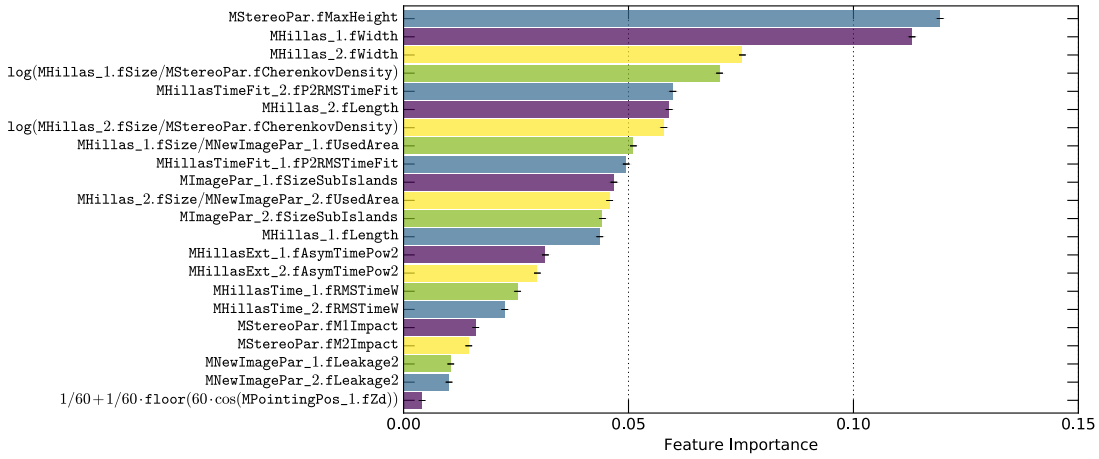

Figure 9.24: Feature importance of the selected features for the classification between gamma and hadron events.

9.6.3 Performance of the Gamma/Hadron Separation

The performance of the Random Forest classification, i. e. the gamma/hadron separation, has been evaluated with typical measures. For the definition of these performance measures, the reader is referred to section 3.3. In gamma-ray astronomy, it is common to name the false positive rate *hadron efficiency* and the true positive rate *gamma efficiency*. The Receiver Operating Characteristic (ROC) curve presents the gamma efficiency as a function of the hadron efficiency, as shown in Figure 9.25. The Area Under this Curve (AUC) indicates the performance of the classification – the closer to 1, the better. The optimized classification model has achieved an AUC score of 0.984, and the default model a score 0.981. Correspondingly, they perform rather similar, but the optimized model performs slightly better.

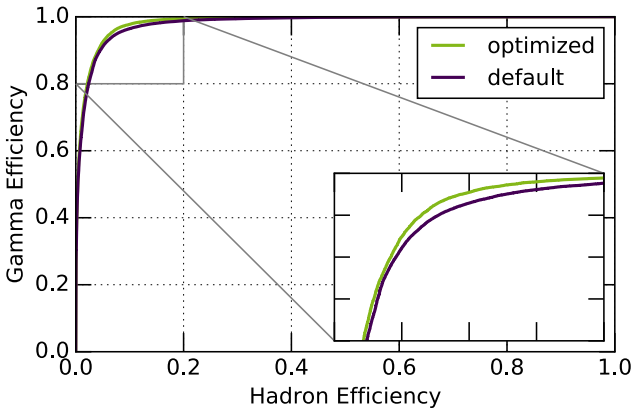

Figure 9.25: ROC curve for the default and the optimized Random Forest classification model. The optimized model performs slightly better, and achieves larger AUC scores.

Further performance measures are the purity and the efficiency. In contrast to the ROC curve and the AUC score, these values are dependent on a `Hadronness` cut. This cut has been chosen according to the gamma efficiency that is achieved. The performance measures purity and hadron efficiency are depicted in Figure 9.26 dependent on the reconstructed energy for various gamma efficiencies.

The purity decreases and the hadron efficiency increase, if the gamma efficiency is increased, i. e. the more gamma events are in the resulting sample, the less pure, and the more hadron events remain in the sample. For gamma efficiencies over 70%, purities over 0.8 and hadron efficiencies of at most 0.4 have been achieved. For energies above 100 GeV, the performance of the classification is close to perfect, i. e. the purity amounts to 1 and the hadron efficiency to 0, and the measures are quasi independent from the gamma efficiency. This is i. a. related to the shower images. The higher the energy, the more details of the shower are visible, and thus, the shower images are more diverse, and simple to classify.

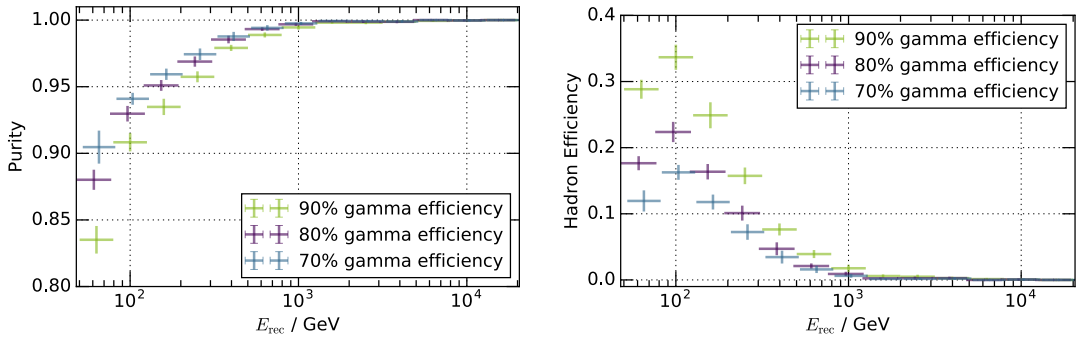

Figure 9.26: Performance of the optimized classification model as a function of the reconstructed energy. The performance has been measured in terms of the purity (left) and the hadron efficiency (right) for different specified gamma efficiencies. The vertical bars represent the uncertainty obtained with a cross validation. The vertical bars are shifted horizontally for visibility purposes.

It is important to note that the `Hadronness` of one model cannot be equated with the `Hadronness` of another model. The executable *odie* calculates the significance according to Li and Ma, 1983 dependent i. a. on the `Hadronness` cut. To determine the significances of both the default and the optimized classification model, it is required to adjust the `Hadronness` cut to allow a fair comparison. As the criterion for this adjustment, the purity has been chosen, since the same purity corresponds to the same introduced systematic uncertainty. In Figure 9.27, the significance as a function of the `Hadronness` cut is depicted for both classification models. The distribution of the default model is much steeper for small cuts than the optimized model. This indicates that the optimized model produces more robust results, since the significance does not change substantially, if the cut is slightly varied.

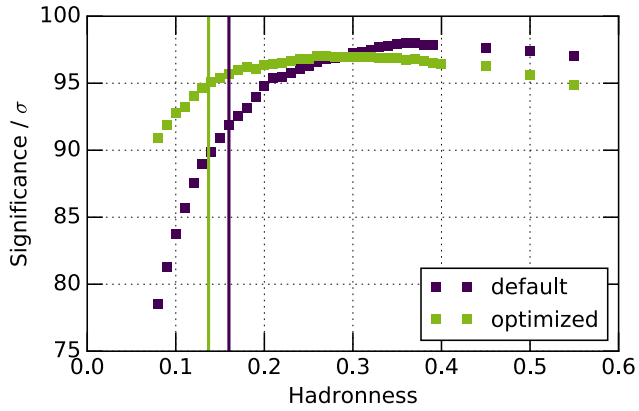

Figure 9.27: Comparison of the significances dependent on the `Hadronness` cut between the default and the optimized classification model. The vertical lines indicate the corresponding `Hadronness` cuts of 0.137 and 0.16, respectively. The optimized model is more robust, and achieves a higher significance for the adjusted cut value.

The resulting significances and event numbers obtained with *odie* are summarized in Table 9.2 for different cuts (see subsection 8.1.2). For the *LE* cuts, the `Hadronness` cut has been adjusted to 0.249, for the *FR* cuts to 0.137 and for the *HE* cuts to 0.077 for the optimized model. For all regimes, an improvement has been achieved with the optimized model.

Table 9.2: Comparison of the results derived with the executable *odie* for different cuts (see subsection 8.1.2 for details). The significance according to Li and Ma, 1983 and the sensitivity have been calculated with the number of events N_{on} in the signal (on) region and N_{off} in the background (off) region for 11.4 h of observation of the Crab Nebula. The optimized classification model achieves better significances for all cuts.

	N_{on}	N_{off}	N_{ex}	Signif. / σ
LE cuts				
Default	14 409	2889.0 ±31.0	11 520.0 ±124.0	119.96
Optimized	13 891	2574.0 ±29.3	11 317.0 ±121.4	121.58
FR cuts				
Default	3717	93.3 ±5.6	3623.7 ±61.2	91.86
Optimized	4211	147.0 ±7.0	4064.0 ±65.3	94.93
HE cuts				
Default	696	5.3 ±1.3	690.7 ±26.4	42.26
Optimized	757	7.7 ±1.6	749.3 ±27.6	43.64

9.6.4 Verification with Crab Nebula Observations

An evaluation and validation of the default and the optimized Random Forest classification models has been performed by determining the spectral energy distributions of the Crab Nebula – the standard candle in astroparticle physics. Figure 9.28 shows the spectral energy distributions in terms of the reconstructed energy, derived with *flute*. The distributions from the default and the optimized model are very similar and compatible with former results for observations of the Crab Nebula obtained with MAGIC (Aleksić et al., 2015, J. Albert et al., 2008b). However, the optimized model produces more robust results in the high-energy regime. To estimate the distributions of the true energy, an unfolding with the macro *CombUnfold* has been conducted. For both models, the same unfolding settings have been used: The strength of the Tikhonov regularization has been chosen such that the noise of the covariance matrices of E_{rec} and E_{MC} is the same. For the choice of the bins of E_{rec} and E_{MC}, i. a. the event distribution of E_{rec} and the collection area has been considered. The unfolded spectral energy distributions are depicted in Figure 9.29. The distributions are very similar, and more stable, even in the high-energy regime.

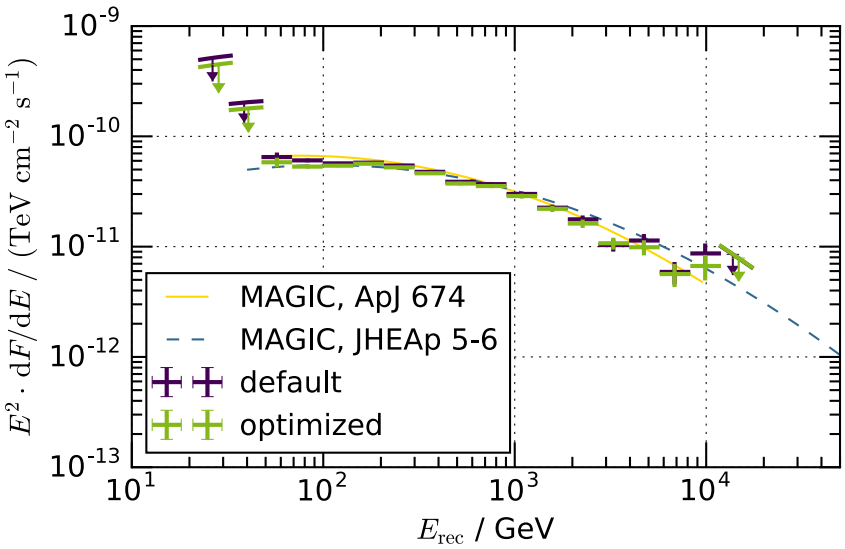

Figure 9.28: Spectral energy distribution obtained with *flute* for the gamma/hadron separation with the default and the optimized model. Both spectra are compatible with former results (Aleksić et al., 2015, J. Albert et al., 2008b). In the high-energy regime, the optimized model produces more robust results.

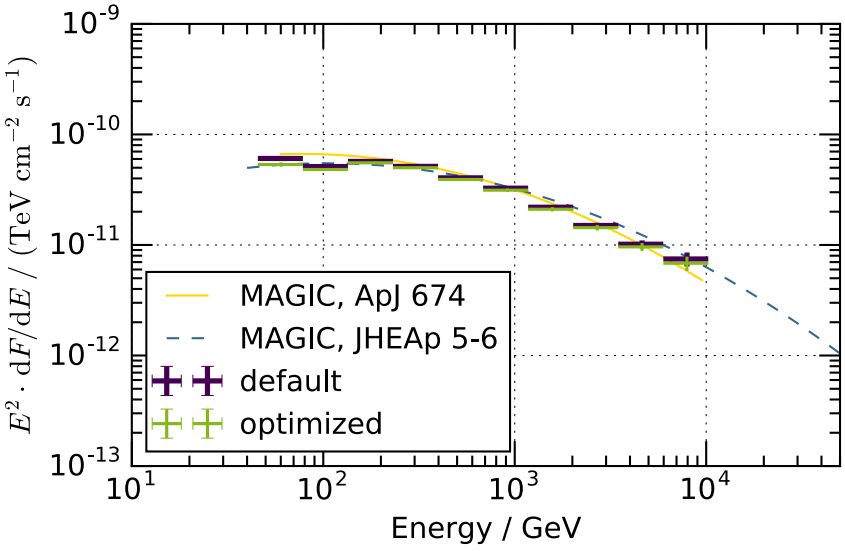

Figure 9.29: Spectral energy distribution, unfolded with *CombUnfold*. The bins in E_{rec} and E_{MC} have been chosen separately. The results are very similar and compatible with former measurements.

9.7 Conclusion and Discussion

In the scope of this chapter, the energy reconstruction and the gamma/hadron separation have been revised. For stereoscopic observations, the energy reconstruction is by default performed using a lookup table for each telescope separately, and the results thereof are averaged to derive the final reconstructed energy. The drawbacks of using only a small fraction of available information, and not fully exploiting the benefits of stereoscopic observations have been eliminated by the application of a Random Forest regression on full stereoscopic information. For this regression, a set of appropriate features has been determined in an adapted backward elimination procedure. By comparing different feature sets, the impact on the performance has been demonstrated and proved the importance of the choice of features in the optimization procedure. Both the feature generation and the sophisticated reduction of the number of features turned out to be crucial. The importance of the feature generation might not be apparent, since basically no new information is added, but specific multivariate combinations of features can support the Random Forest. While multiple relevant features have been generated, the set still comprised irrelevant features. These features unnecessarily inflate the feature space, which is often problematic. A deflation of this space improves the description of the sample's density and reduces the dimensionality.

A verification of the Random Forest regression has been conducted by analyzing observations of the Crab Nebula. In order to perform the regression with the found settings in *MARS*, some modifications in the executables *coach* and *melibea* have been implemented. Comparing the spectral energy distributions derived with the lookup table method and the Random Forest regression, it turns out that both energy reconstructions produce similar distributions, compatible with former results of the spectral energy distribution of the Crab Nebula. However, the Random Forest regression produces more stable results in the high-energy regime, and reconstructs events up to higher energies. This is i. a. related to different pre-cuts prior to the generation of the reconstruction model. In case of the lookup tables, events with large total charges and events partly outside the camera are not included in the model. These events typically exhibit high energies. But also in the low-energy regime, the lookup tables are problematic, since they do not produce the same distribution for recorded data and simulations, in contrast to the Random Forest regression.

Within the MAGIC experiment, it is common to evaluate the performance of the energy reconstruction in terms of the bias and the resolution, determined with a Gaussian fit and several high-quality cuts applied. However, the distribution of the relative differences between the true and the reconstructed energy is slightly asymmetric and skewed. Correspondingly, a Gaussian fit falsifies the performance, and it has been shown that the calculation of the median and an interquantile range is better suited and much more robust. Moreover, the applied high-quality cuts are much more strict than those applied to default analyses. Thus, less strict cuts have been applied here, and more representative performance measures have been calculated.

The standard gamma/hadron separation in *MARS* does not provide performance measures, estimating e. g. the robustness of the built classification model, or the hadron contamination of the resulting sample. To compare the default classification model to the optimized one, the AUC score has been used as performance measure, indicating compatible and robust results. The verification of both models with observations of the Crab Nebula lead to compatible spectral energy distributions and significances as well, although the significance was slightly higher for the optimized model. However, considering the significance as a function of the `Hadronness` cut for both models, it turns out that the optimized model is more robust, since the distribution of the default model is steeper, i. e. the significance changes substantially, if the cut is slightly varied.

10 Concluding Remarks and Future Prospects

A tremendous feeling of peace came over him.
He knew that at last, for once and for ever, it was now all, finally, over.

— Douglas Adams, *The Hitchhiker's Guide to the Galaxy*

This thesis is concerned with the examination of Active Galactic Nuclei (AGN) within the field of gamma-ray and multi-wavelength astronomy in conjunction with data mining applications. The extension of the population of AGNs enables a tremendous refinement of our knowledge of gamma-ray emitting objects. Thereby, it is possible to probe for instance the acceleration and emission processes of AGNs and their subtypes, or the extragalactic background light. For this purpose, the information about the class affiliation of a cataloged source, the precise source localization and its corresponding counterparts is important.

To possibly extend the AGN population, the multi-wavelength data of a list of AGN candidates from the *Fermi* catalog have been investigated to increase the quality of the prediction. The capability and performance of this method has been proven in the scope of this thesis. However, this procedure is quite complex and inefficient, and does not fully exploit modern data mining techniques.

For that reason, a new approach to search for AGN candidates in the *Fermi* catalog has been developed, and the suitability of the concept has been proven. This method exploits multi-wavelength information in conjunction with data mining methods to determine both the class affiliation and the corresponding counterpart. The examination of the data mining model provided also information concerning the physical properties and behaviors of the source.

Looking at future prospects, the findings and developed methods are perfectly suited to also search for potential dark matter candidates. The most promising candidates are Weakly Interacting Massive Particles (WIMPs), whose expected annihilation could produce gamma rays. Gamma-ray emitting objects featuring a dark matter origin are thought to emit a constant flux in the high-energy regime, but should not exhibit counterparts at lower energies. Moreover, it is assumed that their spectra are hard and feature a cutoff related to the WIMP mass. To search for such objects, a multi-step approach could be followed: A classification model discriminating between AGNs and pulsars (the two main classes in the *Fermi*-LAT catalog) is generated based on the *Fermi*-LAT catalog. The score of an unassociated gamma-ray source with dark matter origin should indicate neither an AGN nor a pulsar. Further classification models, searching for AGN candidates and their counterparts as the above-described newly developed method, are created for multiple wavelengths. All scores of a particular gamma-ray source should indicate no compatible counterparts for a source to be considered as dark matter candidate. Certainly, if candidates result from this approach, it is not the proof of dark matter, but it provides a candidate list of interesting sources to be examined further. For instance, a comparison of the spectra of the candidates is instructive for dark matter, since all gamma-ray emitting objects of dark matter origin should exhibit a common spectral slope.

The classification task of the previous search for AGNs has been refined to search for specific blazars, a subtype of AGNs, and their counterparts. Despite the higher complexity, the capability of the method has been successfully demonstrated.
In recent years, topics like data mining and big data increasingly became the focus of general interest, resulting in the rapidly growing development of modern techniques. This progress also gained ground in astronomy and astrophysics, offering a variety of new possibilities. Many data centers collect and distribute astronomical data, such as catalogs and tables published in journals. Moreover, they provide innovative software tools to exploit large scientific data and processing capabilities. The Virtual Observatory (VO) is a fusion of such data centers to a worldwide network. Recent approaches of the VO involve the use of hierarchical data methods and fast catalog-matching. The newly developed methods to search for AGN and blazar candidates and their corresponding counterparts are perfectly suited for an integration into the VO. The VO allows an easy access to multiple catalogs from various instruments at several wavelengths, and moreover, it provides tools and computing power for the data preparation and generation of the models. On the other hand, the VO profits from the catalog cross-matching, and the fast class and counterpart affiliation the developed procedures provide. The future prospects of such a collaboration are immense.
A very promising source of the derived blazar candidates has been observed with the MAGIC telescopes in the very-high-energy regime. The analysis of these observations resulted in the derivation of flux upper limits, which have been set in context with observations at other wavelengths, allowing conclusions on the properties of the object.
The MAGIC analysis chain offers some possibilities to utilize data mining techniques as well, for instance to reconstruct the energy or to discriminate between gamma rays and hadrons. So far, the promising approach to reconstruct the energy of stereoscopic observations with data mining techniques has not been exploited, and thus, it has been newly set up. Both stated applicabilities have been revised, optimized and evaluated, and their benefits have been illustrated. Especially, the choice of the feature set to generate the models turned out to be extremely important.
Therefore, the development of new features is a very relevant and promising topic for future optimizations. For this purpose, the feature generation should not start only at the level of cleaned shower images, but at the level of individual pixels or even at the level of the pixel's time series. One possibility is to calculate the standard image parameters for different thresholds of the image cleaning (without changing the criterion if an image survives the cleaning). Like this, faint parts of the shower can be included without adding events, which are probably due to noise. This new information is especially interesting for the discrimination between gamma rays and hadrons, since the outer parts of the shower are highly discriminative.

A Multi-Wavelength Analysis

A.1 Coordinates of Associated Sources

Table A.1: Coordinates of associated 3FGL sources.

3FGL Name	Associated Name	RA	Dec
3FGL J0001.2-0748	PMN J0001-0746	0.32512	-7.77417
3FGL J0001.4+2120	TXS 2358+209	0.38488	21.22674
3FGL J0002.2-4152	1RXS J000135.5-415519	0.39723	-41.91437
3FGL J0003.2-5246	RBS 0006	0.83158	-52.79081
3FGL J0003.8-1151	PMN J0004-1148	1.02048	-11.81622
3FGL J0004.7-4740	PKS 0002-478	1.14856	-47.60545
3FGL J0006.4+3825	S4 0003+38	1.48823	38.33754
3FGL J0008.0+4713	MG4 J000800+4712	1.99988	47.20217
3FGL J0008.6-2340	RBS 0016	2.1475	-23.65775
3FGL J0009.1+0630	CRATES J000903.95+062821.5	2.26646	6.47264
3FGL J0009.3+5030	NVSS J000922+503028	2.34388	50.50803
3FGL J0009.6-3211	IC 1531	2.39824	-32.27701
3FGL J0012.4+7040	TXS 0008+704	2.88293	70.75878
3FGL J0013.2-3954	PKS 0010-401	3.24962	-39.90724
3FGL J0013.9-1853	RBS 0030	3.48354	-18.9019
3FGL J0014.0-5025	RBS 0032	3.54675	-50.37572
3FGL J0014.6+6119	4C +60.01	3.7033	61.29543
3FGL J0014.7+5802	1RXS J001442.2+580201	3.67546	58.0337
3FGL J0015.7+5552	GB6 J0015+5551	3.91729	55.86242
3FGL J0016.3-0013	S3 0013-00	4.0462	-0.25346
3FGL J0017.2-0643	PMN J0017-0650	4.28908	-6.84264
3FGL J0017.6-0512	PMN J0017-0512	4.39924	-5.2116
3FGL J0018.4+2947	RBS 0042	4.61562	29.79178
3FGL J0018.9-8152	PMN J0019-8152	4.83596	-81.881
3FGL J0019.1-5645	PMN J0019-5641	4.86071	-56.69525
3FGL J0019.4+2021	PKS 0017+200	4.90773	20.36268
3FGL J0021.6-6835	PKS 0021-686	6.02792	-68.3485
3FGL J0021.6-2553	CRATES J002132.55-255049.3	5.38563	-25.84703
3FGL J0022.1-1855	1RXS J002209.2-185333	5.53862	-18.89303
3FGL J0022.1-5141	1RXS J002159.2-514028	5.50025	-51.67344
3FGL J0022.5+0608	PKS 0019+058	5.63517	6.13452
3FGL J0023.5+4454	B3 0020+446	5.89768	44.94327
3FGL J0024.4+0350	GB6 J0024+0349	6.188333	3.817639
3FGL J0026.7-4603	1RXS J002636.3-460101	6.64879	-46.01922
3FGL J0028.6+7507	GB6 J0028+7506	7.05479	75.10372
3FGL J0028.8+1951	TXS 0025+197	7.12424	20.00743
3FGL J0029.1-7045	PKS 0026-710	7.17312	-70.75447
3FGL J0030.2-1646	1RXS J003019.6-164723	7.58167	-16.78972
3FGL J0030.3-4223	PKS 0027-426	7.57327	-42.41278
3FGL J0030.7-0209	PKS B0027-024	7.6326	-2.19893
3FGL J0031.3+0724	NVSS J003119+072456	7.83217	7.41578
3FGL J0032.3-2852	PMN J0032-2849	8.13792	-28.82231
3FGL J0033.6-1921	KUV 00311-1938	8.3925	-19.35925
3FGL J0035.2+1513	RX J0035.2+1515	8.81135	15.25115
3FGL J0035.9+5949	1ES 0033+595	8.96935	59.83461
3FGL J0037.9+1239	NVSS J003750+123818	9.46187	12.63856
3FGL J0038.0-2501	PKS 0035-252	9.5614	-24.98395
3FGL J0038.0+0012	NVSS J003808+001336	9.53543	0.22682
3FGL J0039.0-2218	PMN J0039-2220	9.78421	-22.33372
3FGL J0039.1-0939	TXS 0036-099	9.77622	-9.71302
3FGL J0039.1+4330	NVSS J003907+433015	9.78296	43.50431
3FGL J0040.3+4049	B3 0037+405	10.0575	40.83464
3FGL J0040.5-2339	PMN J0040-2340	10.10379	-23.66689
3FGL J0041.9+3639	RX J0042.0+3641	10.535	36.6875
3FGL J0042.0+2318	PKS 0039+230	10.51894	23.33363
3FGL J0043.5-0444	1RXS J004333.7-044257	10.89217	-4.71681
3FGL J0043.7-1117	1RXS J004349.3-111612	10.95288	-11.26867
3FGL J0043.8+3425	GB6 J0043+3426	10.95353	34.44059

Table A.1: – continued from previous page

3FGL Name	Associated Name	RA	Dec
3FGL J0045.2-3704	PKS 0042-373	11.30025	-37.09667
3FGL J0045.3+2126	GB6 J0045+2127	11.33042	21.46114
3FGL J0045.7+1217	GB6 J0045+1217	11.43059	12.28661
3FGL J0046.7-8419	PKS 0044-84	11.11192	-84.37781
3FGL J0047.0+5658	GB6 J0047+5657	11.75179	56.96178
3FGL J0047.9+5447	1RXS J004754.5+544758	11.96608	54.79581
3FGL J0048.0+3950	B3 0045+395	11.98008	39.816
3FGL J0048.0+2236	NVSS J004802+223525	12.01092	22.59006
3FGL J0049.4-5401	PMN J0049-5402	12.45354	-54.04536
3FGL J0049.4-4149	1RXS J004939.9-415133	12.41625	-41.85917
3FGL J0049.7+0237	PKS 0047+023	12.43015	2.61772
3FGL J0049.8-5737	PKS 0047-579	12.4978	-57.64093
3FGL J0050.0-4458	PMN J0049-4457	12.31933	-44.95319
3FGL J0050.4-0449	PKS 0047-051	12.58973	-4.87239
3FGL J0050.6-0929	PKS 0048-09	12.67216	-9.48478
3FGL J0051.0-0649	PKS 0048-071	12.78421	-6.83395
3FGL J0051.2-6241	1RXS J005117.7-624154	12.81942	-62.70117
3FGL J0054.8-2455	FRBA J0054-2455	13.69471	-24.92486
3FGL J0055.2-1213	TXS 0052-125	13.79909	-12.29919
3FGL J0056.3-2116	PMN J0056-2117	14.1345	-21.28556
3FGL J0056.3-0935	TXS 0053-098	14.08367	-9.60831
3FGL J0057.9-0542	PKS 0055-059	14.52111	-5.66452
3FGL J0058.0-3233	PKS 0055-328	14.50929	-32.57243
3FGL J0058.3+3315	MG3 J005830+3311	14.63362	33.18812
3FGL J0059.1-5701	PKS 0056-572	14.69409	-56.98652
3FGL J0059.2-0152	1RXS J005916.3-015030	14.82054	-1.83822
3FGL J0059.6+0003	PKS 0056-00	14.77298	0.11434
3FGL J0100.2+0745	GB6 J0100+0745	15.08662	7.76428
3FGL J0102.3+4217	GB6 J0102+4214	15.61313	42.23861
3FGL J0102.8+5825	TXS 0059+581	15.69068	58.40309
3FGL J0103.4+5336	1RXS J010325.9+533721	15.85817	53.62036
3FGL J0103.7+1323	NVSS J010345+132346	15.94079	13.39614
3FGL J0105.1-2415	PKS 0102-245	16.24252	-24.27457
3FGL J0105.3+3928	GB6 J0105+3928	16.28833	39.47092
3FGL J0107.0-1208	PMN J0107-1211	16.79913	-12.18989
3FGL J0108.5-0035	PKS 0105-008	17.11184	-0.62338
3FGL J0108.7+0134	4C +01.02	17.16155	1.58342
3FGL J0109.1+1816	MG1 J010908+1816	17.28408	18.26875
3FGL J0109.8+6132	TXS 0106+612	17.4431	61.55846
3FGL J0109.9-4020	RBS 0158	17.48575	-40.34753
3FGL J0110.2+6806	4C +67.04	17.55364	68.09478
3FGL J0110.9-1254	1RXS J011050.0-125455	17.70838	-12.91769
3FGL J0111.5+0535	1RXS J011130.5+053612	17.87579	5.6075
3FGL J0112.1+2245	S2 0109+22	18.02427	22.74411
3FGL J0112.8+3207	4C +31.03	18.20972	32.13818
3FGL J0113.0-3554	PMN J0113-3551	18.31604	-35.86331
3FGL J0113.4+4948	S4 0110+49	18.36253	49.80668
3FGL J0114.8+1326	GB6 J0114+1325	18.71991	13.42708
3FGL J0115.7+0356	PMN J0115+0356	18.9188	3.9454
3FGL J0115.8+2519	RX J0115.7+2519	18.942375	25.3315
3FGL J0116.0-1134	PKS 0113-118	19.05217	-11.60429
3FGL J0116.2-2744	1RXS J011555.6-274428	18.98117	-27.74219
3FGL J0116.3-6153	SUMSS J011619-615343	19.08117	-61.8955
3FGL J0117.8-2113	PKS 0115-214	19.45325	-21.18518
3FGL J0118.8-2142	PKS 0116-219	19.73859	-21.69171
3FGL J0118.9-1457	1RXS J011905.4-145906	19.76925	-14.98292
3FGL J0120.4-2700	PKS 0118-272	20.13193	-27.02351
3FGL J0121.7+5154	NVSS J012133+515557	20.39025	51.93261
3FGL J0122.8+3423	1ES 0120+340	20.78599	34.34685
3FGL J0123.7-2312	1RXS J012338.2-231100	20.90996	-23.18292
3FGL J0125.2-0627	PMN J0124-0624	21.21033	-6.41719
3FGL J0125.4-2548	PKS 0122-260	21.32849	-25.81789
3FGL J0126.1-2227	PKS 0123-226	21.56251	-22.376
3FGL J0127.1-0818	PMN J0127-0821	21.81796	-8.35806
3FGL J0127.2+0325	NVSS J012713+032259	21.80805	3.38353
3FGL J0127.9+2551	4C +25.05	21.6783	25.98369
3FGL J0128.5+4430	GB6 J0128+4439	22.17224	44.655
3FGL J0130.8+1441	4C +14.06	22.48061	14.77995
3FGL J0131.2+6120	1RXS J013106.4+612035	22.78028	61.34267
3FGL J0131.3+5548	TXS 0128+554	22.80758	55.75364
3FGL J0132.5-0802	PKS 0130-083	23.17136	-8.06801
3FGL J0132.6-1655	PKS 0130-17	23.1812	-16.91348
3FGL J0133.0-4413	SUMSS J013306-441422	23.27679	-44.23958
3FGL J0133.2-5159	PKS 0131-522	23.27401	-52.0011
3FGL J0133.3+4324	B3 0129+431	23.18386	43.42574
3FGL J0134.3-3842	PMN J0134-3843	23.63346	-38.72594
3FGL J0134.5+2638	1RXS J013427.2+263846	23.61792	26.64583
3FGL J0135.0+6927	TXS 0130+691	23.66984	69.41969
3FGL J0136.5+3905	B3 0133+388	24.13542	39.09989
3FGL J0137.0+4752	OC 457	24.24414	47.85808
3FGL J0137.6-2430	PKS 0135-247	24.409853	-24.514971
3FGL J0137.8+5813	1RXS J013748.0+581422	24.46025	58.23644
3FGL J0139.9+8735	NVSS J013913+873754	24.80571	87.63186
3FGL J0141.4-0929	PKS 0139-09	25.35763	-9.4788
3FGL J0143.7-5845	SUMSS J014347-584550	25.94842	-58.76394

Table A.1: – continued from previous page

3FGL Name	Associated Name	RA	Dec
3FGL J0144.6+2705	TXS 0141+268	26.13981	27.0842
3FGL J0145.1-2732	PKS 0142-278	26.26414	-27.55954
3FGL J0145.6+8600	NVSS J014929+860114	27.37429	86.02056
3FGL J0146.4-6746	SUMSS J014554-674646	26.47837	-67.78031
3FGL J0147.0-5204	PKS 0144-522	26.70242	-52.04264
3FGL J0148.3+5200	GB6 J0148+5202	27.08438	52.03494
3FGL J0148.6+0128	PMN J0148+0129	27.14081	1.48373
3FGL J0150.5-5447	PMN J0150-5450	27.68467	-54.83469
3FGL J0151.0-3609	PMN J0151-3605	27.86437	-36.10486
3FGL J0151.0+0537	PMN J0151+0540	27.75768	5.67611
3FGL J0151.6+2205	PKS 0149+21	28.07525	22.11881
3FGL J0152.2+3707	B2 0149+37	28.05092	37.26824
3FGL J0152.6+0148	PMN J0152+0146	28.16505	1.78816
3FGL J0152.8+7517	1RXS J015308.4+751756	28.2805	75.29525
3FGL J0153.4+7114	TXS 0149+710	28.35771	71.2518
3FGL J0154.0+0824	GB6 J0154+0823	28.51154	8.39752
3FGL J0154.9+4433	GB6 J0154+4433	28.72695	44.56054
3FGL J0156.3+3913	MG4 J015630+3913	29.13088	39.24192
3FGL J0156.9-4742	2MASS J01564603-4744174	29.19179	-47.73817
3FGL J0157.0-5301	1RXS J015658.6-530208	29.24179	-53.03328
3FGL J0157.9-4615	PMN J0157-4614	29.46304	-46.23978
3FGL J0158.6-3931	PMN J0158-3932	29.65883	-39.53431
3FGL J0159.4+1046	RX J0159.5+1047	29.8933	10.78493
3FGL J0159.8-2741	PMN J0159-2739	29.93062	-27.67739
3FGL J0200.9-6635	PMN J0201-6638	30.28229	-66.63683
3FGL J0202.3+0851	TXS 0159+085	30.60996	8.82047
3FGL J0202.5+4206	B3 0159+418	30.68189	42.08787
3FGL J0203.1-0227	RX J0202.9-0223	30.71796	-2.38922
3FGL J0203.6+3043	NVSS J020344+304238	30.93413	30.71067
3FGL J0204.0+7234	S5 0159+723	30.8891	72.54824
3FGL J0204.2+2420	B2 0201+24	31.08979	24.29742
3FGL J0204.8+3212	B2 0202+31	31.27052	32.20836
3FGL J0205.0+1510	4C +15.05	31.21006	15.2364
3FGL J0205.2-1700	PKS 0202-17	31.24031	-17.02218
3FGL J0206.4-1150	PMN J0206-1150	31.60869	-11.84437
3FGL J0207.9-3846	PKS 0205-391	31.81504	-38.95085
3FGL J0208.0-6838	PKS 0206-688	31.96198	-68.6321
3FGL J0208.6+3522	MS 0205.7+3509	32.159	35.38689
3FGL J0209.4-5229	1RXS J020922.2-522920	32.34004	-52.48967
3FGL J0209.5+4449	1RXS J020917.6+444951	32.32138	44.82956
3FGL J0210.7-5101	PKS 0208-512	32.6925	-51.01719
3FGL J0211.2+1051	MG1 J021114+1051	32.80491	10.85967
3FGL J0211.2-0649	NVSS J021116-064422	32.82071	-6.73886
3FGL J0211.7+5402	TXS 0207+538	32.73485	54.08675
3FGL J0212.8-3504	RBS 0292	33.127	-35.05833
3FGL J0213.0+2245	MG3 J021252+2246	33.22015	22.74785
3FGL J0213.1-2720	PMN J0212-2719	33.2305	-27.305
3FGL J0214.4+5143	TXS 0210+515	33.57473	51.74776
3FGL J0214.7-5823	PMN J0214-5822	33.54304	-58.36858
3FGL J0216.1-7016	PMN J0215-7014	33.91404	-70.25003
3FGL J0216.6-1019	PMN J0216-1017	34.16198	-10.28417
3FGL J0217.0-6635	RBS 0300	34.21187	-66.61169
3FGL J0217.1-0833	PKS 0214-085	34.26109	-8.34788
3FGL J0217.2+0837	ZS 0214+083	34.32135	8.61775
3FGL J0217.3+6209	TXS 0213+619	34.26046	62.19278
3FGL J0217.5+7349	S5 0212+73	34.37839	73.82573
3FGL J0217.8+0143	PKS 0215+015	34.45398	1.74714
3FGL J0218.9+3642	MG3 J021846+3641	34.70849	36.67849
3FGL J0219.0+2440	87GB 021610.9+243205	34.7517	24.75571
3FGL J0221.1+3556	B0218+357	35.27279	35.93715
3FGL J0222.1-1616	PKS 0219-164	35.50302	-16.2546
3FGL J0222.6+4301	3C 66A	35.66505	43.0355
3FGL J0222.9-1117	1RXS J022314.6-111741	35.80946	-11.29397
3FGL J0223.3+6820	NVSS J022304+682154	35.76771	68.36522
3FGL J0223.5+6313	TXS 0219+628	35.87336	63.12147
3FGL J0224.1-7941	PMN J0223-7940	35.91254	-79.67058
3FGL J0225.2-2602	PMN J0225-2603	36.29508	-26.0555
3FGL J0226.3+0941	PMN J0226+0937	36.55717	9.62397
3FGL J0226.5-4442	RBS 0318	36.66204	-44.68956
3FGL J0227.2+0201	RX J0227.2+0201	36.81908	2.03347
3FGL J0228.0+2248	NVSS J022744+224834	36.93484	22.8095
3FGL J0228.3-5545	PKS 0226-559	37.09004	-55.76764
3FGL J0228.5+6703	GB6 J0229+6706	37.34433	67.10961
3FGL J0228.7-3106	PMN J0228-3102	37.05412	-31.04458
3FGL J0229.3-3643	PKS 0227-369	37.36854	-36.73245
3FGL J0230.6-5757	PKS 0229-581	37.78867	-57.91836
3FGL J0230.8+4032	B3 0227+403	37.69046	40.54807
3FGL J0232.8+2016	1ES 0229+200	38.20257	20.28818
3FGL J0232.9+2606	B2 0230+25	38.23485	26.16193
3FGL J0236.7-6136	PKS 0235-618	39.22186	-61.60422
3FGL J0237.5-3603	RBS 0334	39.39183	-36.05786
3FGL J0237.9+2848	4C +28.07	39.468357	28.802497
3FGL J0238.3-3904	1RXS J023800.5-390505	39.50267	-39.08469
3FGL J0238.4-3117	1RXS J023832.6-311658	39.63529	-31.28278
3FGL J0238.6+1636	AO 0235+164	39.66221	16.61647

Table A.1: – continued from previous page

3FGL Name	Associated Name	RA	Dec
3FGL J0241.3+6542	TXS 0237+655	40.34071	65.71992
3FGL J0242.3+1059	OD 166	40.62154	11.01687
3FGL J0243.5+7119	S5 0238+711	40.87871	71.33831
3FGL J0244.4-8224	PMN J0251-8226	42.7895	-82.44147
3FGL J0244.8-5818	RBS 0351	41.16788	-58.33181
3FGL J0245.4+2410	B2 0242+23	41.32023	24.0931
3FGL J0245.9-4651	PKS 0244-470	41.50046	-46.85483
3FGL J0249.1+8438	NVSS J024948+843556	42.45175	84.59903
3FGL J0250.6+1713	NVSS J025037+171209	42.65817	17.20247
3FGL J0250.6+5630	NVSS J025047+562935	42.69829	56.49317
3FGL J0251.5-5959	PKS 0250-602	42.85933	-60.00175
3FGL J0252.3+3830	B2 0249+38	43.28703	38.59028
3FGL J0252.8-2218	PKS 0250-225	43.19981	-22.32374
3FGL J0253.0-0125	FBQS J0253-0124	43.31496	-1.40143
3FGL J0253.1-5438	PKS 0252-549	43.37159	-54.69762
3FGL J0253.5+3216	MG3 J025334+3217	43.39021	32.28914
3FGL J0253.8+5104	NVSS J025357+510256	43.49003	51.04902
3FGL J0255.8+0532	PMN J0255+0533	43.95629	5.56531
3FGL J0256.3+0335	PKS B0253+033	44.11742	3.55764
3FGL J0257.8-1216	PMN J0257-1211	44.42085	-12.20038
3FGL J0258.0+2030	MG3 J025805+2029	44.53046	20.50044
3FGL J0259.5+0746	PKS 0256+075	44.86282	7.79435
3FGL J0301.4-1652	PMN J0301-1652	45.31926	-16.87919
3FGL J0301.8-7157	PKS 0301-721	45.41041	-71.94295
3FGL J0301.8-2721	NVSS J030158-272754	45.49196	-27.46519
3FGL J0302.0+5335	GB6 J0302+5331	45.59473	53.52958
3FGL J0302.5-7915	PMN J0303-7914	45.83721	-79.24911
3FGL J0303.0+3150	B2 0259+31	45.75636	31.84588
3FGL J0303.4-2407	PKS 0301-243	45.86037	-24.11986
3FGL J0303.6+4716	4C +47.08	45.89684	47.27119
3FGL J0303.7-6211	PKS 0302-623	45.96096	-62.19043
3FGL J0304.3-2836	RBS 0385	46.06821	-28.53836
3FGL J0304.9+6817	TXS 0259+681	46.09168	68.36041
3FGL J0305.2-1607	PKS 0302-16	46.31385	-16.13671
3FGL J0308.6+0408	NGC 1218	47.10927	4.11092
3FGL J0309.0+1029	PKS 0306+102	47.2651	10.48787
3FGL J0309.5-0749	NVSS J030943-074427	47.43017	-7.74097
3FGL J0309.9-6057	PKS 0308-611	47.48375	-60.97752
3FGL J0310.4-5015	1RXS J031036.0-501615	47.64413	-50.27575
3FGL J0310.8+3814	B3 0307+380	47.70783	38.24829
3FGL J0312.7+0133	PKS 0310+013	48.18168	1.55488
3FGL J0312.7+3613	V Zw 326	48.20954	36.25544
3FGL J0314.3-5103	PMN J0314-5104	48.60733	-51.07547
3FGL J0315.5-1026	PKS 0313-107	48.98754	-10.52825
3FGL J0316.1+0904	GB6 J0316+0904	49.05306	9.07869
3FGL J0316.1-2611	RBS 0405	49.06225	-26.13261
3FGL J0316.2-6436	SUMSS J031614-643732	49.06129	-64.62567
3FGL J0316.6+4119	IC 310	49.17907	41.32497
3FGL J0318.7+2134	MG3 J031849+2135	49.69029	21.57686
3FGL J0319.8+4130	NGC 1275	49.95067	41.5117
3FGL J0319.8+1847	RBS 0413	49.96583	18.75956
3FGL J0322.0+2335	MG3 J032201+2336	50.49988	23.60312
3FGL J0323.6-0109	1RXS J032342.6-011131	50.93174	-1.19616
3FGL J0325.2+3410	1H 0323+342	51.17151	34.1794
3FGL J0325.2-5634	1RXS J032521.8-563543	51.34796	-56.59569
3FGL J0325.5+2223	TXS 0322+222	51.40339	22.4001
3FGL J0325.6-1648	RBS 0421	51.42121	-16.77136
3FGL J0326.0-1842	PMN J0325-1843	51.47913	-18.73578
3FGL J0326.2+0225	1H 0323+022	51.55811	2.42077
3FGL J0331.3-6155	PMN J0331-6155	52.82646	-61.92444
3FGL J0332.0+6308	GB6 J0331+6307	52.974	63.13731
3FGL J0333.4+7853	NVSS J033344+785027	53.437	78.84103
3FGL J0333.4+4003	B3 0330+399	53.44508	40.11072
3FGL J0333.6+2916	TXS 0330+291	53.45425	29.27544
3FGL J0333.9+6538	TXS 0329+654	53.48641	65.61561
3FGL J0334.2+3915	4C +39.12	53.57679	39.35684
3FGL J0334.3-4008	PKS 0332-403	53.55689	-40.14039
3FGL J0334.3-3726	PMN J0334-3725	53.56425	-37.42869
3FGL J0335.3-4459	1RXS J033514.5-445929	53.81042	-44.99139
3FGL J0336.5+3210	NRAO 140	54.12545	32.30815
3FGL J0336.9-3622	PKS 0335-364	54.2251	-36.2684
3FGL J0336.9-1304	PKS 0334-131	54.14598	-13.03463
3FGL J0338.1-2443	1RXS J033810.1-244337	54.55221	-24.73058
3FGL J0338.5+1303	RX J0338.4+1302	54.62208	13.03778
3FGL J0339.2-1738	PKS 0336-177	54.80701	-17.60022
3FGL J0339.5-0146	PKS 0336-01	54.87891	-1.77661
3FGL J0340.5-2119	PKS 0338-214	55.14837	-21.32533
3FGL J0342.2+3857	GB6 J0342+3858	55.56779	38.98508
3FGL J0342.6-3006	PKS 0340-302	55.66817	-30.13303
3FGL J0343.2-2534	PKS 0341-256	55.83135	-25.50483
3FGL J0343.3-6443	PMN J0343-6442	55.83592	-64.71536
3FGL J0343.3+3622	OE 367	55.87064	36.37012
3FGL J0345.1-2353	NVSS J034518-235218	56.32625	-23.87219
3FGL J0348.6-2748	PKS 0346-27	57.15894	-27.82043
3FGL J0348.7-1606	PKS 0346-163	57.16363	-16.1716

Table A.1: – continued from previous page

3FGL Name	Associated Name	RA	Dec
3FGL J0349.2-1158	1ES 0347-121	57.34658	-11.99083
3FGL J0349.9-2102	PKS 0347-211	57.49094	-21.04659
3FGL J0351.1+0128	TXS 0348+013	57.73911	1.51827
3FGL J0352.9+5655	GB6 J0353+5654	58.28996	56.90864
3FGL J0353.0-3622	XRS J0353-3623	58.271	-36.38564
3FGL J0353.0-6831	PKS 0352-686	58.23962	-68.52133
3FGL J0354.1+4643	B3 0350+465	58.62505	46.72187
3FGL J0354.6+8011	S5 0346+80	58.69219	80.15801
3FGL J0356.3-6948	PMN J0357-6948	59.37529	-69.81244
3FGL J0357.1+2325	MG3 J035721+2319	59.34004	23.33162
3FGL J0357.1-4957	PKS 0355-500	59.25067	-49.93017
3FGL J0358.7+0633	PMN J0358+0629	59.61312	6.48872
3FGL J0358.8+6002	TXS 0354+599	59.761	60.08946
3FGL J0359.3-2612	PKS 0357-264	59.89034	-26.2587
3FGL J0401.4+2109	TXS 0358+210	60.43819	21.17461
3FGL J0401.8-3144	PKS 0400-319	60.58861	-31.79054
3FGL J0402.1-2618	PKS 0359-264	60.50329	-26.26081
3FGL J0403.7-2442	TXS 0401-248	60.92392	-24.73569
3FGL J0403.9-3604	PKS 0402-362	60.97396	-36.08386
3FGL J0405.5-1307	PKS 0403-13	61.39168	-13.13714
3FGL J0407.1-3825	PKS 0405-385	61.746208	-38.44075
3FGL J0407.5+0740	TXS 0404+075	61.87119	7.70208
3FGL J0409.4+3158	NVSS J040928+320245	62.36842	32.046
3FGL J0409.8-0358	NVSS J040946-040003	62.44408	-4.00106
3FGL J0413.6-5334	PMN J0413-5332	63.30629	-53.53361
3FGL J0416.6-1850	PKS 0414-189	64.15227	-18.85232
3FGL J0416.8+0104	1ES 0414+009	64.21871	1.08997
3FGL J0418.0-0251	PKS B0415-029	64.49263	-2.83867
3FGL J0418.5+3813c	3C 111	64.58866	38.02661
3FGL J0422.1-0642	PMN J0422-0643	65.54498	-6.72926
3FGL J0423.2-0119	PKS 0420-01	65.81584	-1.34252
3FGL J0423.8+4150	4C +41.11	65.98337	41.83409
3FGL J0424.7+0035	PKS 0422+00	66.19518	0.60176
3FGL J0425.0-5331	PMN J0425-5331	66.26783	-53.53297
3FGL J0425.2+6319	1RXS J042523.0+632016	66.34583	63.33778
3FGL J0426.3+6827	4C +68.05	66.70862	68.43137
3FGL J0426.6+0459	4C +04.15	66.57922	4.84049
3FGL J0427.3-3900	PMN J0427-3900	66.84037	-39.01664
3FGL J0428.6-3756	PKS 0426-380	67.16843	-37.93877
3FGL J0429.8+2843	MG2 J042948+2843	67.45829	28.71475
3FGL J0430.2-2508	PMN J0430-2507	67.56679	-25.1275
3FGL J0430.5+1655	MG1 J043022+1655	67.59313	16.918
3FGL J0431.6+7403	GB6 J0431+7403	67.93837	74.05767
3FGL J0433.1+3228	NVSS J043307+322840	68.28212	32.47789
3FGL J0433.6+2905	MG2 J043337+2905	68.407624	29.098744
3FGL J0433.7-6028	PKS 0432-606	68.39225	-60.50412
3FGL J0434.0-5726	SUMSS J043344-572613	68.43712	-57.43711
3FGL J0434.0-2010	TXS 0431-203	68.53296	-20.25475
3FGL J0434.4-2341	PMN J0434-2342	68.62071	-23.7015
3FGL J0434.6+0921	TXS 0431+092	68.67079	9.39686
3FGL J0438.3-1258	PKS 0436-129	69.64592	-12.85093
3FGL J0438.8-4519	PKS 0437-454	69.75356	-45.37293
3FGL J0439.6-3159	1RXS J043931.4-320045	69.88425	-32.01353
3FGL J0439.9-1859	PMN J0439-1900	69.95292	-19.01472
3FGL J0440.3-2500	RBS 0570	70.07721	-24.99281
3FGL J0440.3+1444	TXS 0437+145	70.08808	14.632487
3FGL J0440.8+2751	B2 0437+27B	70.20985	27.84635
3FGL J0442.6-0017	PKS 0440-00	70.66109	-0.29539
3FGL J0444.5+3425	B2 0441+34	71.15057	34.42888
3FGL J0444.6-6012	PMN J0444-6014	71.25629	-60.25
3FGL J0447.8-2119	PKS 0446-212	72.07242	-21.16245
3FGL J0448.6-1632	RBS 0589	72.15675	-16.54531
3FGL J0449.0+1121	PKS 0446+11	72.28196	11.35794
3FGL J0449.4-4350	PKS 0447-439	72.35287	-43.83581
3FGL J0453.2-2808	PKS 0451-28	73.31103	-28.12704
3FGL J0453.2+6321	NVSS J045312+632117	73.30167	63.35492
3FGL J0455.7-4617	PKS 0454-46	73.96155	-46.2663
3FGL J0456.3+2702	MG2 J045613+2702	74.07237	27.03919
3FGL J0456.3-3131	PMN J0456-3135	74.15283	-31.60347
3FGL J0457.0-2324	PKS 0454-234	74.26325	-23.41445
3FGL J0457.0+0643	4C +06.21	74.28212	6.75202
3FGL J0501.2-0157	S3 0458-02	75.30337	-1.98729
3FGL J0501.8+3046	1RXS J050140.8+304831	75.42	30.80861
3FGL J0502.5+0612	PKS 0459+060	75.56436	6.15208
3FGL J0502.7+3438	MG2 J050234+3436	75.62462	34.60969
3FGL J0503.4+4522	1RXS J050339.8+451715	75.91494	45.2832
3FGL J0503.5+6538	1E 0458.1+6530	75.77429	65.56711
3FGL J0505.3+0459	PKS 0502+049	76.3466	4.9952
3FGL J0505.3-0422	S3 0503-04	76.46349	-4.32406
3FGL J0505.5+0416	MG1 J050533+0415	76.39487	4.26516
3FGL J0505.5-1558	TXS 0503-160	76.42375	-15.976944
3FGL J0505.9+6114	NVSS J050558+611336	76.49504	61.22664
3FGL J0506.9-5435	1ES 0505-546	76.74092	-54.58436
3FGL J0507.1-6102	PKS 0506-61	76.68329	-61.16139

Table A.1: – continued from previous page

3FGL Name	Associated Name	RA	Dec
3FGL J0508.0+6736	1ES 0502+675	76.98438	67.62344
3FGL J0508.2-1936	PMN J0508-1936	77.07917	-19.59892
3FGL J0509.3+1012	PKS 0506+101	77.3644	10.19572
3FGL J0509.4+0541	TXS 0506+056	77.35819	5.69315
3FGL J0509.7-0400	1H 0506-039	77.40908	-4.01264
3FGL J0509.7-6418	RBS 0625	77.49125	-64.29472
3FGL J0510.0+1802	PKS 0507+17	77.50987	18.01155
3FGL J0512.2+2918	B2 0509+29	78.17586	29.451
3FGL J0512.9+4038	B3 0509+406	78.21893	40.69545
3FGL J0514.4+5603	TXS 0510+559	78.57792	56.0364
3FGL J0515.3-4557	PKS 0514-459	78.93877	-45.94537
3FGL J0515.5-0123	NVSS J051536-012427	78.901	-1.40772
3FGL J0515.8+1526	GB6 J0515+1527	78.94731	15.45461
3FGL J0516.3+7351	GB6 J0516+7350	79.13017	73.85239
3FGL J0516.7-6207	PKS 0516-621	79.18719	-62.11816
3FGL J0517.4+4540	4C +45.08	79.37041	45.61802
3FGL J0517.5+0902	PMN J0517+0858	79.41693	8.97661
3FGL J0519.2-4542	Pictor A	79.95717	-45.77883
3FGL J0519.3+2746	4C +27.15	79.8876	27.73452
3FGL J0519.5+0852	TXS 0516+087	79.79505	8.81576
3FGL J0521.4-1740	TXS 0519-176	80.34816	-17.62505
3FGL J0521.7+0103	NVSS J052140+010257	80.42004	1.04917
3FGL J0521.7+2113	TXS 0518+211	80.44152	21.21429
3FGL J0521.9-3847	PKS 0520-388	80.45521	-38.84197
3FGL J0522.9-3628	PKS 0521-36	80.7416	-36.45857
3FGL J0525.3-4558	PKS 0524-460	81.38083	-45.96519
3FGL J0525.6-6013	SUMSS J052542-601341	81.42521	-60.22828
3FGL J0525.8-2014	PMN J0525-2010	81.36575	-20.18019
3FGL J0526.0+4253	NVSS J052520+425520	81.33675	42.92225
3FGL J0526.2-4829	PKS 0524-485	81.56946	-48.51022
3FGL J0526.6+6321	GB6 J0526+6317	81.52792	63.29139
3FGL J0528.3+1815	1RXS J052829.6+181657	82.12333	18.2825
3FGL J0529.1+0933	GB6 J0529+0934	82.26071	9.57644
3FGL J0529.2-5917	1RXS J052846.9-592000	82.19542	-59.33347
3FGL J0529.8-7242	PKS 0530-727	82.37518	-72.75792
3FGL J0530.8+1330	PKS 0528+134	82.73507	13.53199
3FGL J0532.0-4827	PMN J0531-4827	82.99421	-48.45997
3FGL J0532.7+0732	OG 050	83.16249	7.54537
3FGL J0533.0-3939	PKS 0531-397	83.23827	-39.68584
3FGL J0533.2+4822	TXS 0529+483	83.31611	48.38134
3FGL J0533.6-8323	PKS 0541-834	83.41038	-83.40997
3FGL J0535.6-2749	PMN J0535-2751	83.96483	-27.86569
3FGL J0537.4-5717	SUMSS J053748-571828	84.45321	-57.30803
3FGL J0538.4-3909	NVSS J053810-390844	84.54317	-39.14514
3FGL J0538.8-4405	PKS 0537-441	84.70984	-44.08581
3FGL J0539.8+1434	TXS 0536+145	84.92652	14.56266
3FGL J0540.0-2837	PKS 0537-286	84.97617	-28.66554
3FGL J0540.4+5823	GB6 J0540+5823	85.12509	58.394
3FGL J0540.5-5416	PKS 0539-543	85.19096	-54.30594
3FGL J0542.2-8737	SUMSS J054923-874001	87.34946	-87.66697
3FGL J0542.5-0907c	PMN J0542-0913	85.73282	-9.22528
3FGL J0543.9-5531	1RXS J054357.3-553206	85.98838	-55.53536
3FGL J0550.6-3217	PKS 0548-322	87.66898	-32.27143
3FGL J0553.5-2036	NVSS J055333-203417	88.38808	-20.57164
3FGL J0556.0-4353	SUMSS J055618-435146	89.07812	-43.86281
3FGL J0558.1-3838	EXO 0556.4-3838	89.52671	-38.64211
3FGL J0558.6-7459	PKS 0600-749	89.692	-74.98472
3FGL J0600.9-3943	PKS 0558-396	90.13046	-39.61714
3FGL J0601.0+3837	B2 0557+38	90.26193	38.64126
3FGL J0601.2-7036	PKS 0601-70	90.29717	-70.60239
3FGL J0602.2+5314	GB6 J0601+5315	90.50178	53.26665
3FGL J0602.8-4016	SUMSS J060251-401845	90.71304	-40.31269
3FGL J0603.8+2155	4C +22.12	90.96482	21.9938
3FGL J0604.1-4817	1ES 0602-482	91.03917	-48.29056
3FGL J0604.7-4849	1RXS J060432.7-485007	91.13625	-48.83528
3FGL J0606.4-4729	ESO 254- G 017	91.64912	-47.49867
3FGL J0607.4+4739	TXS 0603+476	91.8469	47.66304
3FGL J0608.0-0835	PKS 0605-08	91.99875	-8.58055
3FGL J0608.1-1522	PMN J0608-1520	92.00638	-15.34361
3FGL J0609.4-0248	NVSS J060915-024754	92.31258	-2.79839
3FGL J0611.1-6100	PKS 0609-609	92.62646	-60.97703
3FGL J0611.2+4323	NVSS J061107+432404	92.7795	43.40133
3FGL J0611.7+2759	GB6 J0611+2803	92.93283	28.06453
3FGL J0612.8+4122	B3 0609+413	93.21327	41.37706
3FGL J0615.4-3116	PKS 0613-312	93.83004	-31.28933
3FGL J0617.2+5701	87GB 061258.1+570222	94.32051	57.02123
3FGL J0617.6-1717	CRATES J061733.67-171522.8	94.39029	-17.25633
3FGL J0618.0+7819	1REX J061757+7816.1	94.48722	78.26872
3FGL J0618.2-2429	PMN J0618-2426	94.59437	-24.44383
3FGL J0618.9-1138	TXS 0616-116	94.76709	-11.68191
3FGL J0620.4+2644	RX J0620.6+2644	95.16692	26.72553
3FGL J0622.4-2606	PMN J0622-2605	95.59196	-26.09578
3FGL J0622.9+3326	B2 0619+33	95.71759	33.43623
3FGL J0623.3+3043	GB6 J0623+3045	95.81746	30.74892

Table A.1: – continued from previous page

3FGL Name	Associated Name	RA	Dec
3FGL J0625.2+4440	GB6 J0625+4440	96.32611	44.66712
3FGL J0626.0-5436	PMN J0625-5438	96.46771	-54.64739
3FGL J0626.6-4259	1RXS J062635.9-425810	96.64958	-42.96958
3FGL J0627.0-3529	PKS 0625-35	96.77815	-35.48757
3FGL J0627.9-1517	NVSS J062753-152003	96.97175	-15.33433
3FGL J0629.0-6248	PKS 0628-627	97.23953	-62.81243
3FGL J0629.4-1959	PKS 0627-199	97.34901	-19.98881
3FGL J0630.3+6906	GB6 J0629+6900	97.34267	69.00542
3FGL J0630.9-2406	TXS 0628-240	97.74804	-24.11281
3FGL J0631.2+2019	TXS 0628+203	97.75443	20.34978
3FGL J0634.7-2334	PMN J0634-2335	98.74584	-23.58665
3FGL J0635.7-7517	PKS 0637-75	98.94378	-75.27134
3FGL J0638.6+7324	S5 0633+73	99.8415	73.41612
3FGL J0640.0-1252	TXS 0637-128	100.02998	-12.88766
3FGL J0641.8-0319	TXS 0639-032	100.46305	-3.34683
3FGL J0643.2+0859	PMN J0643+0857	100.86019	8.96056
3FGL J0643.4-5358	PMN J0643-5358	100.83421	-53.97967
3FGL J0644.3-6713	PKS 0644-671	101.11717	-67.21586
3FGL J0644.6-2853	1RXS J064444.2-285120	101.18417	-28.85556
3FGL J0645.9-3914	PKS 0644-390	101.62883	-39.06087
3FGL J0646.4-5452	PMN J0646-5451	101.62917	-54.85361
3FGL J0647.0-5134	1ES 0646-515	101.79208	-51.59611
3FGL J0647.1-4415	SUMSS J064648-441929	101.7005	-44.32486
3FGL J0647.6-6058	PMN J0647-6058	101.92029	-60.968
3FGL J0648.1+1606	1RXS J064814.1+160708	102.05875	16.11889
3FGL J0648.1-3045	PKS 0646-306	102.05873	-30.73879
3FGL J0648.8+1516	RX J0648.7+1516	102.19854	15.27356
3FGL J0648.8-1740	TXS 0646-176	102.11874	-17.73484
3FGL J0649.6-3138	1RXS J064933.8-313914	102.39083	-31.65389
3FGL J0650.4-1636	PKS 0648-16	102.60242	-16.6277
3FGL J0650.5+2055	1RXS J065033.9+205603	102.64679	20.93244
3FGL J0650.7+2503	1ES 0647+250	102.69371	25.04989
3FGL J0651.3+4014	RX J0651.0+4013	102.77257	40.22725
3FGL J0652.0-4808	1RXS J065201.0-480858	103.00417	-48.14944
3FGL J0653.6+2817	GB6 J0653+2816	103.4345	28.26314
3FGL J0654.4+4514	B3 0650+453	103.59881	45.23987
3FGL J0654.4+5042	GB6 J0654+5042	103.59205	50.70663
3FGL J0654.5+0926	RX J0654.3+0925	103.61317	9.42628
3FGL J0656.2-0323	TXS 0653-033	104.04634	-3.38522
3FGL J0656.4+4232	4C +42.22	104.04443	42.61743
3FGL J0658.3-5832	PMN J0658-5840	104.55763	-58.67422
3FGL J0658.6+0636	NVSS J065844+063711	104.68738	6.61986
3FGL J0700.0+1709	TXS 0657+172	105.00636	17.15603
3FGL J0700.2+1304	GB6 J0700+1304	105.05954	13.07356
3FGL J0700.3-6310	SUMSS J065958-631238	104.993	-63.21067
3FGL J0700.6-6610	PKS 0700-661	105.13017	-66.17939
3FGL J0701.4-4634	PKS 0700-465	105.39395	-46.57684
3FGL J0702.7-1952	TXS 0700-197	105.67875	-19.85612
3FGL J0703.4-3914	1RXS J070312.7-391417	105.803	-39.23833
3FGL J0706.1-4849	PMN J0705-4847	106.49454	-48.79014
3FGL J0706.5+3744	GB6 J0706+3744	106.63208	37.74344
3FGL J0707.0+7741	NVSS J070651+774137	106.71546	77.69386
3FGL J0707.2+6101	TXS 0702+612	106.75257	61.16989
3FGL J0708.9+2239	GB6 J0708+2241	107.24287	22.6932
3FGL J0709.7-0256	PMN J0709-0255	107.43773	-2.92153
3FGL J0710.3+5908	1H 0658+595	107.62529	59.139
3FGL J0710.5+4732	S4 0707+47	107.6921	47.53643
3FGL J0712.2-6436	MRC 0712-643	108.25842	-64.48503
3FGL J0712.6+5033	GB6 J0712+5033	108.18201	50.55631
3FGL J0713.9+1933	MG2 J071354+1934	108.482	19.58345
3FGL J0718.7-4319	PMN J0718-4319	109.68183	-43.33047
3FGL J0719.3+3307	B2 0716+33	109.83092	33.11936
3FGL J0720.0-4010	1RXS J071939.2-401153	109.91333	-40.19806
3FGL J0721.4+0404	PMN J0721+0406	110.34963	4.11228
3FGL J0721.9+7120	S5 0716+71	110.4727	71.34343
3FGL J0723.2-0728	1RXS J072259.5-073131	110.74867	-7.52633
3FGL J0723.7+2050	GB6 J0723+2051	110.95143	20.85854
3FGL J0724.1+2857	GB6 J0723+2859	110.97845	28.9916
3FGL J0725.2+1425	4C +14.23	111.32003	14.42049
3FGL J0725.8-0054	PKS 0723-008	111.461	-0.91571
3FGL J0726.6-4727	PMN J0726-4728	111.60938	-47.48142
3FGL J0728.0+4828	GB6 J0727+4827	111.99975	48.45556
3FGL J0729.5-3127	NVSS J072922-313128	112.34567	-31.5245
3FGL J0730.2-1141	PKS 0727-11	112.57963	-11.68683
3FGL J0730.3+6720	GB6 J0731+6718	112.85583	67.31308
3FGL J0730.5+3307	1RXS J073026.0+330727	112.60854	33.12297
3FGL J0730.5-0537	TXS 0728-054	112.61849	-5.59636
3FGL J0730.5-6606	PMN J0730-6602	112.70658	-66.03869
3FGL J0732.2-4638	PKS 0731-465	113.18486	-46.67474
3FGL J0733.5+5153	NVSS J073326+515355	113.36167	51.89886
3FGL J0733.8+5021	TXS 0730+504	113.46884	50.36918
3FGL J0733.8+4108	GB6 J0733+4111	113.44504	41.18889
3FGL J0734.3-7709	PKS 0736-770	113.68133	-77.18692
3FGL J0738.1+1741	PKS 0735+17	114.53081	17.70528

Table A.1: – continued from previous page

3FGL Name	Associated Name	RA	Dec
3FGL J0739.4+0137	PKS 0736+01	114.82514	1.61795
3FGL J0742.4-8133c	SUMSS J074220-813139	115.58825	-81.52736
3FGL J0742.6+5444	GB6 J0742+5444	115.66579	54.74018
3FGL J0742.6-5623	PMN J0743-5619	115.83542	-56.32617
3FGL J0744.1-3804	PMN J0743-3804	115.93675	-38.06567
3FGL J0744.3+7434	MS 0737.9+7441	116.02192	74.566
3FGL J0744.8-4028	PMN J0744-4032	116.15929	-40.53806
3FGL J0746.4+2540	B2 0743+25	116.60781	25.81726
3FGL J0746.6-0706	PMN J0746-0709	116.61454	-7.16419
3FGL J0746.6-4756	PMN J0746-4755	116.67604	-47.91531
3FGL J0746.9+8511	NVSS J074715+851208	116.81433	85.20228
3FGL J0747.2-3311	PKS 0745-330	116.83201	-33.17971
3FGL J0747.4+0904	RX J0747.3+0905	116.84233	9.09681
3FGL J0748.0-1639	TXS 0745-165	117.01285	-16.66396
3FGL J0748.3+2401	S3 0745+24	117.15045	24.0067
3FGL J0748.5+7910	JVAS J0750+7909	117.68027	79.15472
3FGL J0748.8+4929	NVSS J074837+493040	117.15738	49.51141
3FGL J0749.4+1059	TXS 0746+110	117.36411	10.9592
3FGL J0750.6+1232	OI 280	117.71686	12.51801
3FGL J0753.1+5353	4C +54.15	118.25577	53.88323
3FGL J0754.4-1148	TXS 0752-116	118.61024	-11.78804
3FGL J0754.8+4824	GB1 0751+485	118.69029	48.39743
3FGL J0756.3-6433	SUMSS J075625-643031	119.10529	-64.50883
3FGL J0757.0+0956	PKS 0754+100	119.27768	9.94301
3FGL J0758.1+1130	TXS 0755+117	119.53191	11.61279
3FGL J0758.7+3747	NGC 2484	119.61712	37.78661
3FGL J0758.9+2705	SDSS J075846.99+270515.5	119.69581	27.08766
3FGL J0800.9+4401	B3 0757+441	120.28448	44.01949
3FGL J0802.0+1005	NVSS J080204+100639	120.52033	10.11103
3FGL J0803.3-0339	TXS 0800-034	120.79925	-3.59944
3FGL J0804.0-3629	NVSS J080405-362919	121.02237	-36.48861
3FGL J0804.4+0418	MG1 J080357+0420	120.98518	4.35076
3FGL J0805.0-0622	1RXS J080458.3-062432	121.24292	-6.40889
3FGL J0805.2-0112	PKS B0802-010	121.3037	-1.18717
3FGL J0805.4+7534	RX J0805.4+7534	121.36087	75.57367
3FGL J0805.4+6144	TXS 0800+618	121.32575	61.73992
3FGL J0806.6+5933	SBS 0802+596	121.60808	59.51858
3FGL J0807.1-0541	PKS 0804-05	121.79007	-5.6872
3FGL J0807.1+7744	NVSS J080637+774607	121.65613	77.76864
3FGL J0807.9+4946	OJ 508	122.16528	49.84348
3FGL J0808.2-0751	PKS 0805-07	122.06473	-7.85275
3FGL J0809.5+4045	S4 0805+41	122.23605	40.87914
3FGL J0809.5+5342	87GB 080551.6+535010	122.42389	53.6903
3FGL J0809.6+3456	B2 0806+35	122.41203	34.92702
3FGL J0809.8+5218	1ES 0806+524	122.45494	52.31618
3FGL J0811.2-7529	PMN J0810-7530	122.76321	-75.50772
3FGL J0811.3+0146	OJ 014	122.86128	1.78117
3FGL J0812.0+0237	PMN J0811+0237	123.00775	2.62586
3FGL J0812.9+5555	NVSS J081251+555422	123.21435	55.90606
3FGL J0813.3+6509	GB6 J0812+6508	123.17018	65.15305
3FGL J0814.1-1012	NVSS J081411-101208	123.54875	-10.20289
3FGL J0814.5+2943	EXO 0811.2+2949	123.5887	29.67252
3FGL J0814.7+6428	GB6 J0814+6431	123.66329	64.52278
3FGL J0816.4-1311	PMN J0816-1311	124.11329	-13.19789
3FGL J0816.5+2049	SDSS J081649.78+205106.4	124.20743	20.85179
3FGL J0816.7-2421	PMN J0816-2421	124.16838	-24.35183
3FGL J0816.7+5739	SBS 0812+578	124.0947	57.65254
3FGL J0817.8-0935	TXS 0815-094	124.45729	-9.55848
3FGL J0818.2+4223	S4 0814+42	124.56667	42.37928
3FGL J0818.8+2751	5C 07.119	124.82857	27.79186
3FGL J0820.4+3640	MG2 J082018+3640	125.08412	36.66791
3FGL J0820.9-1258	PKS 0818-128	125.23937	-12.9831
3FGL J0822.9+4041	B3 0819+408	125.73982	40.69716
3FGL J0824.1+2434	B2 0821+24	126.13754	24.64531
3FGL J0824.9+5551	OJ 535	126.19682	55.87852
3FGL J0824.9+3916	4C +39.23	126.23118	39.27831
3FGL J0825.4-0213	PMN J0825-0204	126.32121	-2.0785
3FGL J0825.8-3217	PKS 0823-321	126.46405	-32.30645
3FGL J0825.9-2230	PKS 0823-223	126.50655	-22.50756
3FGL J0826.0+0307	PKS 0823+033	126.45974	3.15681
3FGL J0827.2-0711	PMN J0827-0708	126.77542	-7.14622
3FGL J0828.5+5217	TXS 0824+524	126.97374	52.29953
3FGL J0828.8-2420	NVSS J082841-241850	127.17383	-24.314
3FGL J0829.3+0901	TXS 0826+091	127.37635	8.9726
3FGL J0829.6-1137	NVSS J082939-114103	127.41354	-11.68425
3FGL J0830.3-5855	PMN J0829-5856	127.37987	-58.93353
3FGL J0830.7+2408	OJ 248	127.71703	24.18328
3FGL J0831.9+0430	PKS 0829+046	127.95365	4.49419
3FGL J0832.6+4914	OJ 448	128.09674	49.22251
3FGL J0834.1+4223	OJ 451	128.47452	42.40051
3FGL J0834.7+4403	B3 0831+442	128.7425	44.06059
3FGL J0835.4+0930	GB6 J0835+0936	128.93008	9.62167
3FGL J0836.3+2143	MG2 J083615+2138	129.06757	21.65099
3FGL J0836.5-2020	PKS 0834-20	129.1634	-20.2832

Table A.1: – continued from previous page

3FGL Name	Associated Name	RA	Dec
3FGL J0839.5+0102	PKS 0837+012	129.95671	1.07409
3FGL J0839.6+3538	FIRST J083943.3+354001	129.9307	35.66705
3FGL J0839.6+1803	TXS 0836+182	129.87801	18.04643
3FGL J0840.8+1315	3C 207	130.19828	13.20655
3FGL J0841.3-3554	NVSS J084121-355506	130.34017	-35.91836
3FGL J0841.4+7053	S5 0836+71	130.35152	70.89505
3FGL J0842.0-6055	PMN J0842-6053	130.61062	-60.89689
3FGL J0843.9+5311	NVSS J084411+531250	131.04872	53.21407
3FGL J0845.1-5458	PMN J0845-5458	131.26029	-54.96911
3FGL J0846.7-0651	TXS 0845-068	131.9864	-7.05469
3FGL J0846.9-2336	PMN J0847-2337	131.7565	-23.61711
3FGL J0847.1+1134	RX J0847.1+1133	131.80388	11.56396
3FGL J0849.1+6607	GB6 J0848+6605	132.22758	66.10264
3FGL J0849.3+0458	TXS 0846+051	132.38563	4.91886
3FGL J0849.5-2912	NVSS J084922-291149	132.34246	-29.19694
3FGL J0849.9-3540	PMN J0849-3541	132.4401	-35.68369
3FGL J0849.9+5108	SBS 0846+513	132.49157	51.1414
3FGL J0850.0+4855	GB6 J0850+4855	132.50221	48.91461
3FGL J0850.2-1214	PMN J0850-1213	132.54015	-12.22649
3FGL J0850.2+3500	RX J0850.5+3455	132.65074	34.92299
3FGL J0851.8+5531	SDSS J085135.93+552834.5	132.89971	55.47626
3FGL J0852.6-5756	PMN J0852-5755	133.16125	-57.92506
3FGL J0853.0-3654	NVSS J085310-365820	133.29383	-36.97233
3FGL J0854.2+4408	B3 0850+443	133.54121	44.14175
3FGL J0854.8+2006	OJ 287	133.70365	20.10851
3FGL J0855.2-0718	PKS 0852-07	133.78978	-7.25082
3FGL J0856.5+2057	TXS 0853+211	134.16558	20.96206
3FGL J0856.7-1105	PMN J0856-1105	134.17418	-11.08734
3FGL J0858.1-1951	PKS 0855-19	134.52235	-19.84359
3FGL J0858.1-3130	1RXS J085802.6-313043	134.51083	-31.51194
3FGL J0859.1+6219	1RXS J085930.5+621737	134.87767	62.29177
3FGL J0902.4+2050	NVSS J090226+205045	135.61213	20.84623
3FGL J0903.1+4649	S4 0859+47	135.76663	46.85115
3FGL J0904.3+4240	S4 0900+42	136.06512	42.63466
3FGL J0904.8-5734	PKS 0903-57	136.22158	-57.58494
3FGL J0904.8-3516	NVSS J090442-351423	136.1765	-35.23978
3FGL J0904.9+2739	GB6 J0905+2748A	136.26688	27.80492
3FGL J0905.5+1358	MG1 J090534+1358	136.39579	13.96842
3FGL J0906.3-0906	PMN J0906-0905	136.57538	-9.09556
3FGL J0909.0+2310	RX J0908.9+2311	137.252606	23.186937
3FGL J0909.1+0121	PKS 0906+01	137.29205	1.35989
3FGL J0909.6+0157	PKS 0907+022	137.41603	2.00146
3FGL J0909.8-0229	PKS 0907-023	137.43718	-2.5251
3FGL J0910.5+3329	Ton 1015	137.65431	33.49012
3FGL J0910.7+3858	FBQS J091052.0+390202	137.71682	39.0339
3FGL J0910.9+2248	TXS 0907+230	137.67555	22.80988
3FGL J0911.8+3351	MG2 J091151+3349	137.94901	33.82133
3FGL J0912.2+4126	B3 0908+416B	138.04839	41.43593
3FGL J0912.4+2800	1RXS J091211.9+275955	138.04672	27.99109
3FGL J0912.6-2757	PMN J0912-2752	138.13161	-27.87145
3FGL J0912.7+1556	SDSS J091230.61+155528.0	138.12755	15.92445
3FGL J0912.9-2104	MRC 0910-208	138.25092	-21.05586
3FGL J0915.0+5844	SDSS J091608.57+584434.0	139.03574	58.7428
3FGL J0915.8+2933	Ton 0396	138.96834	29.55668
3FGL J0916.3+3857	S4 0913+39	139.20377	38.90782
3FGL J0917.3-0344	NVSS J091714-034315	139.31079	-3.72092
3FGL J0920.9+4442	S4 0917+44	140.24358	44.69833
3FGL J0921.0-2258	NVSS J092057-225721	140.23946	-22.95606
3FGL J0921.8+6215	OK 630	140.40096	62.26449
3FGL J0922.4-0529	TXS 0919-052	140.59864	-5.48533
3FGL J0922.8-3959	PKS 0920-39	140.69341	-39.99307
3FGL J0923.1+3853	B2 0920+39	140.81022	38.82775
3FGL J0923.3+4127	B3 0920+416	140.88044	41.42429
3FGL J0924.0+2816	B2 0920+28	140.96468	28.25695
3FGL J0924.2+0534	RBS 0771	141.00434	5.56261
3FGL J0925.6+5959	NVSS J092542+595812	141.42878	59.97127
3FGL J0925.7+3129	B2 0922+31B	141.43188	31.453
3FGL J0926.3+5409	NVSS J092638+541126	141.66199	54.19073
3FGL J0927.9-2037	PKS 0925-203	141.96593	-20.5809
3FGL J0928.5+4048	1RXS J092837.8+404858	142.15595	40.81254
3FGL J0928.7+7300	GB6 J0929+7304	142.42565	73.06793
3FGL J0928.9-3530	NVSS J092849-352947	142.20788	-35.49647
3FGL J0929.4+5013	GB6 J0929+5013	142.31433	50.22666
3FGL J0930.0+4951	1ES 0927+500	142.65663	49.84043
3FGL J0930.2+8612	S5 0916+864	142.4294	86.20591
3FGL J0934.1+3933	GB6 J0934+3926	143.52779	39.44226
3FGL J0937.7+5008	GB6 J0937+5008	144.30136	50.1478
3FGL J0939.2-1732	TXS 0936-173	144.82998	-17.52661
3FGL J0939.9-2831	TXS 0937-282	145.02025	-28.49161
3FGL J0940.7-6102	MRC 0939-608	145.19754	-61.12408
3FGL J0940.9-1337	TXS 0938-133	145.26062	-13.5975
3FGL J0941.6+2727	MG2 J094148+2728	145.45048	27.47745
3FGL J0942.1-0756	PMN J0942-0800	145.58942	-7.99811
3FGL J0945.9+5756	GB6 J0945+5757	146.426	57.96325

Table A.1: – continued from previous page

3FGL Name	Associated Name	RA	Dec
3FGL J0946.2+0103	RXS J094620.5+010459	146.58424	1.08112
3FGL J0946.5+1017	TXS 0943+105	146.64612	10.28504
3FGL J0947.1-2542	1RXS J094709.2-254056	146.78963	-25.68336
3FGL J0948.6+4041	4C +40.24	147.23059	40.66238
3FGL J0948.8+0021	PMN J0948+0022	147.23883	0.37377
3FGL J0950.1+4554	RX J0950.2+4553	147.54922	45.8889
3FGL J0953.0-0839	PMN J0953-0840	148.26129	-8.67177
3FGL J0953.1-7657c	1RXS J095306.1-765755	148.26675	-76.96789
3FGL J0954.2+4913	1ES 0950+495	148.5411	49.24988
3FGL J0956.6+2515	OK 290	149.20781	25.25446
3FGL J0956.7-6441	AT20G J095612-643928	149.05075	-64.65781
3FGL J0957.4+4728	OK 492	149.58196	47.41884
3FGL J0957.5-1351	PMN J0957-1350	149.32576	-13.83366
3FGL J0957.6+5523	4C +55.17	149.4091	55.38271
3FGL J0958.3-0318	1RXS J095806.4-031729	149.52458	-3.29447
3FGL J0958.4-6752	1RXS J095812.8-675241	149.55333	-67.87806
3FGL J0958.6+6534	S4 0954+65	149.69685	65.56523
3FGL J0958.6-2447	TXS 0956-244	149.58529	-24.73322
3FGL J0959.7+2124	RX J0959.4+2123	149.87453	21.38918
3FGL J1001.0+2913	GB6 J1001+2911	150.29252	29.19376
3FGL J1002.3+2220	1RXS J100235.8+221609	150.64342	22.27081
3FGL J1003.6+2608	PKS 1000+26	150.92597	26.08694
3FGL J1005.0-4959	PMN J1006-5018	151.55837	-50.30374
3FGL J1006.7+3453	EXO 1004.0+3509	151.73527	34.91255
3FGL J1006.7-2159	PKS 1004-217	151.69339	-21.989
3FGL J1007.4-3334	PKS 1005-333	151.88078	-33.55187
3FGL J1007.8+0026	PKS 1005+007	152.04766	0.5
3FGL J1007.9+0621	MG1 J100800+0621	152.0034	6.35589
3FGL J1008.9-2910	PMN J1008-2912	152.18796	-29.21217
3FGL J1009.0-3137	PKS 1006-313	152.21062	-31.65153
3FGL J1010.2-3120	1RXS J101015.9-311909	152.56654	-31.31897
3FGL J1010.8-0158	PKS 1008-01	152.71914	-2.01584
3FGL J1012.2+0631	NRAO 350	153.05562	6.51589
3FGL J1012.6+2439	MG2 J101241+2439	153.17242	24.6565
3FGL J1012.7+4229	B3 1009+427	153.18453	42.49917
3FGL J1013.5+3440	OL 318	153.45673	34.76411
3FGL J1014.2+4115	GB6 J1014+4112	153.57446	41.20492
3FGL J1015.0+4925	1H 1013+498	153.76725	49.43353
3FGL J1015.2-4512	PMN J1014-4508	153.70981	-45.14476
3FGL J1016.0+0513	TXS 1013+054	154.01307	5.21732
3FGL J1016.0-0635	NVSS J101626-063624	154.11224	-6.60733
3FGL J1016.1+5555	TXS 1012+560	153.93508	55.85015
3FGL J1018.1+1904	NVSS J101808+190614	154.53338	19.10408
3FGL J1018.3+3542	B2 1015+35B	154.54578	35.71096
3FGL J1018.4-3119	PKS 1016-311	154.61981	-31.39829
3FGL J1018.5+0530	TXS 1015+057	154.61603	5.50832
3FGL J1018.8+5913	TXS 1015+594	154.74388	59.19108
3FGL J1020.0+6323	GB6 J1019+6319	154.96199	63.33378
3FGL J1021.8+8023	NVSS J102201+802350	155.50825	80.39742
3FGL J1022.3-4234	PMN J1022-4232	155.57458	-42.54139
3FGL J1022.8-0113	RX J1022.7-0112	155.68219	-1.21736
3FGL J1023.1+3952	4C +40.25	155.79819	39.80427
3FGL J1023.7+3000	RX J1023.6+3001	155.91564	30.01604
3FGL J1023.9-4335	RX J1023.9-4336	155.98413	-43.60042
3FGL J1024.1-3232	PKS 1021-323	156.00177	-32.57113
3FGL J1024.8+0105	PMN J1024+0056	156.15735	0.93716
3FGL J1025.1+2333	MG2 J102456+2332	156.22349	23.54277
3FGL J1026.4-8542	PKS 1029-85	156.6465	-85.72086
3FGL J1026.5+7423	GB6 J1027+7428	156.85061	74.47392
3FGL J1026.9-1750	1RXS J102658.5-174905	156.74404	-17.81633
3FGL J1027.0+0609	NVSS J102703+060934	156.76413	6.15942
3FGL J1027.7+6316	RX J1027.4+6317	156.85396	63.29807
3FGL J1028.0+1829	GB6 J1027+1831	156.93764	18.52743
3FGL J1028.5-0235	PMN J1028-0237	157.14183	-2.61658
3FGL J1030.4-2030	NVSS J103040-203032	157.66858	-20.50908
3FGL J1031.0+7440	S5 1027+74	157.84177	74.69954
3FGL J1031.2+5053	1ES 1028+511	157.82716	50.89328
3FGL J1031.6+6021	TXS 1028+605	157.93648	60.34177
3FGL J1032.5+6623	SDSS J103239.06+662323.2	158.1628	66.3898
3FGL J1032.7+3735	B3 1029+378	158.16974	37.64084
3FGL J1033.2+4116	S4 1030+41	158.26545	41.2684
3FGL J1033.8+6051	S4 1030+61	158.46429	60.85204
3FGL J1035.2+5545	GB6 J1035+5542	158.93978	55.71513
3FGL J1037.0-2934	PKS 1034-293	159.317	-29.56745
3FGL J1037.4-3742	PKS 1034-374	159.22267	-37.73752
3FGL J1037.5-2821	PKS B1035-281	159.42691	-28.38447
3FGL J1037.5+5711	GB6 J1037+5711	159.43462	57.19879
3FGL J1038.9-5311	MRC 1036-529	159.669	-53.19525
3FGL J1040.4+0615	GB6 J1040+0617	160.13179	6.28936
3FGL J1040.8+1342	1RXS J104057.7+134216	160.24042	13.70444
3FGL J1040.9-1205	NVSS J104108-120332	160.28612	-12.05844
3FGL J1041.8+3901	B3 1038+392	160.45479	39.02219
3FGL J1042.0-0557	PMN J1042-0558	160.51696	-5.97072
3FGL J1042.1-4126	1RXS J104204.1-412936	160.51707	-41.49333

Table A.1: – continued from previous page

3FGL Name	Associated Name	RA	Dec
3FGL J1043.1+2407	B2 1040+24A	160.78765	24.14317
3FGL J1044.4+8058	S5 1039+81	161.09609	80.91096
3FGL J1045.7-2926	PKS B1043-291	161.41927	-29.45732
3FGL J1046.9-2531	NVSS J104651-253547	161.71417	-25.59589
3FGL J1047.6+7240	GB6 J1047+7238	161.948	72.63694
3FGL J1047.8-6216	PMN J1047-6217	161.92892	-62.28728
3FGL J1048.4+7144	S5 1044+71	162.11508	71.72665
3FGL J1048.6+2338	NVSS J104900+233821	162.25106	23.63894
3FGL J1049.8+1425	MG1 J104945+1429	162.44303	14.49405
3FGL J1051.4+3941	RBS 0909	162.85577	39.72381
3FGL J1051.5-6517	PKS 1049-650	162.84801	-65.3024
3FGL J1051.8+0105	NVSS J105151+010312	162.96598	1.053
3FGL J1052.8-3741	PMN J1053-3743	163.24208	-37.72178
3FGL J1053.7+4929	GB6 J1053+4930	163.43387	49.49889
3FGL J1054.5+2210	87GB 105148.6+222705	163.6276	22.18191
3FGL J1057.3-2341	PKS B1054-234	164.35175	-23.70048
3FGL J1057.6-2754	RX J1057.8-2753	164.46158	-27.903
3FGL J1058.1+7010	S5 1053+70	164.22341	70.19609
3FGL J1058.4+8112	S5 1053+81	164.54806	81.24241
3FGL J1058.5+0133	4C +01.28	164.62336	1.56634
3FGL J1058.5-8003	PKS 1057-79	164.68046	-80.06504
3FGL J1058.6+5627	TXS 1055+567	164.6572	56.46978
3FGL J1059.2-1133	PKS B1056-113	164.80178	-11.57299
3FGL J1059.9+2056	MG2 J105938+2057	164.91268	20.9561
3FGL J1100.5+4020	RX J1100.3+4019	165.08786	40.32435
3FGL J1101.5+4106	RX J1101.3+4108	165.35302	41.14649
3FGL J1103.1+1155	TXS 1100+122	165.76471	11.97128
3FGL J1103.5-2329	1ES 1101-232	165.90671	-23.492
3FGL J1103.9-5357	PKS 1101-536	165.96759	-53.95019
3FGL J1104.3+0730	MG1 J110424+0730	166.1003	7.51477
3FGL J1104.4+3812	Mkn 421	166.11381	38.20883
3FGL J1105.9+2814	MG2 J110606+2812	166.53026	28.21307
3FGL J1106.4-3643	PMN J1106-3647	166.60017	-36.78317
3FGL J1107.4-4447	PKS 1104-445	166.78622	-44.81877
3FGL J1107.5+0223	NVSS J110735+022225	166.89965	2.37349
3FGL J1107.8+1502	RX J1107.7+1502	166.95029	15.03627
3FGL J1109.4+2411	1ES 1106+244	167.31717	24.189
3FGL J1109.4-4815	PMN J1109-4815	167.32883	-48.25553
3FGL J1109.6+3734	NVSS J110938+373609	167.41045	37.60325
3FGL J1110.0+7134	RX J1110.5+7133	167.65666	71.56571
3FGL J1110.4-1835	CRATES J1110-1835	167.66021154	-17.402
3FGL J1112.4+3449	TXS 1109+350	168.16154	34.77753
3FGL J1112.6+1749	1RXS J111224.2+175131	168.10253	17.85606
3FGL J1117.0+2014	RBS 0958	169.27608	20.23539
3FGL J1117.3+2546	RX J1117.6+2548	169.41832	25.81294
3FGL J1117.7-4632	PKS 1116-46	169.61232	-46.57083
3FGL J1117.9+5355	NVSS J111757+535553	169.48851	53.93193
3FGL J1118.2-0411	PMN J1118-0413	169.55179	-4.22339
3FGL J1119.7-3046	1RXS J111941.0-304652	169.91458	-30.78894
3FGL J1120.8+4212	RBS 0970	170.20028	42.20346
3FGL J1121.4-0554	PKS 1118-05	170.35462	-5.89901
3FGL J1123.2-6415	AT20G J112319-641735	170.83117	-64.29325
3FGL J1123.6+7231	RX J1123.8+7230	170.95501	72.50002
3FGL J1124.1+2337	OM 235	171.01127	23.61274
3FGL J1124.9+4932	GB6 J1124+4933	171.22427	49.56938
3FGL J1125.0-2101	PMN J1125-2100	171.28592	-21.01836
3FGL J1125.5-3558	PMN J1125-3556	171.38118	-35.95093
3FGL J1125.8-0745	1RXS J112551.6-074219	171.46662	-7.70586
3FGL J1125.9+2007	4C +20.25	171.49476	20.09843
3FGL J1126.7-3834	PKS 1124-382	171.68389	-38.4789
3FGL J1127.0-1857	PKS 1124-186	171.7683	-18.95484
3FGL J1127.8+3618	MG2 J112758+3620	171.99529	36.34121
3FGL J1128.0+5921	TXS 1125+596	172.05559	59.42078
3FGL J1129.0+3705	MG2 J112910+3702	172.30963	37.05497
3FGL J1129.4-4215	SUMSS J113006-421441	172.52921	-42.24475
3FGL J1129.9-1446	PKS 1127-14	172.52939	-14.82427
3FGL J1131.1+5810	1RXS J113117.8+580911	172.82765	58.14968
3FGL J1131.4+3819	B2 1128+38	172.72201	38.25515
3FGL J1131.9-0503	PKS 1128-047	172.87715	-5.00546
3FGL J1132.7+0034	PKS B1130+008	173.19008	0.57439
3FGL J1132.8+1015	4C +10.33	173.24788	10.39507
3FGL J1136.1-7411	PKS 1133-739	174.04025	-74.26258
3FGL J1136.4+3405	MG2 J113627+3408	174.11393	34.12763
3FGL J1136.6+7009	Mkn 180	174.11004	70.15758
3FGL J1136.6+6736	RX J1136.5+6737	174.12533	67.61789
3FGL J1136.6-6826	PKS 1133-681	174.00921	-68.4515
3FGL J1136.9+2551	RX J1136.8+2551	174.20879	25.84788
3FGL J1138.2+4905	GB6 J1138+4858	174.50871	48.98256
3FGL J1140.4+1529	NVSS J114023+152808	175.09784	15.46937
3FGL J1141.2+6805	1RXS J114118.3+680433	175.32625	68.07584
3FGL J1141.6-1406	1RXS J114142.2-140757	175.42583	-14.1325
3FGL J1142.0+1546	MG1 J114208+1547	175.53223	15.79838
3FGL J1143.0+6123	87GB 114026.7+613850	175.80041	61.36966
3FGL J1145.1+1935	3C 264	176.27087	19.60632

Table A.1: – continued from previous page

3FGL Name	Associated Name	RA	Dec
3FGL J1145.8+4425	B3 1143+446A	176.41049	44.33942
3FGL J1146.8+3958	S4 1144+40	176.74291	39.9762
3FGL J1147.0-3811	PKS 1144-379	176.75571	-38.20306
3FGL J1147.8-0725	PKS 1145-071	176.96481	-7.41143
3FGL J1149.5+2443	RX J1149.5+2439	177.37625	24.6575
3FGL J1150.3+2417	OM 280	177.58005	24.29829
3FGL J1150.5+4155	RBS 1040	177.64482	41.91114
3FGL J1151.4-1346	PMN J1151-1347	177.87488	-13.79744
3FGL J1151.4+5858	TXS 1148+592	177.85275	58.98821
3FGL J1152.3-0841	PKS B1149-084	178.07171	-8.68425
3FGL J1153.4+4932	OM 484	178.35194	49.51912
3FGL J1153.4+4033	B3 1151+408	178.47775	40.61462
3FGL J1153.7-2555	PMN J1153-2553	178.36008	-25.90189
3FGL J1154.0-3243	PKS 1151-324	178.52569	-32.71194
3FGL J1154.2-0010	1RXS J115404.9-001008	178.51898	-0.1694
3FGL J1154.3+6023	RX J1154.0+6022	178.5189	60.37245
3FGL J1155.4-3417	NVSS J115520-341718	178.83513	-34.28842
3FGL J1155.9+6136	SDSS J115548.40+613553.8	178.95169	61.5983
3FGL J1156.7-2250	NVSS J115633-225004	179.13825	-22.83444
3FGL J1158.8+0941	GB6 J1158+0937	179.72657	9.61993
3FGL J1158.9+0818	RX J1158.8+0819	179.71917	8.33194
3FGL J1159.2-2141	PMN J1159-2142	179.83929	-21.71247
3FGL J1159.3-2226	PKS 1156-221	179.79695	-22.47692
3FGL J1159.5+2914	Ton 599	179.88264	29.24551
3FGL J1159.6-0723	PMN J1159-0723	179.88288	-7.39986
3FGL J1200.8+1228	GB6 J1200+1230	180.16684	12.51757
3FGL J1200.9+2010	TXS 1158+204	180.23797	20.14572
3FGL J1203.1+6029	SBS 1200+608	180.76461	60.52199
3FGL J1203.2+3847	NVSS J120257+385147	180.73792	38.86332
3FGL J1203.5-3925	PMN J1203-3926	180.82413	-39.4385
3FGL J1204.0+1144	1RXS J120413.0+114549	181.05049	11.76539
3FGL J1204.3-0708	1RXS J120417.0-070959	181.06944	-7.16917
3FGL J1205.4+0412	MG1 J120448+0408	181.21533	4.13725
3FGL J1205.8-2636	PKS 1203-26	181.38838	-26.56791
3FGL J1207.6-4537	PMN J1207-4531	181.92667	-45.52167
3FGL J1207.6-2232	NVSS J120738-223250	181.90879	-22.54747
3FGL J1208.2-7810c	PKS 1205-778	182.07583	-78.16336
3FGL J1208.7+5442	TXS 1206+549	182.22607	54.69949
3FGL J1209.4+4119	B3 1206+416	182.34495	41.32816
3FGL J1209.8+1810	MG1 J120953+1809	182.46569	18.16855
3FGL J1212.6+5135	1RXS J121301.8+512942	183.25333	51.49323
3FGL J1213.1-2619	RBS 1080	183.34635	-26.3021
3FGL J1213.7+1306	4C +13.46	183.38398	13.12238
3FGL J1215.0+5002	NVSS J121500+500216	183.75327	50.03768
3FGL J1215.1+1658	TXS 1212+171	183.76658	16.91054
3FGL J1217.8+3007	1ES 1215+303	184.46701	30.11684
3FGL J1218.0-0029	PKS 1215-002	184.4947	-0.49619
3FGL J1218.4-0121	PKS 1216-010	184.64554	-1.33176
3FGL J1218.5+6912	NVSS J122044+690522	185.18421	69.08961
3FGL J1218.8-4827	PMN J1219-4826	184.7595	-48.44114
3FGL J1219.7-0314	1RXS J121946.0-031419	184.94045	-3.23999
3FGL J1220.2+3434	GB2 1217+348	185.03456	34.52271
3FGL J1220.2+7105	S5 1217+71	185.01512	71.09198
3FGL J1221.3+3010	PG 1218+304	185.34142	30.17697
3FGL J1221.4+2814	W Comae	185.38204	28.23292
3FGL J1222.4+0414	4C +04.42	185.59396	4.22105
3FGL J1222.7+8041	S5 1221+80	185.91872	80.66787
3FGL J1223.3-3028	NVSS J122337-303246	185.9055	-30.54614
3FGL J1224.5+4957	SBS 1221+503	186.04129	50.03208
3FGL J1224.5+2436	MS 1221.8+2452	186.10082	24.60654
3FGL J1224.6+4332	B3 1222+438	186.21461	43.58869
3FGL J1224.6-8312	PKS 1221-82	186.2266	-83.21947
3FGL J1224.9+2122	4C +21.35	186.22691	21.37955
3FGL J1225.7-7314	PMN J1225-7313	186.397	-73.22772
3FGL J1226.8+0638	1RXS J122645.2+063906	186.68428	6.64811
3FGL J1226.9-1329	PMN J1226-1328	186.72675	-13.47742
3FGL J1228.7+4857	TXS 1226+492	187.2157	48.96702
3FGL J1229.1+0202	3C 273	187.27792	2.05239
3FGL J1229.8-5305	AT20G J122939-530332	187.41637	-53.05894
3FGL J1230.3+2519	ON 246	187.55871	25.30198
3FGL J1230.9+1224	M 87	187.70593	12.39112
3FGL J1231.5+6414	MS 1229.2+6430	187.88078	64.23843
3FGL J1231.7+2847	B2 1229+29	187.93158	28.79717
3FGL J1231.8+1421	GB6 J1231+1421	187.84958	14.35681
3FGL J1233.7-0145	NVSS J123341-014426	188.42221	-1.73993
3FGL J1233.9-5736	AT20G J123407-573552	188.52933	-57.59803
3FGL J1236.6+3901	RX J1236.4+3859	189.09594	39.0003
3FGL J1237.9+6258	1H 1241+626	189.41282	62.97857
3FGL J1238.2-1958	PMN J1238-1959	189.60163	-19.98711
3FGL J1238.3-4543	PMN J1238-4541	189.52525	-45.69164
3FGL J1239.4+0727	PKS 1236+077	189.85245	7.50477
3FGL J1239.5+0443	MG1 J123931+0443	189.88648	4.71812
3FGL J1241.6-1456	RX J1241.8-1455	190.45567	-14.93289
3FGL J1241.9+0639	1ES 1239+069	190.45117	6.60031

Table A.1: – continued from previous page

3FGL Name	Associated Name	RA	Dec
3FGL J1243.1+3627	Ton 116	190.80307	36.46222
3FGL J1243.9-0217	PMN J1243-0218	190.96871	-2.31067
3FGL J1244.1+1615	SDSS J124444.35+161621.7	191.18483	16.27272
3FGL J1244.3-4955	SUMSS J124422-495422	191.09579	-49.90614
3FGL J1244.8+5707	1RXS J124510.5+571020	191.29166	57.1651
3FGL J1246.7-2547	PKS 1244-255	191.69501	-25.79702
3FGL J1247.0+4421	RX J1246.9+4423	191.75303	44.38855
3FGL J1248.0+5130	RX J1248.4+5128	192.14296	51.46886
3FGL J1248.2+5820	PG 1246+586	192.07827	58.34131
3FGL J1249.7+3705	RX J1249.8+3708	192.44583	37.13667
3FGL J1250.5+0217	PKS 1247+025	192.63575	2.2756
3FGL J1251.0-0203	TXS 1248-017	192.82658	-2.03542
3FGL J1251.3+1041	1RXS J125117.4+103914	192.825	10.65167
3FGL J1253.2+5300	S4 1250+53	193.29967	53.01993
3FGL J1253.7+0327	MG1 J125348+0326	193.44588	3.44178
3FGL J1254.1-2203	NVSS J125422-220413	193.59383	-22.07044
3FGL J1254.1+6240	1RXS J125400.1+624303	193.49708	62.716
3FGL J1254.5+2210	TXS 1252+224	193.63874	22.18438
3FGL J1254.9-4423	PKS 1252-441	193.73964	-44.41572
3FGL J1256.1-0547	3C 279	194.04653	-5.78931
3FGL J1256.1-5919	PMN J1256-5919	194.02054	-59.32872
3FGL J1256.3-1146	PMN J1256-1146	194.06646	-11.77706
3FGL J1256.7+5328	TXS 1254+538	194.16096	53.57325
3FGL J1256.9+3649	1RXS J125716.0+364713	194.31907	36.78754
3FGL J1258.0+6120	NVSS J125820+612049	194.58664	61.34597
3FGL J1258.1+3233	ON 393	194.48847	32.49148
3FGL J1258.6-1800	PKS B1256-177	194.65942	-18.00092
3FGL J1258.7-2219	PKS 1256-220	194.72699	-22.32531
3FGL J1258.7+5137	NVSS J125825+514225	194.606	51.70717
3FGL J1259.0-2310	PKS B1256-229	194.78526	-23.1774
3FGL J1259.8-3749	NVSS J125949-374856	194.958	-37.81558
3FGL J1300.2+1416	OW 197	195.08716	14.28848
3FGL J1302.6+5748	TXS 1300+580	195.7186	57.81045
3FGL J1303.0+2435	MG2 J130304+2434	195.76339	24.56547
3FGL J1303.7-4619	PMN J1303-4621	195.91779	-46.35069
3FGL J1304.2-2411	PMN J1304-2412	196.06987	-24.20469
3FGL J1304.3-5535	PMN J1303-5540	195.95508	-55.67542
3FGL J1304.3-4353	1RXS J130421.2-435308	196.08721	-43.88575
3FGL J1304.8-0338	PKS 1302-035	196.18184	-3.76737
3FGL J1304.9-2109	PKS B1302-208	196.24617	-21.11163
3FGL J1305.5+7854	S5 1304+79	196.25006	78.90994
3FGL J1306.8-2146	PKS 1304-215	196.67596	-21.8033
3FGL J1307.6-4300	1RXS J130737.8-425940	196.9075	-42.99444
3FGL J1308.1-6707	PKS 1304-668	197.07183	-67.11797
3FGL J1308.7+3545	5C 12.291	197.09879	35.77699
3FGL J1309.3+4304	B3 1307+433	197.35636	43.08488
3FGL J1309.5+1154	4C +12.46	197.39139	11.90682
3FGL J1310.2-1159	TXS 1307-117	197.55192	-11.96297
3FGL J1310.6+2446	MG2 J131037+2447	197.65956	24.80043
3FGL J1310.6+3222	OP 313	197.61943	32.3455
3FGL J1310.7+5515	TXS 1308+554	197.76338	55.23176
3FGL J1311.0+0036	RX J1311.1+0035	197.77697	0.58612
3FGL J1312.5-2155	PKS 1309-216	198.13146	-21.93981
3FGL J1312.7+4828	GB 1310+487	198.18064	48.475261
3FGL J1312.7-2349	NVSS J131248-235046	198.20312	-23.84631
3FGL J1312.8-0424	PKS B1310-041	198.21209	-4.41386
3FGL J1314.7-4237	MS 13121-4221	198.76413	-42.61378
3FGL J1314.8+2349	TXS 1312+240	198.68253	23.80744
3FGL J1315.1-5329	PMN J1315-5334	198.76767	-53.57667
3FGL J1315.4+1130	1RXS J131531.9+113327	198.88593	11.5588
3FGL J1316.0-3338	PKS 1313-333	199.03327	-33.64977
3FGL J1317.8+3429	S4 1315+34	199.40206	34.42109
3FGL J1318.7-1232	PMN J1318-1235	199.67854	-12.58439
3FGL J1319.3+1402	RX J1319.4+1405	199.88225	14.09254
3FGL J1319.6+7759	NVSS J131921+775823	199.83846	77.97306
3FGL J1321.0+2215	TXS 1318+225	200.29668	22.27003
3FGL J1321.7+8312	S5 1322+83	200.44004	83.27039
3FGL J1322.3+0839	NVSS J132210+084231	200.54243	8.70913
3FGL J1322.6-1619	PMN J1322-1617	200.68167	-16.29011
3FGL J1322.8-0938	PKS B1319-093	200.6538	-9.62717
3FGL J1322.9+0435	RBS 1257	200.75417	4.66427
3FGL J1323.0+2942	4C +29.48	200.75982	29.69285
3FGL J1323.9+1405	RX J1323.9+1406	200.99318	14.09984
3FGL J1325.4-4301	Cen A Core	201.365063	-43.019113
3FGL J1326.1+2931	TXS 1323+298	201.56228	29.55877
3FGL J1326.6-5256	PMN J1326-5256	201.70513	-52.93989
3FGL J1326.8+2211	B2 1324+22	201.75359	22.1806
3FGL J1327.9+2524	NVSS J132758+252750	201.99565	25.4629
3FGL J1328.5-4728	1WGA J1328.6-4727	202.16458	-47.45917
3FGL J1328.9-5607	PMN J1329-5608	202.25471	-56.13406
3FGL J1330.0-3818	Tol 1326-379	202.32996	-38.23847
3FGL J1330.0+4437	1RXS J133021.4+444117	202.58968	44.68898
3FGL J1330.1-7002	PKS 1326-697	202.54579	-70.05367
3FGL J1330.5+3023	3C 286	202.78453	30.50915

Table A.1: – continued from previous page

3FGL Name	Associated Name	RA	Dec
3FGL J1330.6+7002	NVSS J133025+700141	202.60712	70.02806
3FGL J1330.9+5201	87GB 132842.6+521750	202.6775	52.03763
3FGL J1331.1-1328	PMN J1331-1326	202.83462	-13.43503
3FGL J1331.5+1711	TXS 1329+174	202.88933	17.21408
3FGL J1331.8+4718	B3 1330+476	203.18853	47.37296
3FGL J1332.0-0508	PKS 1329-049	203.0186	-5.16203
3FGL J1332.6-1256	PMN J1332-1256	203.16355	-12.9376
3FGL J1332.8+2723	MG2 J133305+2725	203.28121	27.42177
3FGL J1333.7+5057	CLASS J1333+5057	203.4741	50.959976
3FGL J1335.4-2949	1ES 1332-295	203.87396	-29.84417
3FGL J1337.6-1257	PKS 1335-127	204.41576	-12.95686
3FGL J1338.6-2403	PKS 1336-237	204.75728	-24.02056
3FGL J1338.9+6532	87GB 133543.8+654752	204.31692	65.54619
3FGL J1339.0+1153	SDSS J133859.05+115316.7	204.74606	11.88798
3FGL J1339.8-0133	PKS 1337-013	205.01923	-1.6296
3FGL J1340.6+4412	RX J1340.4+4410	205.12416	44.16777
3FGL J1340.6-0408	NVSS J134042-041006	205.17508	-4.16856
3FGL J1341.0+3955	SDSS J134105.10+395945.4	205.27127	39.99595
3FGL J1341.5+5517	SBS 1339+554	205.40096	55.24388
3FGL J1341.9-2053	PKS B1339-206	205.51975	-20.85821
3FGL J1342.7+0945	NVSS J134240+094752	205.66676	9.79787
3FGL J1343.6+5753	1RXS J134357.9+575426	205.99008	57.91178
3FGL J1344.2-1724	PMN J1344-1723	206.06001	-17.39455
3FGL J1344.5-3655	PKS 1341-366	206.09871	-36.94122
3FGL J1345.6+4453	B3 1343+451	206.38822	44.88321
3FGL J1345.8+0704	TXS 1343+073	206.45548	7.10864
3FGL J1345.9-3357	NVSS J134543-335643	206.42937	-33.94536
3FGL J1346.6-6027	Cen B	206.70435	-60.40815
3FGL J1346.9-2958	NVSS J134706-295840	206.77842	-29.97803
3FGL J1347.6-3754	PMN J1347-3750	206.91845	-37.84351
3FGL J1349.6-1133	PKS 1346-112	207.38101	-11.54829
3FGL J1350.8+3035	B2 1348+30B	207.71973	30.58155
3FGL J1351.1+0030	PKS 1348+007	207.76846	0.52206
3FGL J1351.4+1115	RX J1351.3+1115	207.83686	11.24806
3FGL J1351.7-2913	PKS 1348-289	207.94516	-29.2049
3FGL J1353.1-4414	PKS 1349-439	208.23556	-44.21122
3FGL J1353.2+1435	OP 186	208.34517	14.59424
3FGL J1353.5-6640	1RXS J135341.1-664002	208.41729	-66.66599
3FGL J1354.5+3705	FIRST J135426.6+370654	208.61123	37.1152
3FGL J1355.0-1044	PKS 1352-104	208.69383	-10.68407
3FGL J1356.3-4029	SUMSS J135625-402820	209.10688	-40.47247
3FGL J1357.5+0125	RX J1357.6+0128	209.41126	1.47045
3FGL J1357.6+7643	S5 1357+76	209.48072	76.72251
3FGL J1359.0+5544	87GB 135720.6+555936	209.77393	55.7415
3FGL J1359.2+0204	PKS 1356+022	209.86312	1.99849
3FGL J1359.9-3746	PMN J1359-3746	209.95717	-37.76686
3FGL J1400.7-5605	PMN J1400-5605	210.17496	-56.08203
3FGL J1404.8+6554	NVSS J140450+655428	211.20656	65.90878
3FGL J1404.8+0401	MS 1402.3+0416	211.21208	4.03397
3FGL J1406.0-2508	NVSS J140609-250808	211.53996	-25.13561
3FGL J1406.6+1644	RBS 1350	211.74664	16.70219
3FGL J1407.7-4256	SUMSS J140739-430231	211.91587	-43.04217
3FGL J1408.8-0751	PKS B1406-076	212.23534	-7.87407
3FGL J1410.4+2821	RX J1410.4+2821	212.62316	28.34883
3FGL J1412.0+5249	SBS 1410+530	212.95602	52.81673
3FGL J1413.2-6518	Circinus galaxy	213.29146	-65.33922
3FGL J1415.0-1001	PKS B1412-096	213.83681	-9.93287
3FGL J1415.2+4832	RX J1415.5+4830	213.90334	48.50847
3FGL J1416.0+1325	PKS B1413+135	213.99507	13.33992
3FGL J1416.1-2417	NVSS J141612-241812	214.05071	-24.30347
3FGL J1417.8+2540	1E 1415.6+2557	214.48611	25.72395
3FGL J1418.4-0233	NVSS J141826-023336	214.60969	-2.55948
3FGL J1418.5+3543	NVSS J141828+354250	214.6191	35.71374
3FGL J1418.9+7731	1RXS J141901.8+773229	214.75196	77.54064
3FGL J1419.1-5156	PMN J1419-5155	214.89675	-51.91633
3FGL J1419.5-0836	NVSS J141922-083830	214.84379	-8.64189
3FGL J1419.5+0449	SDSS J141927.49+044513.7	214.86457	4.75383
3FGL J1419.8+3819	B3 1417+385	214.94422	38.36347
3FGL J1419.9+5425	OQ 530	214.94416	54.38744
3FGL J1421.0-1122	PMN J1420-1118	215.25067	-11.30589
3FGL J1422.4+3227	OQ 334	215.62658	32.38623
3FGL J1422.8-7828	PKS 1418-782	215.93146	-78.49303
3FGL J1422.8+5801	1ES 1421+582	215.66206	58.03208
3FGL J1424.3+0434	TXS 1421+048	216.03959	4.58113
3FGL J1424.6-6807	PKS 1420-679	216.23149	-68.1328
3FGL J1424.9+3615	FBQS J142455.5+361536	216.23132	36.26008
3FGL J1426.2+3402	RGB J1426+340	216.53216	34.07397
3FGL J1427.0+2347	PKS 1424+240	216.75163	23.80001
3FGL J1427.6-3305	PKS B1424-328	216.92234	-33.09208
3FGL J1427.8-3215	NVSS J142750-321515	216.95837	-32.25439
3FGL J1427.9-4206	PKS B1424-418	216.98457	-42.1054
3FGL J1428.5+4240	H 1426+428	217.13608	42.67239
3FGL J1434.1+4203	B3 1432+422	218.52373	42.05444
3FGL J1434.6+6640	1RXS J143442.0+664031	218.674	66.67408

Table A.1: – continued from previous page

3FGL Name	Associated Name	RA	Dec
3FGL J1434.6+1951	OQ 253	218.66581	19.86687
3FGL J1435.2+2023	TXS 1433+205	218.84143	20.35498
3FGL J1436.8+5639	RBS 1409	219.2405	56.6569
3FGL J1436.8+2322	PKS B1434+235	219.17075	23.35091
3FGL J1438.7+3710	B2 1436+37B	219.72338	37.1765
3FGL J1439.2+3931	PG 1437+398	219.8229	39.54524
3FGL J1440.0-3955	1RXS J143949.8-395524	219.96267	-39.92153
3FGL J1440.1+4955	GB6 J1439+4958	219.94573	49.96818
3FGL J1440.2-1538	PKS 1437-153	219.98697	-15.53071
3FGL J1440.4-3845	1RXS J144037.4-384658	220.15767	-38.78194
3FGL J1440.9+0610	PMN J1440+0610	220.22058	6.17117
3FGL J1442.0+4348	SDSS J144207.15+434836.6	220.5298	43.81019
3FGL J1442.6+5156	3C 303	220.76158	52.02702
3FGL J1442.8+1200	1ES 1440+122	220.70117	12.01119
3FGL J1443.9+2502	PKS 1441+25	220.98705	25.02903
3FGL J1444.0-3907	PKS 1440-389	220.98833	-39.14436
3FGL J1445.0-0328	RBS 1424	221.27625	-3.43678
3FGL J1446.1-1628	PKS B1443-162	221.4724	-16.48378
3FGL J1446.8-1831	NVSS J144644-182922	221.687	-18.48958
3FGL J1448.0+3608	RBS 1432	222.00241	36.14213
3FGL J1450.4+0911	TXS 1448+093	222.62987	9.17443
3FGL J1450.9+5200	SDSS J145059.99+520111.7	222.74996	52.01992
3FGL J1451.2+6355	RX J1451.4+6354	222.86667	63.90639
3FGL J1454.0+1622	CLASS J1454+1623	223.586896	16.40677
3FGL J1454.2-3751	PKS 1451-375	223.61421	-37.79254
3FGL J1454.5+5124	TXS 1452+516	223.61302	51.40937
3FGL J1457.4-3539	PKS 1454-354	224.3613	-35.65277
3FGL J1458.7+3719	B3 1456+375	224.68664	37.33934
3FGL J1500.6+4750	TXS 1459+480	225.20273	47.85432
3FGL J1500.9+2238	MS 1458.8+2249	225.25766	22.6351
3FGL J1503.7+4759	SDSS J150347.99+475930.9	225.94996	47.99192
3FGL J1503.7-6426	AT20G J150350-642539	225.95892	-64.42764
3FGL J1503.7-1540	RBS 1457	225.91917	-15.68719
3FGL J1504.4+1029	PKS 1502+106	226.10408	10.49422
3FGL J1504.5-8242	1RXS J150537.1-824233	226.40458	-82.70917
3FGL J1505.0-3432	PMN J1505-3432	226.25983	-34.54922
3FGL J1505.1+0326	PKS 1502+036	226.27699	3.44189
3FGL J1506.1+3728	B2 1504+37	226.53971	37.5142
3FGL J1506.3+4332	NVSS J150617+433413	226.57351	43.57046
3FGL J1506.6+0811	PMN J1506+0814	226.68529	8.23353
3FGL J1507.4+1725	NVSS J150716+172103	226.81841	17.3508
3FGL J1507.6-3710	NVSS J150720-370903	226.83671	-37.15083
3FGL J1508.6+2709	RBS 1467	227.17757	27.15216
3FGL J1508.7-4956	PMN J1508-4953	227.16242	-49.88383
3FGL J1509.7+5556	SBS 1508+561	227.44988	55.93814
3FGL J1509.9-2951	TXS 1507-296	227.53783	-29.85925
3FGL J1510.9-0542	PKS 1508-05	227.7233	-5.71873
3FGL J1511.8-0513	NVSS J151148-051345	227.95246	-5.22922
3FGL J1512.2+0202	PKS 1509+022	228.0656	2.05472
3FGL J1512.2-2255	1RXS J151213.1-225515	228.0535	-22.91897
3FGL J1512.3+8005	1RXS J151026.3+795946	227.63646	80.00147
3FGL J1512.8-0906	PKS 1510-08	228.21055	-9.09995
3FGL J1513.1-1014	PKS 1511-100	228.43706	-10.20007
3FGL J1513.5-3233	PKS 1510-324	228.41246	-32.58319
3FGL J1514.1+2940	MG2 J151421+2930	228.58704	29.51767
3FGL J1514.5-4750	PMN J1514-4748	228.66677	-47.80829
3FGL J1514.8+4446	NVSS J151436+445003	228.65265	44.83446
3FGL J1514.8-3623	PMN J1514-3617	228.67033	-36.28458
3FGL J1516.7+3648	MG2 J151646+3650	229.20525	36.83961
3FGL J1516.9+1926	PKS 1514+197	229.23665	19.53694
3FGL J1517.6+6524	1H 1515+660	229.44833	65.42331
3FGL J1517.6-2422	AP Librae	229.42422	-24.37208
3FGL J1518.0-2732	TXS 1515-273	229.51504	-27.52539
3FGL J1520.3+4209	B3 1518+423	230.16504	42.18581
3FGL J1520.8-0348	NVSS J152048-034850	230.20382	-3.81418
3FGL J1521.1+5543	SDSS J152034.98+554256.9	230.14576	55.71582
3FGL J1521.8+4340	B3 1520+437	230.45672	43.61091
3FGL J1522.1+3144	B2 1520+31	230.54163	31.73733
3FGL J1522.6-2730	PKS 1519-273	230.65698	-27.503
3FGL J1525.2-5905	PMN J1524-5903	231.213	-59.06097
3FGL J1531.0+5737	87GB 152947.5+574636	232.74245	57.60702
3FGL J1531.8+4704	1RXS J153148.8+465912	232.95615	46.98716
3FGL J1532.0+3018	RX J1531.9+3016	233.00934	30.27471
3FGL J1532.7-1319	TXS 1530-131	233.18906	-13.31947
3FGL J1533.2+1852	RX J1533.1+1854	233.29688	18.90809
3FGL J1533.5+3416	RX J1533.3+3416	233.35105	34.27784
3FGL J1534.4+5323	1ES 1533+535	233.75334	53.3437
3FGL J1534.5+0128	PKS 1532+01	233.71856	1.51784
3FGL J1535.0+3721	RGB J1534+372	233.69667	37.26523
3FGL J1535.7+3920	RX J1535.4+3922	233.87105	39.37964
3FGL J1536.6+8331	NVSS J153556+832614	233.98333	83.43736
3FGL J1537.8-8000	PMN J1537-7958	234.42054	-79.96767
3FGL J1539.5+2746	MG2 J153938+2744	234.91307	27.74395
3FGL J1539.8-1128	PMN J1539-1128	234.92171	-11.47647

Table A.1: – continued from previous page

3FGL Name	Associated Name	RA	Dec
3FGL J1540.1+8155	1ES 1544+820	235.066	81.91825
3FGL J1540.8+1449	4C +14.60	235.20621	14.79608
3FGL J1541.8+1105	MG1 J154207+1110	235.5025	11.177
3FGL J1542.9+6129	GB6 J1542+6129	235.73727	61.49871
3FGL J1543.5+0451	CGCG 050-083	235.89136	4.87204
3FGL J1546.0+0818	1RXS J154604.6+081912	236.51792	8.32056
3FGL J1546.6+1812	MG1 J154628+1817	236.59883	18.28759
3FGL J1547.1-2801	1RXS J154711.8-280222	236.79917	-28.03944
3FGL J1548.8-2250	PMN J1548-2251	237.20729	-22.85069
3FGL J1549.0+6309	SDSS J154958.45+631021.2	237.49358	63.17258
3FGL J1549.4+0237	PKS 1546+027	237.37265	2.61699
3FGL J1549.5+1709	MG1 J154930+1708	237.37196	17.14111
3FGL J1549.7-0658	NVSS J154952-065907	237.46821	-6.98553
3FGL J1550.3+7409	87GB 155014.9+741816	237.36575	74.15925
3FGL J1550.5+0526	4C +05.64	237.64695	5.4529
3FGL J1552.1+0852	TXS 1549+089	238.01359	8.84648
3FGL J1553.3-2421	PKS 1550-242	238.38183	-24.36842
3FGL J1553.5-3118	1RXS J155333.4-311841	238.39004	-31.30836
3FGL J1553.5+1256	PKS 1551+130	238.38624	12.9477
3FGL J1554.4+2010	1ES 1552+203	238.60054	20.1904
3FGL J1555.7+1111	PG 1553+113	238.92935	11.1901
3FGL J1557.4-7040	PKS 1552-705	239.40017	-70.67453
3FGL J1558.9-6432	PMN J1558-6432	239.70952	-64.54157
3FGL J1558.9+5625	TXS 1557+565	239.7012	56.42059
3FGL J1559.7+8512	WN B1609.6+8517	240.13046	85.16353
3FGL J1559.8-2525	NVSS J160005-252439	240.0225	-25.41094
3FGL J1559.9+2319	87GB 155744.0+232525	239.9675	23.28244
3FGL J1600.3-5810	MRC 1556-580	240.05079	-58.18356
3FGL J1603.7+1106	MG1 J160340+1106	240.92471	11.09685
3FGL J1603.9-4903	PMN J1603-4904	240.96113	-49.06808
3FGL J1604.4-4442	PMN J1604-4441	241.12925	-44.69221
3FGL J1604.6+5714	GB6 J1604+5714	241.15564	57.24352
3FGL J1606.1+5630	RBS 1558	241.58695	56.50486
3FGL J1607.0+1551	4C +15.54	241.77679	15.85958
3FGL J1607.9-2040	NVSS J160756-203942	241.98712	-20.66181
3FGL J1608.6+1029	4C +10.45	242.19251	10.48549
3FGL J1610.6-3956	PMN J1610-3958	242.59116	-39.98287
3FGL J1610.8-6649	PMN J1610-6649	242.69358	-66.817
3FGL J1612.4-3100	NVSS J161219-305937	243.08312	-30.99383
3FGL J1613.8+3410	OS 319	243.4211	34.21331
3FGL J1615.8+4712	TXS 1614+473	243.92167	47.1866
3FGL J1616.4+4631	MG4 J161600+4632	244.01569	46.54034
3FGL J1616.8+4111	B3 1615+412	244.27633	41.11306
3FGL J1617.3-2519	PMN J1617-2537	244.33571	-25.62325
3FGL J1617.4-5846	MRC 1613-586	244.32483	-58.80169
3FGL J1617.7-7717	PKS 1610-77	244.45532	-77.28846
3FGL J1617.8+5137	TXS 1616+517	244.37277	51.67234
3FGL J1621.1-2331	PKS 1617-235	245.25065	-23.70076
3FGL J1625.0+5651	SBS 1623+569	246.13408	56.87444
3FGL J1625.7-2527	PKS 1622-253	246.44538	-25.46065
3FGL J1625.9+4125	B3 1624+414	246.42567	41.31142
3FGL J1626.0-2951	PKS 1622-29	246.52509	-29.85749
3FGL J1626.1+3512	RGB J1626+352	246.60772	35.22821
3FGL J1626.4-7640	PKS 1619-765	246.65879	-76.64872
3FGL J1628.2+7703	6C B163030.4+771303	247.13791	77.11391
3FGL J1630.6+8232	NGC 6251	248.13321	82.53789
3FGL J1630.7+5222	TXS 1629+524	247.67967	52.36069
3FGL J1630.8+1047	MG1 J163119+1051	247.82824	10.86735
3FGL J1635.2+3809	4C +38.41	248.81455	38.13458
3FGL J1636.7+2624	NVSS J163651+262657	249.21425	26.44922
3FGL J1637.1+1314	1RXS J163717.1+131418	249.31973	13.24411
3FGL J1637.6-3449	NVSS J163750-344915	249.46246	-34.82094
3FGL J1637.7+4715	4C +47.44	249.43804	47.29273
3FGL J1637.8+7325	RX J1637.9+7326	249.50692	73.43767
3FGL J1637.9+5719	OS 562	249.55607	57.33999
3FGL J1639.8+4125	MG4 J163918+4127	249.81587	41.47603
3FGL J1640.6+3945	NRAO 512	250.12347	39.77945
3FGL J1640.9+1142	TXS 1638+118	250.24539	11.7345
3FGL J1641.8-0619	TXS 1639-062	250.50907	-6.35658
3FGL J1642.9+3950	3C 345	250.74504	39.81028
3FGL J1643.6-0642	NVSS J164328-064619	250.87062	-6.77219
3FGL J1645.2-5747	AT20G J164513-575122	251.30596	-57.85622
3FGL J1645.9+6336	TXS 1645+635	251.49397	63.50303
3FGL J1647.1-6438	PMN J1647-6437	251.90725	-64.63339
3FGL J1647.4+4950	SBS 1646+499	251.89547	49.8335
3FGL J1648.5-4829	PMN J1648-4826	252.19967	-48.43856
3FGL J1649.4+5238	87GB 164812.2+524023	252.35387	52.58728
3FGL J1650.2-5044	PMN J1650-5044	252.56871	-50.74617
3FGL J1650.8+0830	MG1 J165034+0824	252.65651	8.41451
3FGL J1651.6+7219	RX J1651.6+7218	252.91646	72.307
3FGL J1653.9+3945	Mkn 501	253.46757	39.76017
3FGL J1656.0+2044	MG2 J165546+2043	253.944	20.75653
3FGL J1656.2-3303	Swift J1656.3-3302	254.07022	-33.03641
3FGL J1656.8-2010	1RXS J165655.0-201049	254.22917	-20.18028

Table A.1: – continued from previous page

3FGL Name	Associated Name	RA	Dec
3FGL J1656.9+6008	87GB 165604.4+601702	254.20102	60.20457
3FGL J1657.7+4807	4C +48.41	254.44533	48.14251
3FGL J1658.3+6149	NVSS J165808+615001	254.53471	61.83394
3FGL J1659.4+2631	4C +26.51	254.85062	26.49359
3FGL J1659.7-3132	NVSS J165949-313047	254.95433	-31.51317
3FGL J1700.1+6829	TXS 1700+685	255.03872	68.50193
3FGL J1702.6+3116	RX J1702.6+3115	255.66061	31.26209
3FGL J1703.6-6211	MRC 1659-621	255.90225	-62.21111
3FGL J1704.0+7646	NVSS J170357+764611	255.99058	76.76978
3FGL J1705.5+7134	GB6 J1704+7138	256.19582	71.63823
3FGL J1709.6+4318	B3 1708+433	257.42119	43.31237
3FGL J1711.5-5029	PMN J1711-5028	257.92792	-50.47083
3FGL J1711.6+8846	1RXS J171643.8+884414	259.1825	88.73722
3FGL J1712.6+2932	RX J1712.8+2931	258.20323	29.52136
3FGL J1714.1-2029	1RXS J171405.2-202747	258.52167	-20.46306
3FGL J1715.7+6837	S4 1716+68	259.05807	68.61076
3FGL J1716.7-8112	1RXS J171712.6-811501	259.29804	-81.25267
3FGL J1717.4-5157	PMN J1717-5155	259.39667	-51.92222
3FGL J1717.8-3342	TXS 1714-336	259.40012	-33.70245
3FGL J1718.1-3056	PMN J1718-3056	259.52171	-30.9375
3FGL J1719.2+1744	PKS 1717+177	259.80437	17.75179
3FGL J1719.3+1206	1RXS J171921.2+120711	259.8395	12.12286
3FGL J1722.7+6104	GB6 J1722+6105	260.66691	61.09994
3FGL J1722.7+1014	TXS 1720+102	260.68576	10.2266
3FGL J1723.5-5609	PMN J1723-5614	260.93513	-56.24872
3FGL J1723.7-7713	PKS 1716-771	260.96146	-77.23092
3FGL J1723.9+4004	S4 1722+40	261.02262	40.07679
3FGL J1725.0+1152	1H 1720+117	261.26809	11.87096
3FGL J1725.3+5853	7C 1724+5854	261.39588	58.86107
3FGL J1727.1+4531	S4 1726+45	261.86521	45.51104
3FGL J1728.0+1217	PKS 1725+123	262.02938	12.26097
3FGL J1728.3+5013	I Zw 187	262.0776	50.21957
3FGL J1728.5+0428	PKS 1725+044	262.103958	4.451361
3FGL J1730.5+0023	PKS 1728+004	262.64583	0.41075
3FGL J1730.6+3711	GB6 J1730+3714	262.69602	37.24865
3FGL J1731.8-3001	NVSS J173146-300309	262.94538	-30.05253
3FGL J1733.0-1305	PKS 1730-13	263.26127	-13.08043
3FGL J1734.3+3858	B2 1732+38A	263.58574	38.96429
3FGL J1735.4-1118	PMN J1735-1117	263.86325	-11.29292
3FGL J1736.0+2033	NVSS J173605+203301	264.02217	20.55028
3FGL J1736.4+0634	MG1 J173624+0632	264.11912	6.52986
3FGL J1739.0+8716	NVSS J173722+871744	264.34567	87.29575
3FGL J1739.4+4955	S4 1738+49	264.86413	49.9176
3FGL J1740.3+5211	4C +51.37	265.15407	52.19539
3FGL J1740.3+4736	S4 1738+47	264.98804	47.63288
3FGL J1740.4+5347	NVSS J174036+534623	265.15219	53.77329
3FGL J1741.9-2539	NVSS J174154-253743	265.47725	-25.62881
3FGL J1742.2+5947	RGB J1742+597	265.63333	59.75187
3FGL J1743.9+1934	S3 1741+19	265.99097	19.58584
3FGL J1744.3-0353	PKS 1741-03	265.99523	-3.83462
3FGL J1744.9-1725	1RXS J174459.5-172640	266.24792	-17.44444
3FGL J1745.4-0754	TXS 1742-078	266.36294	-7.88443
3FGL J1745.7+3952	B2 1743+39C	266.40729	39.85858
3FGL J1747.1+0139	PMN J1746+0141	266.71879	1.69656
3FGL J1748.0+3405	MG2 J174803+3403	267.02425	34.06699
3FGL J1748.6+7005	S4 1749+70	267.13683	70.09744
3FGL J1749.1+4322	B3 1747+433	267.2515	43.36425
3FGL J1751.5+0939	OT 081	267.88674	9.6502
3FGL J1753.5-5010	PMN J1753-5015	268.41046	-50.25414
3FGL J1754.1+3212	RX J1754.1+3212	268.54915	32.2064
3FGL J1754.3-6424	PMN J1754-6423	268.6745	-64.39639
3FGL J1756.3+5523	1RXS J175615.5+552217	269.06618	55.3717
3FGL J1756.9+7032	MS 1757.7+7034	269.30417	70.56044
3FGL J1757.1+1533	87GB 175437.6+153548	269.22126	15.58912
3FGL J1757.4+6536	7C 1757+6536	269.3495	65.60281
3FGL J1759.1-4822	PMN J1758-4820	269.74354	-48.35344
3FGL J1800.5+7827	S5 1803+784	270.19035	78.46778
3FGL J1801.5+4403	S4 1800+44	270.38465	44.07275
3FGL J1802.6-3940	PMN J1802-3940	270.67783	-39.66886
3FGL J1804.1+0341	TXS 1801+036	270.98451	3.68544
3FGL J1806.7+6949	3C 371	271.71117	69.82447
3FGL J1807.8+6427	7C 1807+6428	271.88404	64.49081
3FGL J1807.8-5011	PMN J1808-5011	272.058	-50.19819
3FGL J1808.0+4652	RGB J1808+468	272.005	46.82806
3FGL J1809.4+2040	RX J1809.3+2041	272.35608	20.69194
3FGL J1809.7+2909	MG2 J180948+2910	272.439125	29.172194
3FGL J1810.8+1609	87GB 180835.5+160714	272.70904	16.13894
3FGL J1811.2+0340	NVSS J181118+034113	272.82538	3.68719
3FGL J1813.6+0614	TXS 1811+062	273.38921	6.26168
3FGL J1813.6+3143	B2 1811+31	273.39668	31.73823
3FGL J1816.9-4944	PMN J1816-4943	274.23371	-49.729
3FGL J1818.6+0903	MG1 J181841+0903	274.66692	9.06283
3FGL J1819.1+2134	MG2 J181902+2132	274.77167	21.54308
3FGL J1819.1+4259	NVSS J181927+425800	274.86613	42.96681

Table A.1: – continued from previous page

3FGL Name	Associated Name	RA	Dec
3FGL J1820.3+3625	NVSS J182021+362343	275.08787	36.39536
3FGL J1822.1-7051	PMN J1823-7056	275.87358	-70.93325
3FGL J1823.4+6857	7C 1823+6856	275.88689	68.96461
3FGL J1823.6-3453	NVSS J182338-345412	275.91087	-34.90333
3FGL J1824.2+5649	4C +56.27	276.02945	56.85041
3FGL J1824.4+4310	1RXS J182418.7+430954	276.07208	43.16556
3FGL J1825.2-5230	PKS 1821-525	276.30783	-52.51629
3FGL J1828.9-2417	1RXS J182853.8-241746	277.22417	-24.29611
3FGL J1829.4+5402	1RXS J182925.7+540255	277.35117	54.04986
3FGL J1829.6+4844	3C 380	277.38242	48.74616
3FGL J1829.8+1328	MG1 J183001+1323	277.50313	13.40392
3FGL J1830.0-4439	PMN J1830-4441	277.5035	-44.68658
3FGL J1830.1+0617	TXS 1827+062	277.52475	6.3211
3FGL J1831.0-2714	PMN J1831-2714	277.75025	-27.23556
3FGL J1832.4-5659	PMN J1832-5659	278.12908	-56.98906
3FGL J1833.6-2103	PKS 1830-211	278.4162	-21.06105
3FGL J1835.4+1349	TXS 1833+137	278.89717	13.81481
3FGL J1836.3+3137	RX J1836.2+3136	279.08846	31.60744
3FGL J1838.5-6006	SUMSS J183806-600033	279.52883	-60.00939
3FGL J1838.8+4802	GB6 J1838+4802	279.70482	48.04285
3FGL J1841.2+2910	MG3 J184126+2910	280.34054	29.16139
3FGL J1841.7+3218	RX J1841.7+3218	280.446	32.31086
3FGL J1842.3-5841	1RXS J184230.6-584202	280.6275	-58.70056
3FGL J1842.8+6810	S4 1842+68	280.64017	68.15701
3FGL J1844.1+5709	TXS 1843+571	281.21329	57.16128
3FGL J1844.3+1547	NVSS J184425+154646	281.10558	15.77944
3FGL J1848.1-4230	PMN J1848-4230	282.02575	-42.50739
3FGL J1848.4+3216	B2 1846+32A	282.09204	32.31739
3FGL J1848.9+4247	RGB J1848+427	282.19642	42.76094
3FGL J1849.2+6705	S4 1849+67	282.31697	67.09491
3FGL J1849.3-1645	1RXS J184919.7-164726	282.316	-16.7797
3FGL J1849.4-4312	PMN J1849-4314	282.35796	-43.23703
3FGL J1849.5+2751	MG2 J184929+2748	282.38204	27.80028
3FGL J1852.4+4856	S4 1851+48	283.11895	48.92986
3FGL J1855.1-6008	PMN J1854-6009	283.71542	-60.15647
3FGL J1858.4-2509	PMN J1858-2511	284.57958	-25.18069
3FGL J1902.9-6745	PMN J1903-6749	285.75513	-67.82656
3FGL J1903.2+5541	TXS 1902+556	285.79836	55.67735
3FGL J1904.5+3627	MG2 J190411+3627	286.04945	36.44971
3FGL J1908.8-0130	NVSS J190836-012642	287.15375	-1.44525
3FGL J1910.8+2855	1RXS J191053.2+285622	287.72167	28.93944
3FGL J1911.2-2006	PKS B1908-201	287.79022	-20.11531
3FGL J1911.4-1908	PMN J1911-1908	287.87392	-19.13986
3FGL J1912.0-0804	PMN J1912-0804	288.02975	-8.07261
3FGL J1912.6-1223	TXS 1909-124	288.12308	-12.38372
3FGL J1912.9-8008	PKS 1903-80	288.16675	-80.16832
3FGL J1913.5-3631	PMN J1913-3630	288.33696	-36.50553
3FGL J1913.9+4441	1RXS J191401.9+443849	288.50792	44.64694
3FGL J1917.7-1921	1H 1914-194	289.43675	-19.35878
3FGL J1918.0+3750	1RXS J191810.2+375315	289.5425	37.8875
3FGL J1918.2-4110	PMN J1918-4111	289.56738	-41.19192
3FGL J1921.2-1232	TXS 1918-126	290.34971	-12.53175
3FGL J1921.9-1607	PMN J1921-1607	290.46467	-16.12014
3FGL J1923.5-2104	TXS 1920-211	290.88412	-21.07593
3FGL J1924.8-2914	PKS B1921-293	291.21273	-29.2417
3FGL J1924.9+2817	NVSS J192502+281542	291.25942	28.26175
3FGL J1925.7+1228	TXS 1923+123	291.42007	12.46058
3FGL J1926.8+6154	1RXS J192649.5+615445	291.70788	61.91176
3FGL J1927.7+6118	S4 1926+61	291.87684	61.29247
3FGL J1931.1+0937	RX J1931.1+0937	292.78846	9.62122
3FGL J1932.6-4537	PKS 1929-457	293.18703	-45.61054
3FGL J1933.4+0727	1RXS J193320.3+072616	293.33458	7.43778
3FGL J1935.5+8355	6C B194425+834912	294.41524	83.94138
3FGL J1936.9-4719	PMN J1936-4719	294.23379	-47.33061
3FGL J1937.0-3956	PKS 1933-400	294.31757	-39.9671
3FGL J1939.6-4925	SUMSS J193946-492539	294.94238	-49.42769
3FGL J1941.2-6210	PKS 1936-623	295.3407	-62.18918
3FGL J1941.8+7218	87GB 194202.1+721428	295.36243	72.36173
3FGL J1942.7+1033	1RXS J194246.3+103339	295.69783	10.55772
3FGL J1943.2-3510	1RXS J194306.8-351001	295.77804	-35.16856
3FGL J1944.1-4523	1RXS J194422.6-452326	296.09417	-45.39056
3FGL J1945.9-3115	PKS 1942-313	296.49733	-31.19397
3FGL J1949.0+1312	87GB 194635.4+130713	297.23037	13.24403
3FGL J1949.4-6140	PMN J1949-6137	297.3975	-61.62731
3FGL J1954.8-1122	TXS 1951-115	298.67148	-11.38962
3FGL J1954.9-5640	1RXS J195503.1-564031	298.76292	-56.67528
3FGL J1955.0-1605	1RXS J195500.6-160328	298.752417	-16.060528
3FGL J1955.1+1357	87GB 195252.4+135009	298.79821	13.97118
3FGL J1955.9+0212	NVSS J195547+021514	298.94958	2.25392
3FGL J1957.0-3234	PKS 1953-325	299.24773	-32.42945
3FGL J1958.0-3847	PKS 1954-388	299.49925	-38.75177
3FGL J1958.2-3011	1RXS J195815.6-301119	299.56208	-30.18661
3FGL J1959.1-4245	PMN J1959-4246	299.80579	-42.76886
3FGL J1959.8-4725	SUMSS J195945-472519	299.9395	-47.42203

Table A.1: – continued from previous page

3FGL Name	Associated Name	RA	Dec
3FGL J2000.0+6509	1ES 1959+650	299.99938	65.14851
3FGL J2000.1+4212	MG4 J195957+4213	299.99479	42.22978
3FGL J2000.4-2926	PMN J2000-2931	300.07075	-29.50725
3FGL J2001.0-1750	PKS 1958-179	300.23788	-17.81602
3FGL J2001.1+4352	MG4 J200112+4352	300.30364	43.88134
3FGL J2001.8+7041	TXS 2001+705	300.39192	70.67372
3FGL J2002.7+6303	1RXS J200245.4+630226	300.68833	63.04056
3FGL J2004.8+7003	1RXS J200504.0+700445	301.26583	70.08139
3FGL J2005.2+7752	S5 2007+77	301.37887	77.87865
3FGL J2006.0-2311	TXS 2002-233	301.48588	-23.17444
3FGL J2007.3+6605	TXS 2007+659	301.86988	66.12293
3FGL J2007.7-7728	PKS 2000-776	301.77733	-77.51183
3FGL J2007.8-4429	PKS 2004-447	301.97993	-44.57896
3FGL J2009.3-4849	PKS 2005-489	302.35579	-48.83159
3FGL J2010.3+7228	4C +72.28	302.46793	72.48871
3FGL J2012.0+4629	7C 2010+4619	303.02349	46.48216
3FGL J2012.1-1643	PMN J2012-1646	303.12563	-16.78072
3FGL J2014.3-0047	PMN J2014-0047	303.61917	-0.79
3FGL J2014.5+0648	NVSS J201431+064849	303.62971	6.81386
3FGL J2014.9+1623	4C +16.67	303.76629	16.37431
3FGL J2015.2-0138	PKS 2012-017	303.81316	-1.62571
3FGL J2015.6+3709	MG2 J201534+3710	303.86971	37.1832
3FGL J2016.4-0905	PMN J2016-0903	304.1	-9.05908
3FGL J2017.6-4110	1RXS J201731.2-411452	304.38	-41.24778
3FGL J2018.5+3851	TXS 2016+386	304.62946	38.85561
3FGL J2021.9+0630	87GB 201926.8+061922	305.48125	6.48725
3FGL J2022.2-4515	PMN J2022-4513	305.61	-45.22486
3FGL J2022.5+7612	S5 2023+760	305.64823	76.1906
3FGL J2023.2+3154	4C +31.56	305.82924	31.88397
3FGL J2023.6-1139	PMN J2023-1140	305.90292	-11.66619
3FGL J2024.4-0848	1RXS J202428.9-084810	306.12246	-8.80125
3FGL J2024.4-3254	PKS 2021-330	306.14823	-32.89331
3FGL J2025.2+3340	B2 2023+33	306.29517	33.71673
3FGL J2025.6-0736	PKS 2023-07	306.41942	-7.59797
3FGL J2026.3+7644	1RXS J202633.4+764432	306.63917	76.74222
3FGL J2029.4+4923	MG4 J202932+4925	307.41614	49.43948
3FGL J2030.2-0622	TXS 2027-065	307.56325	-6.37081
3FGL J2031.0+1937	RX J2030.8+1935	307.73805	19.60359
3FGL J2031.8+1223	PKS 2029+121	307.97914	12.32815
3FGL J2033.6+6309	87GB 203249.5+625814	308.4155	63.14444
3FGL J2034.3+1155	TXS 2032+117	308.65462	11.90872
3FGL J2035.3+1055	PKS 2032+107	308.84306	10.93522
3FGL J2036.4+6551	87GB 203539.4+654245	309.08321	65.88744
3FGL J2036.6-3325	1RXS J203650.9-332817	309.20621	-33.47522
3FGL J2036.8-2830	PMN J2036-2830	309.15738	-28.50736
3FGL J2038.8+5113	3C 418	309.65431	51.32018
3FGL J2039.0-1047	TXS 2036-109	309.75296	-10.77825
3FGL J2039.5+5217	1ES 2037+521	309.84804	52.33058
3FGL J2040.0-5734	PKS 2036-577	310.00467	-57.58594
3FGL J2040.2-7115	PKS 2035-714	310.03421	-71.25
3FGL J2041.7-3732	NVSS J204150-373341	310.4605	-37.56147
3FGL J2041.9-7318	SUMSS J204201-731911	310.5065	-73.31975
3FGL J2042.1+2428	MG2 J204208+2426	310.52521	24.44787
3FGL J2046.7-1011	PMN J2046-1010	311.72721	-10.17842
3FGL J2049.0-6801	PKS 2043-682	312.10042	-68.08078
3FGL J2049.7+1002	PKS 2047+098	312.4411	10.054
3FGL J2050.2+0409	PKS 2047+039	312.526	4.130222
3FGL J2051.8-5535	PMN J2052-5533	313.057	-55.55278
3FGL J2055.0+0016	RGB J2054+002	313.73687	0.2605
3FGL J2055.2-0019	1RXS J205528.2-002123	313.86762	-0.35476
3FGL J2056.2-4714	PKS 2052-47	314.06817	-47.24656
3FGL J2056.7+4938	RGB J2056+496	314.17804	49.6685
3FGL J2103.9-6233	PMN J2103-6232	315.90975	-62.54072
3FGL J2103.9-3546	NVSS J210353-354620	315.97112	-35.77222
3FGL J2104.2-0211	NVSS J210421-021239	316.09125	-2.21069
3FGL J2106.1+2505	MG3 J210642+2501	316.66566	25.01598
3FGL J2107.7-4822	PMN J2107-4827	316.93575	-48.46739
3FGL J2108.0+3654	TXS 2106+367	317.02263	36.92411
3FGL J2108.6-8619	1RXS J210959.5-861853	317.54138	-86.31136
3FGL J2108.6-0250	TXS 2106-030	317.18637	-2.84286
3FGL J2109.1-6638	PKS 2104-668	317.21592	-66.62306
3FGL J2110.0+0812	PMN J2110+0810	317.54033	8.16539
3FGL J2110.3-1013	PKS 2107-105	317.50408	-10.34926
3FGL J2110.3+3540	B2 2107+35A	317.38283	35.54933
3FGL J2112.7+0819	1RXS J211242.5+081831	318.17908	8.31006
3FGL J2114.7+3130	B2 2112+31	318.71025	31.50588
3FGL J2115.4+2933	B2 2113+29	318.87256	29.56066
3FGL J2116.1+3339	B2 2114+33	319.0605	33.65567
3FGL J2118.0-3241	NVSS J211754-324326	319.47888	-32.7245
3FGL J2118.4+0013	PMN J2118+0013	319.5725	0.22133
3FGL J2119.2-3313	PMN J2118-3316	319.722	-33.28133
3FGL J2121.0+1901	OX 131	320.25253	19.02452
3FGL J2123.6+0533	OX 036	320.93549	5.58947
3FGL J2126.5-4605	PKS 2123-463	321.62793	-46.09664

Table A.1: – continued from previous page

3FGL Name	Associated Name	RA	Dec
3FGL J2126.5-3926	PMN J2126-3921	321.60504	-39.35639
3FGL J2127.7+3612	B2 2125+35	321.92929	36.21828
3FGL J2130.8-2745	RBS 1751	322.76358	-27.78286
3FGL J2131.5-0915	RBS 1752	322.8975	-9.25653
3FGL J2131.8-2516	RBS 1755	322.96375	-25.26633
3FGL J2132.4-5420	PMN J2132-5420	323.03479	-54.34378
3FGL J2133.3+2533	87GB 213100.1+251534	323.30983	25.48283
3FGL J2133.8+6648	NVSS J213349+664706	323.45463	66.78522
3FGL J2134.1-0152	PKS 2131-021	323.54296	-1.88812
3FGL J2135.3-5008	PMN J2135-5006	323.83371	-50.11394
3FGL J2139.4-4235	MH 2136-428	324.85067	-42.589
3FGL J2141.6-6412	PMN J2141-6411	325.44346	-64.18733
3FGL J2141.7-3734	PKS 2138-377	325.46854	-37.48694
3FGL J2142.2-2546	PMN J2142-2551	325.56636	-25.85739
3FGL J2143.1-3928	PMN J2143-3929	325.76192	-39.49022
3FGL J2143.5+1744	OX 169	325.8981	17.73022
3FGL J2144.2+3132	MG3 J214415+3132	326.06346	31.56089
3FGL J2144.9-3356	PMN J2145-3357	326.25471	-33.95456
3FGL J2145.7+0717	MS 2143.4+0704	326.46792	7.32422
3FGL J2146.6-1344	NVSS J214637-134359	326.654	-13.73353
3FGL J2146.7-1527	PKS 2143-156	326.59575	-15.42886
3FGL J2147.2+0929	PKS 2144+092	326.79235	9.4963
3FGL J2147.3-7536	PKS 2142-75	326.80304	-75.60367
3FGL J2149.6+1915	TXS 2147+191	327.44713	19.34617
3FGL J2149.7+0323	PKS B2147+031	327.42446	3.38095
3FGL J2150.2-1411	TXS 2147-144	327.56458	-14.18056
3FGL J2151.6-2744	PMN J2151-2742	327.84127	-27.70639
3FGL J2151.8-3025	PKS 2149-306	327.98135	-30.46492
3FGL J2152.4+1735	S3 2150+17	328.10341	17.57717
3FGL J2152.9-0045	RBS 1792	328.27229	-0.70855
3FGL J2154.0-1137	PMN J2153-1136	328.45933	-11.60392
3FGL J2156.0+1818	RX J2156.0+1818	329.00679	18.31089
3FGL J2156.9-0855	FIRST J215650.3-085535	329.2097	-8.9265
3FGL J2157.5+3126	B2 2155+31	329.3701	31.45038
3FGL J2158.0-1501	PKS 2155-152	329.52617	-15.01926
3FGL J2158.8-3013	PKS 2155-304	329.71694	-30.22559
3FGL J2159.2-2841	NVSS J215910-284115	329.79567	-28.68817
3FGL J2159.8+1025	TXS 2157+102	330.03305	10.50217
3FGL J2200.2+2139	TXS 2157+213	330.05916	21.63251
3FGL J2200.9-2412	NVSS J220036-241428	330.15279	-24.24131
3FGL J2201.7+5047	NRAO 676	330.4314	50.81566
3FGL J2202.4-8339	PKS 2155-83	330.58221	-83.63653
3FGL J2202.7+4217	BL Lacertae	330.68038	42.27777
3FGL J2203.4+1725	PKS 2201+171	330.86206	17.43007
3FGL J2203.7+3143	4C +31.63	330.8124	31.76063
3FGL J2204.4+0439	4C +04.77	331.07358	4.66722
3FGL J2206.9-0031	PMN J2206-0031	331.68038	-0.51736
3FGL J2207.8-5345	PKS 2204-54	331.93222	-53.77606
3FGL J2212.0+2355	PKS 2209+236	333.02486	23.92793
3FGL J2212.3-7039	PMN J2211-7039	332.98454	-70.65372
3FGL J2212.6+2801	MG3 J221240+2759	333.16293	27.99401
3FGL J2213.1-2532	PKS 2210-25	333.26041	-25.49169
3FGL J2213.6-4755	SUMSS J221330-475426	333.37662	-47.90722
3FGL J2217.0+2421	B2 2214+24B	334.25342	24.36277
3FGL J2219.2+1806	MG1 J221916+1806	334.80872	18.10988
3FGL J2220.3+2812	RX J2220.4+2814	335.119708	28.232111
3FGL J2221.6-5225	PMN J2221-5224	335.37208	-52.42436
3FGL J2222.3-3500	PKS 2220-351	335.77471	-34.92977
3FGL J2224.6-1122	PKS 2221-116	336.03318	-11.4392
3FGL J2225.8-0454	3C 446	336.44691	-4.95039
3FGL J2227.8+0040	PMN J2227+0037	336.99221	0.61817
3FGL J2229.7-0833	PKS 2227-08	337.41702	-8.54845
3FGL J2230.5-7817	PKS 2225-785	337.62804	-78.26564
3FGL J2230.6-4419	PKS 2227-445	337.735	-44.275
3FGL J2232.5+1143	CTA 102	338.1517	11.73081
3FGL J2232.9-2021	1RXS J223249.5-202232	338.20167	-20.37278
3FGL J2233.5-1235	PKS 2231-127	338.47896	-12.50964
3FGL J2234.1-2655	PMN J2234-2656	338.53483	-26.94467
3FGL J2235.3-4835	PKS 2232-488	338.80515	-48.59967
3FGL J2235.6-2319	PMN J2236-2309	339.10933	-23.15739
3FGL J2236.0-3629	NVSS J223554-362901	338.97833	-36.48369
3FGL J2236.0-1706	PKS 2233-173	339.03968	-17.1061
3FGL J2236.2-5049	SUMSS J223605-505521	339.02425	-50.92261
3FGL J2236.3+2829	B2 2234+28A	339.09363	28.48261
3FGL J2236.5-1432	PKS 2233-148	339.14203	-14.55616
3FGL J2237.1-3921	NVSS J223708-392137	339.28379	-39.36061
3FGL J2240.9+4121	B3 2238+410	340.28002	41.33656
3FGL J2243.2-3933	NVSS J224326-393353	340.85892	-39.56467
3FGL J2243.4-2541	PKS 2240-260	340.86004	-25.74186
3FGL J2243.6-1230	RBS 1888	340.92292	-12.52722
3FGL J2243.9+2021	RGB J2243+203	340.9781	20.35105
3FGL J2244.1+4057	TXS 2241+406	341.05305	40.95378
3FGL J2246.2+1547	NVSS J224604+154437	341.52079	15.74375
3FGL J2246.7-5205	RBS 1895	341.67537	-52.11128

Table A.1: – continued from previous page

3FGL Name	Associated Name	RA	Dec
3FGL J2247.8+4413	NVSS J224753+441317	341.97171	44.22092
3FGL J2248.6-3235	PKS 2245-328	342.16119	-32.59783
3FGL J2250.1+3825	B3 2247+381	342.52395	38.41033
3FGL J2250.3-4206	PMN J2250-4206	342.59258	-42.10372
3FGL J2250.7-2806	PMN J2250-2806	342.68538	-28.11093
3FGL J2251.5-4928	SUMSS J225128-492912	342.86975	-49.48667
3FGL J2251.9+4031	MG4 J225201+4030	342.99905	40.51615
3FGL J2254.0+1608	3C 454.3	343.49062	16.14821
3FGL J2254.0+1403	NVSS J225354+140439	343.47598	14.07691
3FGL J2255.1+2411	MG3 J225517+2409	343.81392	24.17014
3FGL J2256.7-2011	PKS 2254-204	344.1717	-20.19459
3FGL J2258.0-2759	PKS 2255-282	344.52488	-27.97256
3FGL J2258.1-8248	PMN J2258-8246	344.49733	-82.78075
3FGL J2258.3-5526	PMN J2258-5526	344.57892	-55.42708
3FGL J2300.3+3136	NVSS J230022+313703	345.09514	31.61788
3FGL J2304.6+3704	1RXS J230437.1+370506	346.15333	37.08528
3FGL J2305.3-4219	SUMSS J230512-421859	346.30154	-42.31639
3FGL J2307.4-1208	1RXS J230722.5-120520	346.84167	-12.08836
3FGL J2307.7+1449	MG1 J230734+1449	346.89153	14.83845
3FGL J2309.6-3633	1RXS J230940.6-363241	347.421	-36.54386
3FGL J2311.0+3425	B2 2308+34	347.7722	34.4197
3FGL J2311.0+0204	NVSS J231101+020504	347.75546	2.08492
3FGL J2312.9-6923	SUMSS J231347-692332	348.44592	-69.39242
3FGL J2314.0+1443	RGB J2313+147	348.48891	14.73978
3FGL J2315.7-5018	PKS 2312-505	348.93458	-50.31108
3FGL J2316.8-5209	SUMSS J231701-521003	349.257	-52.16772
3FGL J2317.3-4534	1RXS J231733.0-453348	349.38125	-45.5675
3FGL J2318.6+1912	TXS 2315+189	349.59546	19.24783
3FGL J2319.2-4207	PKS 2316-423	349.77463	-42.11339
3FGL J2321.2-6439	PMN J2321-6438	350.42592	-64.63528
3FGL J2321.6+4438	B3 2319+444	350.58483	44.76176
3FGL J2321.9+2732	4C +27.50	350.49943	27.54623
3FGL J2321.9+3204	B2 2319+31	350.47898	32.06878
3FGL J2322.5+3436	TXS 2320+343	350.68338	34.60386
3FGL J2322.9-4917	SUMSS J232254-491629	350.72629	-49.27497
3FGL J2323.5-0315	PKS 2320-035	350.88314	-3.28473
3FGL J2323.9+4211	1ES 2321+419	350.96696	42.18299
3FGL J2324.7-4040	1ES 2322-409	351.18613	-40.68036
3FGL J2324.7+0801	PMN J2324+0801	351.18879	8.03497
3FGL J2325.2+3957	B3 2322+396	351.32446	39.96014
3FGL J2325.3-3557	CTS 0490	351.36917	-35.965
3FGL J2325.4-4758	PKS 2322-482	351.36196	-48.00472
3FGL J2325.6+1650	NVSS J232538+164641	351.40882	16.77853
3FGL J2327.7+0941	PKS 2325+093	351.88992	9.66929
3FGL J2328.4-4034	PKS 2325-408	352.08021	-40.586
3FGL J2329.2+3754	NVSS J232914+375414	352.30937	37.90414
3FGL J2329.3-4955	PKS 2326-502	352.337	-49.92797
3FGL J2329.9-4734	PKS 2326-477	352.32377	-47.50531
3FGL J2330.4-3726	PKS 2327-376	352.64917	-37.41056
3FGL J2330.5+1104	4C +10.73	352.67022	11.0052
3FGL J2330.8-2144	PMN J2331-2148	352.76679	-21.80428
3FGL J2334.1+0732	TXS 2331+073	353.55345	7.60765
3FGL J2334.8+1432	NVSS J233453+143214	353.72433	14.53743
3FGL J2335.1-0133	PKS 2332-017	353.83505	-1.51933
3FGL J2336.5+2356	B2 2334+23	354.17583	23.92472
3FGL J2336.5-7620	PMN J2336-7620	354.11496	-76.34378
3FGL J2336.5-4116	PKS 2333-415	354.1416	-41.25611
3FGL J2338.1-0229	PKS 2335-027	354.48891	-2.51601
3FGL J2338.7-7401	1RXS J233919.8-740439	354.83742	-74.07669
3FGL J2339.0+2122	RX J2338.8+2124	354.73494	21.41147
3FGL J2340.7+8016	1RXS J234051.4+801513	355.22608	80.25444
3FGL J2343.6+1551	MG1 J234342+1542	355.92812	15.71749
3FGL J2343.7+3437	1RXS J234332.5+343957	355.89083	34.66789
3FGL J2344.4+0549	1RXS J234354.4+054713	355.98325	5.78444
3FGL J2345.2-1554	PMN J2345-1555	356.30193	-15.91884
3FGL J2346.7+0705	TXS 2344+068	356.66639	7.08523
3FGL J2347.0+5142	1ES 2344+514	356.77015	51.70497
3FGL J2347.9+5436	NVSS J234753+543627	356.97138	54.60756
3FGL J2348.0-1630	PKS 2345-16	357.01087	-16.52001
3FGL J2348.4-5100	SUMSS J234852-510311	357.22037	-51.05325
3FGL J2350.4-3004	NVSS J235034-300603	357.64283	-30.10119
3FGL J2352.0+1752	CLASS J2352+1749	298.752417	17.820639
3FGL J2353.3-4805	SUMSS J235310-480558	358.2925	-48.0995
3FGL J2353.6-3037	PKS 2351-309	358.44774	-30.63014
3FGL J2353.7-3911	NVSS J235342-391442	358.42888	-39.24508
3FGL J2354.1+4605	4C +45.51	358.59033	45.88451
3FGL J2355.5+8154	S5 2353+81	359.09497	81.88118
3FGL J2356.0+4037	NVSS J235612+403648	359.05308	40.61306
3FGL J2357.3-0150	PKS 2354-021	359.35474	-1.87097
3FGL J2357.4-1716	RBS 2066	359.37492	-17.30097
3FGL J2357.8-5310	PKS 2355-534	359.47194	-53.18714
3FGL J2358.0-4552	PKS 2355-461	359.50896	-45.92181
3FGL J2358.2-1022	PKS 2355-106	359.54534	-10.33573
3FGL J2358.3-2853	PMN J2358-2853	359.57071	-28.89281
3FGL J2358.9+3926	B2 2356+39	359.7494	39.37453
3FGL J2359.3-3038	H 2356-309	359.78296	-30.62794
3FGL J2359.5-2052	TXS 2356-210	359.83137	-20.799

A.2 3FGL Catalog Description

Table A.2: Features extracted from the 3FGL catalog. Description adapted from Acero et al., 2015.

Column	Unit	Description
Source_Name		Official source name
RAJ2000	deg	Right Ascension
DEJ2000	deg	Declination
GLON	deg	Galactic Longitude
GLAT	deg	Galactic Latitude
Conf_68_SemiMajor	deg	Long radius of uncertainty ellipse at 68% confidence
Conf_68_SemiMinor	deg	Short radius of uncertainty ellipse at 68% confidence
Conf_68_PosAng	deg	Position Angle of the 68% long axis from celestial North, positive towards increasing RA (eastward)
Conf_95_SemiMajor	deg	Long radius of uncertainty ellipse at 95% confidence
Conf_95_SemiMinor	deg	Short radius of uncertainty ellipse at 95% confidence
Conf_95_PosAng	deg	Position Angle of the 95% long axis from celestial North, positive towards increasing RA (eastward)
ROI_num		Region of Interest number
Signif_Avg	σ	Source significance (derived from Test Statistic) over the 100 MeV to 300 GeV band
Pivot_Energy	MeV	Energy at which uncertainty on differential flux is minimal
Flux_Density	$cm^{-2}\,MeV^{-1}\,s^{-1}$	Differential flux at Pivot_Energy
Unc_Flux_Density	$cm^{-2}\,MeV^{-1}\,s^{-1}$	1σ uncertainty on differential flux at Pivot Energy
Flux1000	$cm^{-2}\,s^{-1}$	Integral photon flux from 1 to 100 GeV
Unc_Flux1000	$cm^{-2}\,s^{-1}$	1σ uncertaitny on integral photon flux from 1 to 100 GeV
Energy_Flux100	$erg\,cm^{-2}\,s^{-1}$	Energy flux from 100 MeV to 100 GeV obtained by spectral fitting
Unc_Energy_Flux100	$erg\,cm^{-2}\,s^{-1}$	1σ uncertainty on energy flux from 100 MeV to 100 GeV
Signif_Curve		Significance in terms of σ of fit improvement between PowerLaw and LogParabola or PLExpCutoff; a value greater than 4 indicates significant curvature.
SpectrumType		Spectral type (PowerLaw, LogParabola, PLExpCutoff, PLSuperExpCutoff)
Spectral_Index		Best fit photon number power law index: for LogParabola spectra, index at Pivot_Energy; for PL(Super)ExpCutoff spectra, low-energy index
Unc_Spectral_Index		1σ uncertainty on Spectral Index
beta		Curvature parameter for LogParabola; NULL for other spectral types
Unc_beta		1σ uncertainty on beta for LogParabola; NULL for other spectral types
Cutoff	MeV	Cutoff energy for PL(Super)ExpCutoff; NULL for other spectral types
Unc_Cutoff	MeV	1σ uncertainty on cutoff energy for PL(Super)ExpCutoff; NULL for other spectral types
Exp_Index		Exponential index for PLSuperExpCutoff; NULL for other spectral types
Unc_Exp_Index		1σ uncertainty on exponential index for PLSuperExpCutoff; NULL for other spectral types
PowerLaw_Index		Best fit power law index; equal to Spectral Index if SpectrumType is PowerLaw
Flux100_300	$cm^{-2}\,s^{-1}$	Integral photon flux from 100 to 300 MeV
nuFnu100_300	$erg\,cm^{-2}\,s^{-1}$	Spectral energy distribution between 100 and 300 MeV
Sqrt_TS100_300		Square root of the Test Statistic between 100 and 300 MeV
Flux300_1000	$cm^{-2}\,s^{-1}$	Integral photon flux from 300 MeV to 1 GeV
nuFnu300_1000	$erg\,cm^{-2}\,s^{-1}$	Spectral energy distribution between 300 MeV and 1 GeV
Sqrt_TS300_1000		Square root of the Test Statistic between 300 MeV and 1 GeV
Flux1000_3000	$cm^{-2}\,s^{-1}$	Integral photon flux from 1 to 3 GeV
nuFnu1000_3000	$erg\,cm^{-2}\,s^{-1}$	Spectral energy distribution between 1 and 3 GeV
Sqrt_TS1000_3000		Square root of the Test Statistic between 1 and 3 GeV
Flux3000_10000	$cm^{-2}\,s^{-1}$	Integral photon flux from 3 to 10 GeV
nuFnu3000_10000	$erg\,cm^{-2}\,s^{-1}$	Spectral energy distribution between 3 and 10 GeV
Sqrt_TS3000_10000		Square root of the Test Statistic between 3 and 10 GeV
Flux10000_100000	$cm^{-2}\,s^{-1}$	Integral photon flux from 10 to 100 GeV
nuFnu10000_100000	$erg\,cm^{-2}\,s^{-1}$	Spectral energy distribution between 10 and 100 GeV
Sqrt_TS10000_100000		Square root of the Test Statistic between 10 and 100 GeV
Variability_Index		Sum of $2\cdot\log(\text{Likelihood})$ difference between the flux fitted in each time interval and the average flux over the full catalog interval; a value greater than 72.44 over 48 intervals indicates <1% chance of being a steady source
Signif_Peak	σ	Source significance in peak interval
Flux_Peak	$cm^{-2}\,s^{-1}$	Peak integral photon flux from 100 MeV to 100 GeV
Unc_Flux_Peak	$cm^{-2}\,s^{-1}$	1σ uncertainty on peak integral photon flux
Time_Peak	s	Time of center of interval in which peak flux was measured
Peak_Interval	s	Length of interval in which peak flux was measured
Extended_Source_Name		Cross-reference to the ExtendedSources extension for extended sources, if any
0FGL_Name		Name of corresponding 0FGL source, if any
1FGL_Name		Name of corresponding 1FGL source, if any
2FGL_Name		Name of corresponding 2FGL source, if any
1FHL_Name		Name of corresponding 1FHL source, if any
ASSOC_GAM1		Name of likely corresponding 1AGL source
ASSOC_GAM2		Name of likely corresponding 3EG source
ASSOC_GAM3		Name of likely corresponding EGR source
TEVCAT_FLAG		P if positional association with non-extended source in TeVCat E if associated with a more extended source in TeVCat, N if no TeV association
ASSOC_TEV		Name of likely corresponding TeV source from TeVCat
CLASS1		Class designation for associated source
ASSOC1		Name of identified or likely associated source
ASSOC2		Alternate name of identified or likely associated source
Flags		Source flags

B MAGIC Analysis

B.1 Selected Data

3FGLJ2346.7+0705 Data

The following runs have been used for the MAGIC analysis of the source 3FGLJ2346.7+0705 and have been downloaded on superstar level:
05055339, 05055340, 05055436, 05055437, 05055438, 05055439, 05055440, 05055441, 05055471, 05055472, 05055473, 05055474, 05055475, 05055502, 05055503, 05055504, 05055505, 05055506, 05055507, 05055581, 05055582, 05055583, 05055584, 05055629, 05055630, 05055631, 05055632, 05055633, 05055634, 05055635, 05055659, 05055660, 05055670, 05055671, 05055672, 05055699, 05055700, 05057801, 05057802, 05057803, 05057804, 05057838, 05057840, 05057895, 05058673, 05058855, 05058856, 05058857, 05058859.

Additionally, the following subruns have been used for the MAGIC analysis of the source 3FGLJ2346.7+0705 and have been downloaded on star level:
M1: 05057839.001-.020, 05057892.001-.005, .008-.011, .013, .015-.020, .022-.025, 05057893.001-.003, .005-.007, .009-.012, .014-.025, 05057894.001-.011, .013-.025, 05058674.001-.009, 05058858.001-.011, .016-.021.
M2: 05057839.001-.020, 05057892.001-.005, .008-.011, .013-.020, .022-.025, 05057893.001-.003, .005-.007, .010, .012, .015-.025, 05057894.001-.011, .013-.025, 05058674.001-.009, 05058858.001-.010, .015-.020.

Crab Nebula Data

The following runs have been used for the sanity check of the MAGIC analysis of the source 3FGLJ2346.7+0705 and for the Random Forest study, and have been downloaded on superstar level:
05056516, 05056517, 05056587, 05056588, 05057065, 05057066, 05057067, 05057068, 05057069, 05057070, 05057144, 05057145, 05057146, 05057147, 05057148, 05057188, 05057189, 05057190, 05057191, 05057192, 05057649, 05057650, 05057651, 05057652, 05057653, 05057977, 05057978, 05058069, 05058072, 05058073, 05058074, 05058749, 05059266, 05059267, 05059268.

Additionally, the following subruns have been used for the sanity check of the MAGIC analysis of the source 3FGLJ2346.7+0705 and the Random Forest study, and have been downloaded on star level:
M1: 05057193.002-.004, 05059212.001-.012.
M2: 05057193.002-.004, 05059212.001-.012.

Data used as Hadrons

The following runs have been used for the training of the gamma / hadron separation for the MAGIC analysis of the source 3FGLJ2346.7+0705 and have been downloaded on superstar level:
1ES 0229+200: 05057396, 05057397, 05058048, 05058049, 05058461, 05058462, 05058923, 05058924, 05058925, 05058926.
M15: 05055321, 05055322, 05055323, 05055324, 05055624, 05055625, 05055626, 05056900, 05056901, 05056902, 05056903.
S3 0218+35: 05056153, 05056154, 05056155, 05056156, 05056617, 05056618, 05057172, 05057173.
Triangulum II: 05055980, 05055981, 05056416, 05056417, 05057051, 05057052, 05057865, 05057602, 05057603.

The following runs have been used for the Random Forest study and have been downloaded on superstar level:
1ES 0229+200: 05058048, 05058049, 05058461, 05058462, 05058925, 05058926.
M15: 05055321, 05055322, 05055625, 05055626, 05056900, 05056901.
S3 0218+35: 05056153, 05056154, 05056617, 05056618, 05057172, 05057173.
Triangulum II: 05056416, 05056417, 05057051, 05057052, 05057602, 05057603, 05057865.

B.2 Input Cards

coach.rc

```
RF.numTree:    100
RF.trainRatio: 0.95
RF.createTestSample: TRUE
RF.zdmin: 5.
RF.zdmax: 35.
RFLoop.FilterCuts.Continue0.Condition: MHillas_1.fSize<50.
RFLoop.FilterCuts.Continue1.Condition: MHillas_2.fSize<50.
RFLoop.FilterCuts.Continue2.Condition: MHillas_1.fSize>50000.
RFLoop.FilterCuts.Continue3.Condition: MHillas_2.fSize>50000.
RFLoop.FilterCuts.Continue4.Condition: MNewImagePar_1.fLeakage1>0.15
RFLoop.FilterCuts.Continue5.Condition: MNewImagePar_2.fLeakage1>0.15
RFLoop.FilterCuts.Continue6.Condition: MStereoPar.fValid<0.5

#####

RFLoop.GHCuts.Continue0.Condition: MImagePar_1.fNumIslands>1
RFLoop.GHCuts.Continue1.Condition: MImagePar_2.fNumIslands>1
RF.ReZenithing: TRUE
RF.numAzBins:   1
RF.numZdBins:   30
RF.NumTryGH: 3
RF.NdSizeGH: 5
RF.NumVariableGH: 12

RF.GHVariable1:  0.025+0.05*floor(log10(MHillas_1.fSize)/0.05)
RF.GHVariable2:  0.025+0.05*floor(log10(MHillas_2.fSize)/0.05)
RF.GHVariable3:  MHillas_1.fWidth
RF.GHVariable4:  MHillas_2.fWidth
RF.GHVariable5:  MHillas_1.fLength
RF.GHVariable6:  MHillas_2.fLength
RF.GHVariable7:  MStereoPar.fM1Impact
RF.GHVariable8:  MStereoPar.fM2Impact
RF.GHVariable9:  MStereoPar.fMaxHeight
RF.GHVariable10: sqrt(MHillasTimeFit_1.fP1Grad*MHillasTimeFit_1.fP1Grad)
RF.GHVariable11: sqrt(MHillasTimeFit_2.fP1Grad*MHillasTimeFit_2.fP1Grad)
RF.GHVariable12: (0.5/30.)+(1./30.)*floor(cos(MPointingPos_1.fZd*0.0174532925)/(1./30.))

#####

RFLoop.EnergyCuts.Continue0.Condition: MStereoPar.fCherenkovRadius<4000
RFLoop.EnergyCuts.Continue1.Condition: MStereoPar.fTheta2>0.1
RFLoop.EnergyCutsUnphys.Continue0.Condition: MStereoPar.fValid<1
RFLoop.EnergyCutsUnphys.Continue1.Condition: MStereoPar.fCherenkovDensity<0
MEnergyTable.SizeBinning 19
MEnergyTable.MinSize 25
MEnergyTable.MaxSize 200000
MEnergyTable.ImpactBinning 50
MEnergyTable.MinImpact 0
MEnergyTable.MaxImpact 3.5
MEnergyTable.ZdCorrectionFormula 0.97*pow(x,-0.3)/(1-pow(1-x,2.25))
MEnergyTable.LeakageCorrectionFormula_1 1-4*x*x
```

```
MEnergyTable.LeakageCut_1 0.2
MEnergyTable.LeakageCorrectionFormula_2 1-4*x*x
MEnergyTable.LeakageCut_2 0.2
MEnergyTable.BCorrectionFormula 0.93+0.2*sqrt(1.-pow(x,2))
MEnergyTable.MinEvtPerBin 3

########

RF.NumTryDisp: 3
RF.NdSizeDisp: 5
RF.NumVariableDisp: 11

RF.Disp1Variable1:  log10(MHillas_1.fSize)
RF.Disp1Variable2:  log10(MHillas_2.fSize)
RF.Disp1Variable3:  MStereoPar.fM1Impact
RF.Disp1Variable4:  MStereoPar.fM2Impact
RF.Disp1Variable5:  MPointingPos_1.fZd
RF.Disp1Variable6:  MStereoPar.fMaxHeight
RF.Disp1Variable7:  sqrt(MHillasTimeFit_1.fP1Grad*MHillasTimeFit_1.fP1Grad)
RF.Disp1Variable8:  sqrt(MHillasTimeFit_2.fP1Grad*MHillasTimeFit_2.fP1Grad)
RF.Disp1Variable9:  MHillas_1.fWidth
RF.Disp1Variable10: MHillas_1.fLength
# To be estimated in regression:
RF.Disp1Variable11: MHillasSrc_1.fDist*0.0033703

RF.Disp2Variable1:  log10(MHillas_1.fSize)
RF.Disp2Variable2:  log10(MHillas_2.fSize)
RF.Disp2Variable3:  MStereoPar.fM1Impact
RF.Disp2Variable4:  MStereoPar.fM2Impact
RF.Disp2Variable5:  MPointingPos_1.fZd
RF.Disp2Variable6:  MStereoPar.fMaxHeight
RF.Disp2Variable7:  sqrt(MHillasTimeFit_1.fP1Grad*MHillasTimeFit_1.fP1Grad)
RF.Disp2Variable8:  sqrt(MHillasTimeFit_2.fP1Grad*MHillasTimeFit_2.fP1Grad)
RF.Disp2Variable9:  MHillas_2.fWidth
RF.Disp2Variable10: MHillas_2.fLength
# To be estimated in regression:
RF.Disp2Variable11: MHillasSrc_2.fDist*0.0033703
```

odie.rc

```
Odie.sourceRa: 23.777760757
Odie.sourceDec: 7.085236944
Odie.minZenith: 0.
Odie.maxZenith: 35.
Odie.analysisEpoch: Jul13
Odie.stereo: TRUE
Odie.deadTime: 26.e-6
Odie.skipNormalization: FALSE
Odie.normAbove: 0.1
Odie.normBelow: 0.25
Odie.nWobbleOff: 3
Odie.doExtensionFit: FALSE

# For low energy (LE) analysis
Odie.eRange: LE
Odie.signalCut: 0.0165
Odie.psf40: 0.053
Odie.psftailfract: 0.531
Odie.psftailsigma: 0.105
Odie.psfaccuracy: 0.167

# For full range (FR) analysis
Odie.eRange: FR
Odie.signalCut: 0.0091
Odie.psf40: 0.038
Odie.psftailfract: 0.669
Odie.psftailsigma: 0.066
Odie.psfaccuracy: 0.212

# For high energy (HE) analysis
Odie.eRange: HE
Odie.signalCut: 0.0081
Odie.psf40: 0.031
Odie.psftailfract: 0.651
Odie.psftailsigma: 0.055
Odie.psfaccuracy: 0.323
```

flute.rc

```
flute.sourceRa: 23.777760757
flute.sourceDec: 7.085236944
flute.deadTimePerEvent: 26.e-6
flute.minZd: 0.
flute.maxZd: 35.
flute.minSize: 50
flute.bckgMode: 1
flute.numSimultaneousBgPositions: 3
flute.posContainer: MStereoParDisp
flute.nBinsTheta2: 40
flute.maxTheta2: 0.4
flute.normRangeTheta2Min: 0.1
flute.normRangeTheta2Max: 0.4
flute.nBinsEnergyEst: 10
flute.minEnergyEst: 5.
flute.maxEnergyEst: 50000.
flute.estTrueFactor: 1.41
flute.ReDoEnergyAverage: FALSE
flute.nBinsAz: 4
flute.EminLC: 100.
flute.EmaxLC: inf
flute.LCbinning: night-wise
flute.FindCutsFromEfficiency: TRUE
flute.Theta2Effi: 0.75
flute.HadEffi: 0.9
flute.allowedHadRange: 0.15, 0.95
flute.allowedTheta2Range: 0.01, 0.2
flute.SourceRedshift: 0.
flute.AssumedSpectrum: pow(x/300., -2)
flute.maximumRelError: 0.5
flute.maximumRelErrorLC: 0.5
flute.AllowNegativeExcessInUpperLimit: TRUE
flute.SetMinAllowedUpperLimitAsFractionOfBackground: 0.03
flute.PropagateEffAreaError: TRUE
flute.NameEnergyEst: MEnergyEst
flute.AeffAtmCorr: FALSE
flute.AtmCorrDisplay: FALSE
```

caspar.rc

```
Caspar.cuts: STANDARD
Caspar.deadTime: 26.e-6
Caspar.analysisEpoch: Jul13
Caspar.zenithRange: 0., 35.
Caspar.stereo: TRUE
Caspar.nWobblePosition: 4
Caspar.wobbleMap: TRUE
Caspar.rangeRaDec: 2.0, 2.0

# For low energy (LE) analysis
Caspar.eRange: LE
Caspar.psf40: 0.081

# For full range (FR) analysis
Caspar.eRange: FR
Caspar.psf40: 0.057
```

B.3 Input Cards for Optimized Random Forest Applications

coach.rc

```
RF.numTree:    100
RF.zdmin: 5.
RF.zdmax: 35.

# G/h Separation

RFLoop.GHCuts.Continue0.Condition: MImagePar_1.fNumIslands>1
RFLoop.GHCuts.Continue1.Condition: MImagePar_2.fNumIslands>1
RFLoop.GHCuts.Continue2.Condition: MHillas_1.fSize<50.
RFLoop.GHCuts.Continue3.Condition: MHillas_2.fSize<50.
RFLoop.GHCuts.Continue4.Condition: MHillas_1.fSize>50000.
RFLoop.GHCuts.Continue5.Condition: MHillas_2.fSize>50000.
RFLoop.GHCuts.Continue6.Condition: MNewImagePar_1.fLeakage1>0.15
RFLoop.GHCuts.Continue7.Condition: MNewImagePar_2.fLeakage1>0.15
RFLoop.GHCuts.Continue8.Condition: MStereoPar.fValid<0.5

RF.ReZenithing: TRUE
RF.numAzBins:   1
RF.numZdBins:   30
RF.NumTryGH: 3
RF.NdSizeGH: 5
RF.NumVariableGH: 12

RF.GHVariable1:  0.025+0.05*floor(log10(MHillas_1.fSize)/0.05)
RF.GHVariable2:  0.025+0.05*floor(log10(MHillas_2.fSize)/0.05)
RF.GHVariable3:  MHillas_1.fWidth
RF.GHVariable4:  MHillas_2.fWidth
RF.GHVariable5:  MHillas_1.fLength
RF.GHVariable6:  MHillas_2.fLength
RF.GHVariable7:  MStereoPar.fM1Impact
RF.GHVariable8:  MStereoPar.fM2Impact
RF.GHVariable9:  MStereoPar.fMaxHeight
RF.GHVariable10: sqrt(MHillasTimeFit_1.fP1Grad*MHillasTimeFit_1.fP1Grad)
RF.GHVariable11: sqrt(MHillasTimeFit_2.fP1Grad*MHillasTimeFit_2.fP1Grad)
RF.GHVariable12: (0.5/30.)+(1./30.)*floor(cos(MPointingPos_1.fZd*0.0174532925)/(1./30.))


## Energy reconstruction

RFLoop.EnergyCuts.Continue0.Condition: MHillas_1.fSize<50.
RFLoop.EnergyCuts.Continue1.Condition: MHillas_2.fSize<50.
RFLoop.EnergyCutsUnphys.Continue0.Condition: MStereoPar.fValid<1
RFLoop.EnergyCutsUnphys.Continue1.Condition: MStereoPar.fCherenkovDensity<1e-16

# Method 1: Random Forest Regression

RF.NumTreeEn: 200
RF.NumTryEn: 4
RF.NdSizeEn: 5
RF.NumVariableEn: 19
```

```
RF.EnVariable1: MNewImagePar_1.fLeakage1
RF.EnVariable2: MNewImagePar_2.fLeakage1
RF.EnVariable3: MNewImagePar_1.fLeakage2
RF.EnVariable4: MNewImagePar_2.fLeakage2
RF.EnVariable5: MNewImagePar_1.fConc1
RF.EnVariable6: MNewImagePar_2.fConc1
RF.EnVariable7: MHillasTime_1.fRMSTimeW
RF.EnVariable8: MHillasTime_2.fRMSTimeW
RF.EnVariable9: MStereoPar.fM1Impact
RF.EnVariable10: MStereoPar.fM2Impact
RF.EnVariable11: MStereoPar.fMaxHeight
RF.EnVariable12: log10(MHillas_1.fSize/MStereoPar.fCherenkovDensity)
RF.EnVariable13: log10(MHillas_2.fSize/MStereoPar.fCherenkovDensity)
RF.EnVariable14: sqrt(MHillasTimeFit_1.fP2Grad*MHillasTimeFit_1.fP2Grad)
RF.EnVariable15: sqrt(MHillasTimeFit_2.fP2Grad*MHillasTimeFit_2.fP2Grad)
RF.EnVariable16: MHillas_1.fSize/MNewImagePar_1.fUsedArea
RF.EnVariable17: MHillas_1.fSize/MNewImagePar_2.fUsedArea
RF.EnVariable18: MPointingPos_1.fZd
# To be estimated in regression
RF.EnVariable19: log10(MMcEvt_1.fEnergy)

# Method 2: Lookup Table Method

MEnergyTable.SizeBinning 19
MEnergyTable.MinSize 25
MEnergyTable.MaxSize 200000
MEnergyTable.ImpactBinning 50
MEnergyTable.MinImpact 0
MEnergyTable.MaxImpact 3.5

MEnergyTable.ZdCorrectionFormula 0.97*pow(x,-0.3)/(1-pow(1-x,2.25))
MEnergyTable.LeakageCorrectionFormula_1 1-4*x*x
MEnergyTable.LeakageCut_1 0.2
MEnergyTable.LeakageCorrectionFormula_2 1-4*x*x
MEnergyTable.LeakageCut_2 0.2
MEnergyTable.BCorrectionFormula 0.93+0.2*sqrt(1.-pow(x,2))


# Direction Recontruction

RFLoop.DispCuts.Continue0.Condition: MHillas_1.fSize<50.
RFLoop.DispCuts.Continue1.Condition: MHillas_2.fSize<50.
RFLoop.DispCuts.Continue2.Condition: MHillas_1.fSize>50000.
RFLoop.DispCuts.Continue3.Condition: MHillas_2.fSize>50000.
RFLoop.DispCuts.Continue4.Condition: MNewImagePar_1.fLeakage1>0.15
RFLoop.DispCuts.Continue5.Condition: MNewImagePar_2.fLeakage1>0.15
RFLoop.DispCuts.Continue6.Condition: MStereoPar.fValid<0.5

RF.NumTryDisp: 3
RF.NdSizeDisp: 5
RF.NumVariableDisp: 11

RF.Disp1Variable1:  log10(MHillas_1.fSize)
RF.Disp1Variable2:  log10(MHillas_2.fSize)
RF.Disp1Variable3:  MStereoPar.fM1Impact
```

```
RF.Disp1Variable4:  MStereoPar.fM2Impact
RF.Disp1Variable5:  MPointingPos_1.fZd
RF.Disp1Variable6:  MStereoPar.fMaxHeight
RF.Disp1Variable7:  sqrt(MHillasTimeFit_1.fP1Grad*MHillasTimeFit_1.fP1Grad)
RF.Disp1Variable8:  sqrt(MHillasTimeFit_2.fP1Grad*MHillasTimeFit_2.fP1Grad)
RF.Disp1Variable9:  MHillas_1.fWidth
RF.Disp1Variable10: MHillas_1.fLength
# To be estimated in regression
RF.Disp1Variable11: MHillasSrc_1.fDist*0.0033703

RF.Disp2Variable1:  log10(MHillas_1.fSize)
RF.Disp2Variable2:  log10(MHillas_2.fSize)
RF.Disp2Variable3:  MStereoPar.fM1Impact
RF.Disp2Variable4:  MStereoPar.fM2Impact
RF.Disp2Variable5:  MPointingPos_1.fZd
RF.Disp2Variable6:  MStereoPar.fMaxHeight
RF.Disp2Variable7:  sqrt(MHillasTimeFit_1.fP1Grad*MHillasTimeFit_1.fP1Grad)
RF.Disp2Variable8:  sqrt(MHillasTimeFit_2.fP1Grad*MHillasTimeFit_2.fP1Grad)
RF.Disp2Variable9:  MHillas_2.fWidth
RF.Disp2Variable10: MHillas_2.fLength
# To be estimated in regression
RF.Disp2Variable11: MHillasSrc_2.fDist*0.0033703
```

melibea.rc (extract)

```
MJMelibea.UseEnergyLog: TRUE
MJMelibea.UseEnergyMedian: TRUE
```

B.4 Description of Features

Table B.1: Features available on *star*-level, extracted from *MARS*. Description adapted from source code.

Column	Unit	Description
		Hillas container
Length	mm	Second moment along the major axis of the shower ellipse, weighted with pixel charges
Width	mm	Second moment along the minor axis of the shower ellipse, weighted with pixel charges
Delta	rad	Angle between major shower axis and x-axis
Size	phe	Total charge of the camera image
MeanX	mm	X-coordinate of the center of gravity, determined by the first moment along the x-axis of the camera, weighted with pixel charges
MeanY	mm	Y-coordinate of the center of gravity, determined by the first moment along the y-axis of the camera, weighted with pixel charges
SinDelta		Sine of Delta
CosDelta		Cosine of Delta
		HillasExt container
Asym	mm	Distance between the center of gravity and the pixel with the highest charge
M3Long	mm	Third moment along the major axis of the shower ellipse, weighted with pixel charges
M3Trans	mm	Third moment along the minor axis of the shower ellipse, weighted with pixel charges
MaxDist	mm	Distance between the camera center and the most distant selected pixel
AsymPow2[1]	mm	Second moment along the major axis of the shower ellipse, weighted with squared pixel charges
M3LongPow3	mm	Third moment along the major axis of the shower ellipse, weighted with cubed pixel charges
AsymTimePow2		Approximation of time gradient
BetaAsymT	deg	Angle between center of gravity, weighted with pixel charges and times and connected with camera center and center of gravity, weighted with pixel charges and times squared and connected with camera center
CurvMuonRad		Radius of fitted muon ring
MuonFitChi2		Chi^2 of muon fit
TmeanCore	time slices	Mean arrival time of shower core, weighted with squared pixel charges
MuonRingCenterDist		Distance from center of fitted muon ring to major shower axis
		HillasSrc container
Alpha	deg	Angle between major axis of the shower ellipse and center of gravity connected with the source position in the camera image
Dist	mm	Distance between center of gravity and source position in camera image
CosDeltaAlpha		Cosine of angle between center of gravity connected with the source position in the camera and the major axis of the shower ellipse defined with positive x-component
DCA	mm	
DCADelta	deg	Similar to Delta but in interval (0, 360)
		ImagePar container
NumIslands		Number of separated pixel groups
NumSinglePixels		Number of separated single pixels
SizeSinglePixels	phe	Total charge of single pixels
SizeSubIslands	phe	Total charge of the camera image subtracted by the total charge of the largest separated pixel group
SizeMainIsland	phe	Total charge of the largest separated pixel group
NumSatPixelsHG		Parameter not calculated in *MARS V2-17-2*
NumSatPixelsLG		Parameter not calculated in *MARS V2-17-2*
NumClustersPre		Parameter not calculated in *MARS V2-17-2*
NumClustersAfter		Parameter not calculated in *MARS V2-17-2*
SizeMaxClusterPre		Parameter not calculated in *MARS V2-17-2*
CorePrimariesCharge		Parameter not calculated in *MARS V2-17-2*
CoreSecondariesCharge		Parameter not calculated in *MARS V2-17-2*
ClusterSize		Parameter not calculated in *MARS V2-17-2*
ClusterSizeRMS		Parameter not calculated in *MARS V2-17-2*
TimeJitter		Parameter not calculated in *MARS V2-17-2*
		NewImagePar container
Leakage1		Ratio between charge in most outer ring of pixels and total charge in camera image
Leakage2		Ratio between charge in two most outer rings of pixels and total charge in camera image
InnerLeakage1		Parameter not calculated in *MARS V2-17-2*
InnerLeakage2		Parameter not calculated in *MARS V2-17-2*
InnerSize	phe	Similar to Size due to same pixel sizes in camera nowadays
Conc		Ratio between the two largest charges per pixel and the total charge in camera image
Conc1		Ratio between the largest charge per pixel and the total charge in camera image
UsedArea	mm^2	Area of pixels surviving the image cleaning
CoreArea	mm^2	Area of core pixels
NumUsedPixels		Number of pixels surviving the image cleaning
NumCorePixels		Number of core pixels

Table B.1: – continued from previous page

Column	Unit	Description
		PointingPos container
Zd	deg	Zenith angle of the telescope
Az	deg	Azimuth angle of the telescope
DevZd	deg	Deviation of the zenith angle, determined by the starguider
DevAz	deg	Deviation of the azimuth angle, determined by the starguider
Ra	h	Right ascension the telescope is pointing at
Ha	h	Hour angle the telescope is pointing at
Dec	deg	Declination the telescope is pointing at
DevHa	h	Deviation of the hour angle determined by the starguider
DevDec	deg	Deviation of the declination determined by the starguider
		MoonPos container
MoonZd	deg	Zenith angle of the moon
MoonAz	deg	Azimuth angle of the moon
MoonRa	h	Right ascension of the moon
MoonDec	deg	Declination fo the moon
MoonAngle	deg	Angle between the pointing of the telescope and the moon
MoonPhase		Moon phase
		Concentration container
Conc[9]		Ratio between the ten largest charges per pixel and the total charge in camera image
		SrcPosCam container
X	mm	X-coordinate of the source position projected to the camera image
Y	mm	Y-coordinate of the source position projected to the camera image
		HillasTime container
MeanTime	time slices	Mean arrival time
RMSTime	time slices	Uncorrected sample standard deviation of the arrival time
MeanTimeW	time slices	Mean arrival time, weighted with pixel charges
RMSTimeW	time slices	Uncorrected sample standard deviation of the arrival time, weighted with pixel charges
		HillasTimeFit container
P1Const	time slices	Fitted constant parameter of linear fit
P1Grad	time slices / mm	Fitted linear parameter of linear fit
P1Chi2		Chi^2 of linear fit
P1NDF		Number of degrees of freedom of linear fit
P1RMSTimeFit	time slices	Standard deviation of the temporal progress of the arrival times determined with a linear fit with respect to major axis of the shower ellipse
P2Const	time slices	Fitted constant parameter of squared fit
P2Grad	time slices / mm	Fitted linear parameter of squared fit
P2Curv	time slices / mm^2	Fitted squared parameter of squared
P2Chi2		Chi^2 of squared fit
P2NDF		Number of degrees of freedom of squared fit
P2RMSTimeFit	time slices	Standard deviation of the temporal progress of the arrival times determined with a squared fit with respect to major axis of the shower ellipse
MinRforFit	mm	Lower limit for the fits
MaxRforFit	mm	Upper limit for the fits
		Time container
Mjd	d	Modified Julian Date of the observation
Time	ms	Time of day with respect to Mjd
NanoSec	ns	Nanosecond part of Time

Table B.2: Features available on *superstar*-level, extracted from *MARS*. Description adapted from source code.

Column	Unit	Description
		StereoPar container
Valid		Indicator for successfull reconstruction
DirectionX	deg	Estimated x-coordinate of the shower direction in MAGIC-I camera
DirectionY	deg	Estimated y-coordinate of the shower direction in MAGIC-I camera
DirectionZd	deg	Estimated zenith angle of the shower direction
DirectionAz	deg	Estimated azimuth angle of the shower direction
DirectionDec	deg	Estimated declination of the shower direction
DirectionRA	h	Estimated right ascension of the shower direction
Theta2	deg^2	Squared distance between estimated shower position and source position in camera
CoreX	cm	Estimated x-coordinate of the shower core impact position on ground
CoreY	cm	Estimated y-coordinate of the shower core impact position on ground
M1Impact	cm	Distance between estimated shower core impact position and MAGIC-I telescope
M2Impact	cm	Distance between estimated shower core impact position and MAGIC-II telescope
M1ImpactAz	deg	Azimuth angle of MAGIC-I telescope of the estimated shower core position, projected to plane perpendicular to the shower axis
M2ImpactAz	deg	Azimuth angle of MAGIC-II telescope of the estimated shower core position, projected to plane perpendicular to the shower axis
MaxHeight	cm	Estimated height above the telescope of the shower maximum
XMax	g/cm^2	Depth of maximum number of particles in shower
CherenkovRadius	cm	Radius of Cherenkov ring at ground, produced by an 86 MeV electron at the max height of the shower
CherenkovDensity	a. u.	Scaled ratio between Cherenkov light and area at ground level, produced by an 86 MeV electron at the max height of the shower
Energy		Parameter not calculated by default in *MARS V2-17-2*
EnergyUncertainty		Parameter not calculated by default in *MARS V2-17-2*
EnergyDiscrepancy		Parameter not calculated by default in *MARS V2-17-2*
PhiBaseLineM1	deg	Angle between azimuth angle of MAGIC-I and the connecting line of MAGIC-I and MAGIC-II
PhiBaseLineM2	deg	Angle between azimuth angle of MAGIC-II and the connecting line of MAGIC-I and MAGIC-II
ImageAngle	deg	Angle between the major axes of the telescopes
DispRMS		Parameter not calculated by default in *MARS V2-17-2*
DispDiff2		Parameter not calculated by default in *MARS V2-17-2*
CosBSangle		Cosine of angle between shower axis and geomagnetic field
		EnergyEst container
Energy		Parameter not calculated by default in *MARS V2-17-2*
EnergyRMS		Parameter not calculated by default in *MARS V2-17-2*
Impact		Parameter not calculated by default in *MARS V2-17-2*
Uncertainty		Parameter not calculated by default in *MARS V2-17-2*

Table B.3: Features available on *melibea*-level, extracted from *MARS*. Description adapted from source code.

Column	Unit	Description
		EnergyEst_1/2 container
Energy	GeV	Energy determined by the lookup table
EnergyRMS		Parameter not calculated by default in *MARS V2-17-2*
Impact		Parameter not calculated by default in *MARS V2-17-2*
Uncertainty	GeV	Uncertainty of energy estimation using lookup table
		EnergyEst_SrcDep_1/2 container
Energy		Parameter not calculated by default in *MARS V2-17-2*
EnergyRMS		Parameter not calculated by default in *MARS V2-17-2*
Impact		Parameter not calculated by default in *MARS V2-17-2*
Uncertainty		Parameter not calculated by default in *MARS V2-17-2*
		ImageParDisp_1/2 container
Disp	deg	Estimated distance of reconstructed source position from center of gravity along major shower axis, determined with a Random Forest Regression
DispRMS	deg	Uncertainty of the regression
Xdisp	deg	X-coordinate of reconstructed source position, positive / negative direction along major shower axis according to asymmetry of the shower
Ydisp	deg	Y-coordinate of reconstructed source position, positive / negative direction along major shower axis according to asymmetry of the shower
		StereoParDisp container
Valid		Indicator for successfull reconstruction
DirectionX	deg	Estimated x-coordinate of the shower direction in MAGIC-I camera
DirectionY	deg	Estimated y-coordinate of the shower direction in MAGIC-I camera
DirectionZd	deg	Estimated zenith angle of the shower direction
DirectionAz	deg	Estimated azimuth angle of the shower direction
DirectionDec	deg	Estimated declination of the shower direction
DirectionRA	h	Estimated right ascension of the shower direction
Theta2	deg^2	Squared distance between estimated shower position and source position in camera
CoreX	cm	Estimated x-coordinate of the shower core impact position on ground
CoreY	cm	Estimated y-coordinate of the shower core impact position on ground
M1Impact	cm	Distance between estimated shower core impact position and MAGIC-I telescope
M2Impact	cm	Distance between estimated shower core impact position and MAGIC-II telescope
M1ImpactAz	deg	Azimuth angle of MAGIC-I telescope of the estimated shower core position, projected to plane perpendicular to the shower axis
M2ImpactAz	deg	Azimuth angle of MAGIC-II telescope of the estimated shower core position, projected to plane perpendicular to the shower axis
MaxHeight	cm	Estimated height above the telescope of the shower maximum
XMax	g/cm^2	Depth of maximum number of particles in shower
CherenkovRadius	cm	Radius of Cherenkov ring at ground, produced by an 86 MeV electron at the max height of the shower
CherenkovDensity	a. u.	Scaled ratio between Cherenkov light and area at ground level, produced by an 86 MeV electron at the max height of the shower
Energy	GeV	Average of energies of both telescopes determined with lookup tables and weighted with corresponding uncertainty of energy estimation
EnergyUncertainty	GeV	Uncertainty of average of energies of both telescopes by propagating individual uncertainties
EnergyDiscrepancy		Chi^2 of energies and uncertainties
PhiBaseLineM1		Parameter not calculated by default in *MARS V2-17-2*
PhiBaseLineM2		Parameter not calculated by default in *MARS V2-17-2*
ImageAngle		Parameter not calculated by default in *MARS V2-17-2*
DispRMS	deg	Uncertainty of Disp calculation
DispDiff2	deg^2	Squared distance between reconstructed positions of individual telescopes
CosBSangle		Cosine of angle between shower axis and geomagnetic field
		Hadronness container
Hadronness		Average of score of multiple decision trees in classification of gamma rays and hadrons
HadronnessRMS		Standard deviation of score of multiple decision trees
		EnergyEst container
Energy	GeV	Average of energies of both telescopes determined with lookup tables and weighted with corresponding uncertainty of energy estimation
EnergyRMS		Parameter not calculated by default in *MARS V2-17-2*
Impact		Parameter not calculated by default in *MARS V2-17-2*
Uncertainty	GeV	Uncertainty of average of energies of both telescopes by propagating individual uncertainties

Bibliography

Aab, A. et al. (2015). "The Pierre Auger Cosmic Ray Observatory". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 798.Supplement C, pp. 172–213.

Abdo, A. A. et al. (2010). "Fermi Large Area Telescope First Source Catalog". In: *The Astrophysical Journal Supplement Series* 188.2, p. 405.

Abeysekara, A.U. et al. (2013). "Sensitivity of the high altitude water Cherenkov detector to sources of multi-TeV gamma rays". In: *Astroparticle Physics* 50.Supplement C, pp. 26–32.

Acero, F. et al. (2013). "Hunting for Treasures among the Fermi Unassociated Sources: A multiwavelength Approach". In: *The Astrophysical Journal* 779.2, p. 133.

Acero, F. et al. (2015). "Fermi Large Area Telescope Third Source Catalog". In: *The Astrophysical Journal Supplement Series* 218.2, p. 23.

Acharya, B. S. et al. (2017). "Science with the Cherenkov Telescope Array". In:

Ackermann, M. et al. (2011). "The Radio/Gamma-Ray Connection in Active Galactic Nuclei in the Era of the Fermi Large Area Telescope". In: *The Astrophysical Journal* 741.1, p. 30.

Ackermann, M. et al. (2012a). "A Statistical Approach to Recognizing Source Classes for Unassociated Sources in the First Fermi-LAT Catalog". In: *The Astrophysical Journal* 753.1, p. 83.

Ackermann, M. et al. (2012b). "The Fermi Large Area Telescope on Orbit: Event Classification, Instrument Response Functions, and Calibration". In: *The Astrophysical Journal Supplement Series* 203.1, p. 4.

Ackermann, M. et al. (2015a). "The Spectrum of Isotropic Diffuse Gamma-Ray Emission between 100 MeV and 820 GeV". In: *The Astrophysical Journal* 799.1, p. 86.

Ackermann, M. et al. (2015b). "The Third Catalog of Active Galactic Nuclei Detected by the Fermi Large Area Telescope". In: *The Astrophysical Journal* 810.1, p. 14.

Ageron, M. et al. (2011). "ANTARES: The first undersea neutrino telescope". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 656.1, pp. 11–38.

Aharonian, F. et al. (2006). "Observations of the Crab nebula with HESS". In: *Astronomy & Astrophysics* 457, pp. 899–915.

Alam et al., S. (2015). "The Eleventh and Twelfth Data Releases of the Sloan Digital Sky Survey: Final Data from SDSS-III". In: *The Astrophysical Journal Supplement Series* 219.1, p. 12.

Albert, J. et al. (2008a). "Implementation of the Random Forest method for the Imaging Atmospheric Cherenkov Telescope MAGIC". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 588.3, pp. 424–432.

Albert, J. et al. (2008b). "VHE γ-Ray Observation of the Crab Nebula and its Pulsar with the MAGIC Telescope". In: *Astroparticle Journal* 674, 1037-1055, pp. 1037–1055.

Aleksić, J. et al. (2015). "Measurement of the Crab Nebula spectrum over three decades in energy with the MAGIC telescopes". In: *Journal of High Energy Astrophysics* 5–6, pp. 30–38.

Aleksić, J. et al. (2016a). "The major upgrade of the MAGIC telescopes, Part I: The hardware improvements and the commissioning of the system". In: *Astroparticle Physics* 72, pp. 61–75.

Aleksić, J. et al. (2016b). "The major upgrade of the MAGIC telescopes, Part II: A performance study using observations of the Crab Nebula". In: *Astroparticle Physics* 72, pp. 76–94.

Álvarez Crespo, N. et al. (2016). "Optical archival spectra of blazar candidates of uncertain type in the 3rd Fermi Large Area Telescope Catalog". In: *Astrophysics and Space Science* 361.9, p. 316.

Angelini, L. and C. A. Padgett (2010). *XIMAGE – An X-Ray Image Package – User's Guide*. Version 4.5.1.

Atwood, W. B. et al. (2009). "The Large Area Telescope on the Fermi Gamma-Ray Space Telescope Mission". In: *The Astrophysical Journal* 697.2, p. 1071.

Beckmann, V. and C. Shrader (2012). *Active Galactic Nuclei*. Wiley-VCH Verlag GmbH & Co. KGaA.

Boller, T. et al. (2016). "Second ROSAT all-sky survey (2RXS) source catalogue". In: *Astronomy and Astrophysics* 588, A103, A103.

Bonnefoy, S. et al. (2017). "Performance of the upgraded H.E.S.S. cameras". In: *The 35th International Cosmic Ray Conference*. (Bexco, Busan, Korea). 805. Proceedings of Science.

Breiman, L. and P. Spector (1992). "Submodel Selection and Evaluation in Regression. The X-Random Case". In: *International Statistical Review / Revue Internationale de Statistique* 60.3, pp. 291–319.

Breiman, L. et al. (1984). *Classification and Regression Trees*. New York: Chapman & Hall, p. 358.

Breimann, L. (2001). "Random Forests". In: *Machine Learning* 45.1, pp. 5–32.

Burrows, D. N. et al. (2005). "The Swift X-Ray Telescope". In: *Space Science Reviews* 120, pp. 165–195.

Capalbi, M., M. Perri, B. Saija, and F. Tamburelli (2005). *The Swift-XRT Data Reduction Guide*.

Cherenkov, P.A. (1937). *Visible Radiation Produced by Electrons Moving in a Medium with Velocities Exceeding that of Light*. American Physical Society.

Condon, J. J. et al. (1998). "The NRAO VLA Sky Survey". In: *Astronomical Journal* 115, pp. 1693–1716.

D'Abrusco, R. et al. (2013). "Unveiling the Nature of Unidentified Gamma-Ray Sources. I. A New Method for the Association of Gamma-Ray Blazars". In: *The Astrophysical Journal Supplement Series* 206.2, p. 12.

Degrange, B. and G. Fontaine (2015). "Introduction to high-energy gamma-ray astronomy". In: *Comptes Rendus Physique* 16.6. Gamma-ray astronomy / Astronomie des rayons gamma, pp. 587–599.

D'Elia, V. et al. (2013). "The seven year Swift-XRT point source catalog (1SWXRT)". In: *Astronomy and Astrophysics* 551.

D'Elia, V. et al. (2013). "The seven year Swift-XRT point source catalog (1SWXRT)". In: *Astronomy and Astrophysics* 551, A142, A142.

Doert, M. and M. Errando (2014). "Search for Gamma-ray-emitting Active Galactic Nuclei in the Fermi-LAT Unassociated Sample Using Machine Learning". In: *The Astrophysical Journal* 782.1, p. 41.

Domínguez, A. et al. (2011). "Extragalactic background light inferred from AEGIS galaxy-SED-type fractions". In: *Monthly Notices of the Royal Astronomical Society* 410.4, pp. 2556–2578.

Eckle, K., N. Bissantz, H. Dette, and S. Einecke (2017a). "Silverman-based Testing for the Number of Modes in Multivariate and Deconvolution Density Estimation". In: *Submitted*.

Eckle, K., N. Bissantz, H. Dette, K. Proksch, and S. Einecke (2017b). "Multiscale Inference for a Multivariate Density with Applications to X-Ray Astronomy". In: *Annals of the Institue of Statistical Mathematics*.

Efron, B. and T. Hastie (2016). *Computer Age Statistical Inference: Algorithms, Evidence, and Data Science*. 1st. New York, NY, USA: Cambridge University Press.

Einecke, S. for the MAGIC Collaboration (2015). "The MAGIC Telescope System: Scientific Highlights, Status and Future Perspectives". In: *27th Rencontres de Blois Conference*. (Blois, France).

Einecke, S. (2016). "Search for High-Confidence Blazar Candidates and Their MWL Counterparts in the *Fermi*-LAT Catalog Using Machine Learning". In: *Galaxies* 4.3, p. 14.

Einecke, S., D. Elsässer, W. Rhode, and K. Morik (2017). "Towards Refined Population Studies: High-Confidence Blazar Candidates and their MWL Counterparts using Machine Learning". In: *The 35th International Cosmic Ray Conference*. (Bexco, Busan, Korea). 626. Proceedings of Science.

Einecke, S., K. Proksch, N. Bissantz, F. Clevermann, and W. Rhode (2016). "Uncertainty Limits on Solutions of Inverse Problems over Multiple Orders of Magnitude using Bootstrap Methods: An Astroparticle Physics Example". In: *ArXiv e-prints* 1607.07226.

Errando, M. et al. (2013). "Fill-In *Swift*-XRT Observations of Unassociated Gamma-Ray Sources". Application for Observing Time with *Swift*-XRT in 2013/2014.

Evans, P. A. et al. (2014). "1SXPS: A Deep Swift X-Ray Telescope Point Source Catalog with Light Curves and Spectra". In: *The Astrophysical Journal Supplement Series* 210.1, p. 8.

Fanaroff, B. L. and J. M. Riley (1974). "The Morphology of Extragalactic Radio Sources of High and Low Luminosity". In: *Monthly Notices of the Royal Astronomical Society* 167.1, 31P–36P.

Fornasa, M. and M. A. Sánchez-Conde (2015). "The nature of the Diffuse Gamma-Ray Background". In: *Physics Reports* 598. The nature of the Diffuse Gamma-Ray Background, pp. 1–58.

Fossati, G., L. Maraschi, A. Celotti, A. Comastri, and G. Ghisellini (1998). "A unifying view of the spectral energy distributions of blazars". In: *Monthly Notices of the Royal Astronomical Society* 299.2, pp. 433–448.

Franceschini, A. and Rodighiero, G. (2017). "The extragalactic background light revisited and the cosmic photon-photon opacity". In: *Astronomy & Astrophysics* 603, A34.

Gehrels, N. et al. (2004). "The Swift Gamma-Ray Burst Mission". In: *The Astrophysical Journal* 611.2, p. 1005.

Ghisellini, G., C. Righi, L. Costamante, and F. Tavecchio (2017). "The Fermi blazar sequence". In: *Monthly Notices of the Royal Astronomical Society* 469.1, pp. 255–266.

Gregory, P. C., W. K. Scott, K. Douglas, and J. J. Condon (1996). "The GB6 Catalog of Radio Sources". In: *Astrophysical Journal Supplements* 103, p. 427.

Grieder, P. K. F. (2010). *Extensive Air Showers: High Energy Phenomena and Astrophysical Aspects - A Tutorial, Reference Manual and Data Book*. Berlin, Heidelberg: Springer.

Guyon, I. and A. Elisseeff (2003). "An Introduction to Variable and Feature Selection". In: *Journal of Machine Learning Research* 3, pp. 1157–1182.

Hassan, T., N. Mirabal, J. Contreras González, and I. Oya (2012). "Gamma-Ray Active Galactic Nucleus Type through Machine-Learning Algorithms". In: *Monthly Notices of the Royal Astronomical Society* 428.

Hastie, T., R. Tibshirani, and J. Friedmann (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd. New York, NY, USA: Springer New York Inc.

Healey, S. E. et al. (2007). "CRATES: An All-Sky Survey of Flat-Spectrum Radio Sources". In: *Astrophysical Journal Supplements* 171, pp. 61–71.

Helfand, D. J., R. L. White, and R. H. Becker (2015). "The Last of FIRST: The Final Catalog and Source Identifications". In: *The Astrophysical Journal* 801.1, p. 26.

Hillas, A. M. (1985). "Cerenkov light images of EAS produced by primary gamma rays and by nuclei". In: *Proceedings of the 19th International Cosmic Ray Conference*. Ed. by F. C. Jones. Vol. 3. La Jolla, pp. 445–448.

Holder, J. (2012). "TeV gamma-ray astronomy: A summary". In: *Astroparticle Physics* 39.Supplement C. Cosmic Rays Topical Issue, pp. 61–75.

Ingham, J., K.A. Arnaud, and M.F. Corcoran (2006). *The Xselect User's Guide*. 2.4.

Ishio, K. (2016). Picture of the MAGIC Telescopes.

Jackson, N. et al. (2007). "A survey of polarization in the JVAS/CLASS flat-spectrum radio source surveys - I. The data and catalogue production". In: *Monthly Notices of the Royal Astronomical Society* 376, pp. 371–377.

Kawai, H. et al. (2008). "Telescope Array Experiment". In: *Nuclear Physics B - Proceedings Supplements* 175.Supplement C. Proceedings of the XIV International Symposium on Very High Energy Cosmic Ray Interactions, pp. 221–226.

Kneiske, T. M. and Dole, H. (2010). "A lower-limit flux for the extragalactic background light". In: *Astronomy & Astrophysics* 515, A19.

Knödlseder, J. (2016). "The future of gamma-ray astronomy". In: *Comptes Rendus Physique* 17.6. Gamma-ray astronomy / Astronomie des rayons gamma - Volume 2, pp. 663–678.

Landi, R. et al. (2015). "Swift/XRT counterparts to unassociated Fermi high-energy LAT sources". In: *Astronomy & Astrophysics* 581, A57.

Lefaucheur, J. and S. Pita (2017). "Research and characterisation of blazar candidates among the Fermi/LAT 3FGL catalogue using multivariate classifications ". In: *Astronomy & Astrophysics* 602, A86.

Li, T.-P. and Y.-Q. Ma (1983). "Analysis methods for results in gamma-ray astronomy". In: *Astrophysical Journal* 272, pp. 317–324.

Marscher, A. P. (2016). "Variability of Blazars and Blazar Models over 38 Years". In: *Galaxies* 4.37.

Massaro, E. et al. (2009). "Roma-BZCAT: a multifrequency catalogue of blazars". In: *Astronomy & Astrophysics* 495.2, pp. 691–696.

Massaro, F. et al. (2012). "The WISE Gamma-Ray Strip Parameterization: The Nature of the Gamma-Ray Active Galactic Nuclei of Uncertain Type". In: *The Astrophysical Journal* 750.2, p. 138.

Massaro, F. et al. (2013a). "BL Lac Candidates for TeV Observations". In: *The Astrophysical Journal Supplement Series* 207.1, p. 16.

Massaro, F. et al. (2013b). "Unveiling the Nature of the Unidentified Gamma-Ray Sources. III. Gamma-Ray Blazar-like Counterparts at Low Radio Frequencies". In: *The Astrophysical Journal Supplement Series* 207.1, p. 4.

Massaro, F. et al. (2013c). "Unveiling the Nature of Unidentified Gamma-Ray Sources. II. Radio, Infrared, and Optical Counterparts of the Gamma-Ray Blazar Candidates". In: *The Astrophysical Journal Supplement Series* 206.2, p. 13.

Mazin, D. and Raue, M. (2007). "New limits on the density of the extragalactic background light in the optical to the far infrared from the spectra of all known TeV blazars". In: *Astronomy & Astrophysics* 471.2, pp. 439–452.

Meegan, C. et al. (2009). "The Fermi Gamma-ray Burst Monitor". In: *The Astrophysical Journal* 702.1, p. 791.

Mirabal, N., V. Frías-Martinez, T. Hassan, and E. Frías-Martinez (2012). "Fermi's SIBYL: mining the gamma-ray sky for dark matter subhaloes". In: *Monthly Notices of the Royal Astronomical Society: Letters* 424.1, pp. L64–L68.

Myers, S. T. et al. (2003). "The Cosmic Lens All-Sky Survey - I. Source selection and observations". In: *Monthly Notices of the Royal Astronomical Society* 341, pp. 1–12.

Netzer, H. (2015). "Revisiting the Unified Model of Active Galactic Nuclei". In: *Annual Review of Astronomy and Astrophysics* 53.1, pp. 365–408.

Nieto et al., D. (2010). "Unidentified Fermi Objects as Dark Matter Clump Candidates". Application for Observing Time at MAGIC - Cycle 6.

Nolan, P. L. et al. (2012). "Fermi Large Area Telescope Second Source Catalog". In: *The Astrophysical Journal Supplement Series* 199.2, p. 31.

Padovani, P. et al. (2017). "Active galactic nuclei: what's in a name?" In: *The Astronomy and Astrophysics Review* 25.2.

Paggi, A. et al. (2013). "Unveiling the Nature of the Unidentified Gamma-Ray Sources. IV. The Swift Catalog of Potential X-Ray Counterparts". In: *The Astrophysical Journal Supplement Series* 209.1, p. 9.

Paiano, S., A. Stamerra, K. Satalecka, S. Einecke, and M. Doert (2015). "MAGIC observations of Unassociated Fermi Objects". Application for Observing Time at MAGIC - Cycle 11.

Park, N. et al. (2015). "Performance of the VERITAS experiment". In: *The 34th International Cosmic Ray Conference*. (The Hague, The Netherlands). Proceedings of Science.

Parkinson, P. M. Saz et al. (2016). "Classification and Ranking of Fermi LAT Gamma-ray Sources from the 3FGL Catalog using Machine Learning Techniques". In: *The Astrophysical Journal* 820.1, p. 8.

Pedregosa, F. et al. (2011). "Scikit-learn: Machine Learning in Python". In: *Journal of Machine Learning Research* 12, pp. 2825–2830.

Peng, H., F. Long, and C. Ding (2005). "Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27.8, pp. 1226–1238.

Perley, R. A., C. J. Chandler, B. J. Butler, and J. M. Wrobel (2011). "The Expanded Very Large Array: A New Telescope for New Science". In: *The Astrophysical Journal Letters* 739.1, p. L1.

Petrov, L. et al. (2013). "Australia Telescope Compact Array observations of Fermi unassociated sources". In: *Monthly Notices of the Royal Astronomical Society* 432.2, pp. 1294–1302.

Saripalli, L. (2012). "Understanding the Fanaroff–Riley Radio Galaxy Classification". In: *The Astronomical Journal* 144.3, p. 85.

Schinzel, F. K., L. Petrov, G. B. Taylor, and P. G. Edwards (2017). "Radio Follow-up on All Unassociated Gamma-Ray Sources from the Third Fermi Large Area Telescope Source Catalog". In: *The Astrophysical Journal* 838.2, p. 139.

Temme, F. et al. (2015). "FACT - First Energy Spectrum from a SiPM Cherenkov Telescope". In: *The 34th International Cosmic Ray Conference*. (The Hague, The Netherlands). 707. Proceedings of Science.

Thompson, A. R., B. G. Clark, C. M. Wade, and P. J. Napier (1980). "The Very Large Array". In: *The Astrophysical Journal Supplement Series* 44, pp. 151–167.

Urry, C. M. and P. Padovani (1995). "Unified schemes for radio-loud active galactic nuclei". In: *Publ. Astron. Soc. Pac.* 107, p. 803.

Voges, W. et al. (1999). "The ROSAT all-sky survey bright source catalogue". In: *Astronomy and Astrophysics* 349, pp. 389–405.

Wakely, S. P. and D. Horan (2008). "TeVCat: An online catalog for Very High Energy Gamma-Ray Astronomy". In: *International Cosmic Ray Conference* 3, pp. 1341–1344.

Walter, M. et al. (2007). "The IceCube Neutrino Observatory - Design and Performance". In: *Nuclear Physics B - Proceedings Supplements* 172.Supplement C. Proceedings of the 10th Topical Seminar on Innovative Particle and Radiation Detectors, pp. 13–16.

White, R. L. and R. H. Becker (1992). "A new catalog of 30,239 1.4 GHz sources". In: *Astrophysical Journal Supplements* 79, pp. 331–467.

White, R. L., R. H. Becker, D. J. Helfand, and M. D. Gregg (1997). "A Catalog of 1.4 GHz Radio Sources from the FIRST Survey". In: *The Astrophysical Journal* 475.2, p. 479.

Wright, A. E., M. R. Griffith, B. F. Burke, and R. D. Ekers (1994). "The Parkes-MIT-NRAO (PMN) surveys. 2: Source catalog for the southern survey (delta greater than -87.5 deg and less than -37 deg)". In: *Astrophysical Journal Supplements* 91, pp. 111–308.

Wright, E. L. et al. (2010). "The Wide-field Infrared Survey Explorer (WISE): Mission Description and Initial On-orbit Performance". In: *The Astronomical Journal* 140.6, p. 1868.

Author's Publications

Publications which are associated with this thesis and with main contributions by the author of this thesis:

Eckle, K., N. Bissantz, H. Dette, and S. Einecke (2017a). "Silverman-based Testing for the Number of Modes in Multivariate and Deconvolution Density Estimation". In: *Submitted*.

Eckle, K., N. Bissantz, H. Dette, K. Proksch, and S. Einecke (2017b). "Multiscale Inference for a Multivariate Density with Applications to X-Ray Astronomy". In: *Annals of the Institue of Statistical Mathematics*.

Einecke, S. for the MAGIC Collaboration (2015). "The MAGIC Telescope System: Scientific Highlights, Status and Future Perspectives". In: *27th Rencontres de Blois Conference*. (Blois, France).

Einecke, S. (2016). "Search for High-Confidence Blazar Candidates and Their MWL Counterparts in the *Fermi*-LAT Catalog Using Machine Learning". In: *Galaxies* 4.3, p. 14.

Einecke, S., D. Elsässer, W. Rhode, and K. Morik (2017). "Towards Refined Population Studies: High-Confidence Blazar Candidates and their MWL Counterparts using Machine Learning". In: *The 35th International Cosmic Ray Conference*. (Bexco, Busan, Korea). 626. Proceedings of Science.

Einecke, S., K. Proksch, N. Bissantz, F. Clevermann, and W. Rhode (2016). "Uncertainty Limits on Solutions of Inverse Problems over Multiple Orders of Magnitude using Bootstrap Methods: An Astroparticle Physics Example". In: *ArXiv e-prints* 1607.07226.

Errando, M. et al. (2013). "Fill-In *Swift*-XRT Observations of Unassociated Gamma-Ray Sources". Application for Observing Time with *Swift*-XRT in 2013/2014.

Paiano, S., A. Stamerra, K. Satalecka, S. Einecke, and M. Doert (2015). "MAGIC observations of Unassociated Fermi Objects". Application for Observing Time at MAGIC - Cycle 11.

Temme, F. et al. (2015). "FACT - First Energy Spectrum from a SiPM Cherenkov Telescope". In: *The 34th International Cosmic Ray Conference*. (The Hague, The Netherlands). 707. Proceedings of Science.

Acknowledgment

Ich möchte mich ganz herzlich bei meinem Doktorvater Prof. Dr. Dr. Wolfgang Rhode für die Betreuung meiner Dissertation, meiner Bachelor- und Masterarbeiten, und für die unermüdliche Unterstützung und Förderung in den letzten 7 Jahren bedanken. Vielen Dank auch für das Vertrauen und die Ermöglichung unzähliger Teilnahmen an Konferenzen, Schulen und Kollaborationstreffen.

Vielen Dank auch an Prof. Dr. Carsten Westphal für seine Bereitschaft und Mühen als Zweitgutachter dieser Dissertation.

Ein großer Dank geht an die Deutsche Forschungsgemeinschaft (DFG), die meine Promotionsstelle im Rahmen des Sonderforschungsbereichs 823 finanziert hat und an den Sonderforschungsbereich 876, der mir eine Vielzahl an Fortbildungen ermöglicht hat.

Ebenfalls bedanken möchte ich mich beim Deutschen Akademischen Austauschdienst (DAAD), der mir mehrere Forschungsaufenthalte in New York und die Teilnahme an einer Konferenz in Südkorea ermöglicht hat.

Ein ganz herzlicher Dank an die Astroteilchengruppe Dortmund, ihr habt die letzten Jahre sehr viel angenehmer gemacht! Besonders erwähnen möchte ich Dominik Elsässer für seine zahlreichen Ratschläge, seine Unterstützung und als einen angenehmen Bürokollegen. Nicht zu vergessen ist natürlich die gute Seele unseres Lehrstuhls – Andrea Teichmann – der ich für die Hilfe bei allen bürokratischen und nicht-bürokratischen Dingen danken möchte und eine stets gute Laune.

Vielen lieben Dank an meine Famile und meine Freunde für eure Unterstützung in den letzten Jahren und euer Verständnis, wenn ich nicht viel Zeit hatte und gestresst war. Ich gelobe Besserung!

Allerherzlichsten Dank!

I would like to thank the Instituto de Astrofísica de Canarias for the excellent working conditions at the Observatorio del Roque de los Muchachos in La Palma. The financial support of the German BMBF and MPG, the Italian INFN and INAF, the Swiss National Fund SNF, the ERDF under the Spanish MINECO (FPA2015-69818-P, FPA2012-36668, FPA2015-68278-P, FPA2015-69210-C6-2-R, FPA2015-69210-C6-4-R, FPA2015-69210-C6-6-R, AYA2013-47447-C3-1-P, AYA2015-71042-P, ESP2015-71662-C2-2-P, CSD2009-00064, RYC-2013-14660), and the Japanese JSPS and MEXT is gratefully acknowledged. This work was also supported by the Spanish Centro de Excelencia "Severo Ochoa" SEV-2012-0234 and SEV-2015-0548, and Unidad de Excelencia "María de Maeztu" MDM-2014-0369, by the Croatian Science Foundation (HrZZ) Project 09/176 and the University of Rijeka Project 13.12.1.3.02, by the DFG Collaborative Research Centers SFB823/C4 and SFB876/C3, and by the Polish MNiSzW grant 745/N-HESS-MAGIC/2010/0.

This work made use of data supplied by the UK Swift Science Data Centre at University of Leicester.

This research has made use of data obtained through the High Energy Astrophysics Science Archive Research Center Online Service, provided by the NASA/Goddard Space Flight Center.

This publication makes use of data products from the Wide-field Infrared Survey Explorer, which is a joint project of the University of California, Los Angeles, and the Jet Propulsion Laboratory/California Institute of Technology, and NEOWISE, which is a project of the Jet Propulsion Laboratory/California Institute of Technology. WISE and NEOWISE are funded by the National Aeronautics and Space Administration.

This research has made use of the NASA/IPAC Infrared Science Archive, which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

This research used the facilities of the Canadian Astronomy Data Centre operated by the National Research Council of Canada with the support of the Canadian Space Agency.

This research has made use of the SIMBAD database, operated at CDS, Strasbourg, France.

This research has made use of the NASA/IPAC Extragalactic Database (NED) which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

Part of this work is based on archival data, software or online services provided by the Space Science Data Center - ASI.

Eidesstattliche Versicherung

Ich versichere hiermit an Eides statt, dass ich die vorliegende Abschlussarbeit mit dem Titel “The Data Mining Guide to the Galaxy” selbstständig und ohne unzulässige fremde Hilfe erbracht habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt, sowie wörtliche und sinngemäße Zitate kenntlich gemacht. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

Ort, Datum | Unterschrift

Belehrung

Wer vorsätzlich gegen eine die Täuschung über Prüfungsleistungen betreffende Regelung einer Hochschulprüfungsordnung verstößt, handelt ordnungswidrig. Die Ordnungswidrigkeit kann mit einer Geldbuße von bis zu 50 000 € geahndet werden. Zuständige Verwaltungsbehörde für die Verfolgung und Ahndung von Ordnungswidrigkeiten ist der Kanzler/die Kanzlerin der Technischen Universität Dortmund. Im Falle eines mehrfachen oder sonstigen schwerwiegenden Täuschungsversuches kann der Prüfling zudem exmatrikuliert werden (§ 63 Abs. 5 Hochschulgesetz –HG–).
Die Abgabe einer falschen Versicherung an Eides statt wird mit Freiheitsstrafe bis zu 3 Jahren oder mit Geldstrafe bestraft.
Die Technische Universität Dortmund wird ggf. elektronische Vergleichswerkzeuge (wie z. B. die Software “turnitin”) zur Überprüfung von Ordnungswidrigkeiten in Prüfungsverfahren nutzen.

Die oben stehende Belehrung habe ich zur Kenntnis genommen.

Ort, Datum | Unterschrift