CHAPTER 16

Multiwavelength Extragalactic Surveys: Examples of Data Mining

IRINA VAVILOVA, DR • DARIA DOBRYCHEVA, DR • MAKSYM VASYLENKO, MSC • ANDRII ELYIV, DR • OLGA MELNYK, DR

16.1 INTRODUCTION

Multiwavelength astronomy provides a huge amount of data from different sources over the whole electromagnetic spectrum. Recently it has been supplemented with data from neutrino astronomy and gravitational wave astronomy (multimessenger astronomy). Since the end of the 20th century the all-sky and large-area observational surveys as well as their catalogued databases enriched and continue to enrich our knowledge of the Universe. Astronomy has entered the Big Data era, when these data are combined/compiled into numerous archives. Each archive contains the observational data that were obtained in a specific spectral range, for which ground-based or space-born telescopes were developed (see, for example, Chapter 5 in this book and commentaries by Brunner et al. (2002) on the nature of astronomical data). However, modern astrophysics forces to study astrophysical objects across the whole electromagnetic spectrum, because different physical processes make themselves felt at different wavelengths (different substructures of celestial bodies radiate at different wavelengths). An identification of such substructures and processes requires a very wide knowledge from high-energy astrophysics to the radio decameter astronomy, wherein X-ray, gamma ray, and radio sources need to be identified with their optical counterparts.

16.2 THE AUTOMATED MORPHOLOGICAL CLASSIFICATION FOR THE SDSS GALAXIES

Since 2000, the Sloan Digital Sky Survey (SDSS) (York et al., 2000) has collected the most data that have been amassed in the entire history of astronomy.[1]

[1] See Chapter 5 in this book.

Now, its archive contains about 170 terabytes of information, with most of these data about galaxies. Astronomers who are directly involved in the SDSS identified the problem of morphological galaxy classification "as one of the most cumbersome areas in celestial classification, and the one that has proven the most difficult to automate" (Kasivajhula et al., 2007).

Sense of galaxy morphological classification. Such substructures of galaxies as the central region with active nucleus and supermassive black hole, bulge and bar, spiral arms, halo with a dark matter component, star formation regions, intergalaxy medium, jets, disk, rings, and others features compose the important building blocks, which are related to galaxy morphology, dynamics (mass distribution) and kinematics, and have a decisive role in our understanding of galaxy formation and evolution. Galaxy morphological classification on large-scale datasets allows us to reduce classification errors and to improve statistics of the known morphological types of galaxies at different redshifts. A good introduction to the classification algorithms for astronomical tasks, including the machine learning methods for galaxy morphological classification, one can find in works by Buta and McCall (1999), de la Calleja and Fuentes (2004), Feigelson and Babu (2006), Ball and Brunner (2010), Ivezić et al. (2014) as well as in Chapters 12 and 13 of this book.

To imagine better this interesting scientific problem, we remind in Table 16.1 the numerical morphological Hubble stage, which was introduced by G. de Vaucouleurs in 1959 and improved by him later in "Global Physical Parameters of Galaxies" in 1994. Various morphological types of galaxies are illustrated in Fig. 16.1, where images in optical range were taken from the SDSS. It is important to underline that morphological type (T) of a galaxy correlates with stellar population

Knowledge Discovery in Big Data from Astronomy and Earth Observation. https://doi.org/10.1016/B978-0-12-819154-5.00028-X
Copyright © 2020 Elsevier Inc. All rights reserved.

TABLE 16.1
Numerical Hubble stage.

Hubble stage T	−6	−5	−4	−3	−2	−1	0	1	2
de Vaucouleurs class	cE	E	E^+	$S0^-$	$S0^0$	$S0^+$	$S0/a$	Sa	Sab
Hubble class		E			$S0$		$S0/a$	Sa	$Sa-b$
Hubble stage T	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	
de Vaucouleurs class	Sb	Sbc	Sc	Scd	Sd	Sdm	Sm	Im	
Hubble class	Sb	$Sb-c$		Sc		$Irr I$	$Irr I$		

and star formation history, bulge/disk luminosity ratio, mass concentration, interstellar media (chemical abundance), and nuclear activity properties.

A very good pedagogical review with discussion of the major methods in which galaxies are studied morphologically and structurally is given by Conselice (2014) that "includes the well-established visual method for morphology; Sercic fitting to measure galaxy sizes and surface brightness profile shapes; non-parametric structural methods including the concentration (C), asymmetry (A), clumpiness (S) (CAS method), the Gini/M_{20} parameters, as well as newer structural indices." We remind that concentration C index (the logarithmic ratio of the radii containing 90% and 50% of the light, R_{90}/R_{50}) is tightly related to morphology, A (asymmetry) to merging, and S to star formation rate (CAS parameters).

Visual and automated classification. Notwithstanding, below we would like to underline briefly several works, where different approaches were developed and great efforts were made to identify the morphological types of galaxies, first of all, from the SDSS, in the visual and/or the automated modes. We note that many machine learning methods were actively involved to disentangle this problem.

During the 1990s, the artificial neural network (ANN) algorithm was widely used for automatic morphological classification of galaxies since the very large extragalactic datasets have been constructed. The classification accuracy (or the success rate) of the ANN ranged from 65% to 90%, depending on the mathematical subtleties of the applied methods and the quality of galaxy samples. One of the first such works was made by Storrie-Lombardi et al. (1992) with a feedforward neural network and dealt with classification of 5217 galaxies onto five classes (E, SO, Sa-Sb, Sc-Sd, and Irr) with a 64% accuracy. A detailed comparison of human and neural classifiers was presented by Naim et al. (1995), who used principal component analysis to test various architectures to classify 831 galaxies: the best result was obtained with an r.m.s. deviation of 1.8 T-types. Summarizing the first attempts, Lahav et al. (1996) concluded that "the ANNs can replicate the classification by a human expert almost to the same degree of agreement as that between two human experts, to within 2 T-type units" (see Table 16.1).

Later de la Calleja and Fuentes (2004) developed a method which combines two machine learning algorithms: Locally Weighted Regression and ANN. They tested it with 310 images of galaxies from the NGC catalogue and obtained an accuracy of 95.11% and 90.36%, respectively. Ball et al. (2004) using the supervised ANN derived that it may be applied without human intervention for the SDSS galaxies (correlations between predicted and actual properties were around 0.9 with r.m.s. errors of order 10%). Andrae et al. (2010) classified the SDSS bright galaxies with a probabilistic classification algorithm and obtained that it produces reasonable morphological classes and object-to-class assignments without any prior assumptions.

As for the visual morphological classification conducted during the last years, we note a very powerful study by Banerji et al. (2010), where galaxies classified by the Galaxy Zoo Project into three classes (early types, spirals, "spam" objects) have formed a training sample for morphological classifications of galaxies in the SDSS DR6 (http://data.galaxyzoo.org). These authors convincingly showed that using a set of certain galaxy parameters, the neural network is able to reproduce the human classifications to better than 90% for all these classes and that the Galaxy Zoo catalogue (GZ1) can serve as a training sample.

Totally, hundreds of thousands of volunteers were involved into the Galaxy Zoo project to achieve visual classification of a million galaxies in the SDSS. Most of their results have found good scientific application. For

FIG. 16.1 Galaxies of different morphological types from the SDSS sample.

example, using the raw imaging data from the SDSS that was available in the GZ1, and the hand-picked features from the SDSS, Kates-Harbeck (2012) applied a logistic regression classifier and attained 95.21% classification accuracy. Willett et al. (2013) issued a new catalogue of morphological types from the Galaxy Zoo Project (GZ2) in synergy with the SDSS DR7, which contains more than 16 million morphological classifications of 304,122 galaxies and their finer morphological features (bars, bulges, and the shapes of edge-on disks, as well as parameters of the relative strengths of galactic bulges and spiral arms). Another approach was developed by Nair and Abraham (2010), who prepared the detailed visual classifications for 14,034 galaxies in the SDSS DR4 at $z < 0.1$, which can be used as a good training sample for calibrating the automated galaxy classification algorithms. A morphology catalogue of the SDSS galaxies was generated with the Wndchrm image analysis utility and nearest neighbor classifier by Kuminski and Shamir (2016). These authors pointed out that about 900,000 of the instances classified as spiral galaxies and about 600,000 of those classified as elliptical galaxies have a statistical agreement rate of about 98% with the Galaxy Zoo classification.

"One approach to automated classification is to ask what set of analytic or empirical components (bulge, disk) best represent a galaxy's detected image, and what the expected errors (say in the χ^2 sense) are. The limitation here is that even in perfectly ordinary galaxies, the fitted forms for these components vary, and many galaxies have images that overlap with neighbors or are dotted with brilliant star-forming regions. A quite different approach is taken by neural-network schemes. Here, one defines a set of input values based on the galaxy image, and trains the code using a large set of galaxies classified by eyeball (usually by several sets of eyeballs for a consistency check). The code then finds the set of hidden connections needed to give these outputs, and can apply this mapping to any further data desired. This is thought to be an analog of what the human brain does in learning to recognize patterns, though working backwards, it is not particularly clear just what the code is responding to in the image, except that it looks most like the typical image that it was taught to classify in this way. Neural net classifiers seem to be statistically about as good as human ones, which is especially impressive if one considers that people may fold in all sorts of outside knowledge as to redshifts and pass bands in their estimates" (Keel, 2007).

Recently Murrugarra and Hirata (2017) evaluated the convolutional neural network to classify galaxies from the SDSS onto two classes as ellipticals/spirals using their images and achieved an accuracy around 90%–91%. Using convolutional neural networks, especially the inception method, Wahaono and Azhari (2018) conducted classification into three general categories: ellipticals, spirals, and irregulars. They used 710 images (206 E, 320 Sp, 184 Irr) and obtained that images which underwent image processing showed a rather poor testing accuracy compared to not using any form of image processing. Their best testing accuracy was 78.3%. Both supervised and unsupervised methods were applied by Jain et al. (2016) to study the Galaxy Zoo dataset of 61,578 preclassified galaxies (spiral, elliptical, round, disk). They found that the variation of galaxy images is correlated with brightness and eccentricity, the random forest method gives a best accuracy (67%), meanwhile its combination with regression to predict the probabilities of galaxies associated with each class allows to reach a 94% accuracy.

Examples on binary and ternary morphological classification of the SDSS galaxies. Let us use the well-known fact that galaxy morphological type is correlated with the color indices, luminosity, de Vaucouleurs radius, inverse concentration index (R_{50}/R_{90}), etc. For example, let us combine visual classification and the two-dimensional diagrams of color indices g–i and one of the aforementioned parameters as "color-absolute magnitude," "color-inverse concentration index," "color-de Vaucouleurs radius," and "color-scale radius" for each galaxy with redshifts $0.02 < z < 0.06$, visual $m_r < 17.7$, and absolute $-24^m < M_r < -17^m$ magnitudes from the SDSS DR9. Photometric and spectral parameters of each object as well as their images are available through the SDSS web site. As a result, we can discover possible criteria for separating the galaxies into three classes (Melnyk et al., 2012), i.e., early types (E) – elliptical and lenticular; spiral (S) – $Sa - Scd$ types; late spiral (LS) – $Sd - Sdm$ types and irregular Im/BCG galaxies. One can see in Fig. 16.2 that the "color indices vs inverse concentration indexes" diagrams allow making a ternary morphological galaxy classification with a good accuracy (98% for E, 88% for S, and 57% for LS classes). The combinations of (1) color indices g–i and inverse concentration index $R50/R90$ and (2) color indices g–i and absolute magnitude M_r gives the best result: 143,263 E class, 112,578 S class, 61,177 LS class for the sample of the SDSS galaxies at $z < 0.1$ (Dobrycheva et al., 2017).

We can apply different machine learning methods[2] for providing a binary automated morphological classification for the same sample of the SDSS DR9 galaxies as in the above case of photometric diagrams. Why is it binary one? Because (1) S and LS classes of galaxies could be considered as one class of the late type galaxies L at the Hubble stage and (2) an accuracy for classification of late spirals LS was low enough (57%).

The first step is to prepare a training galaxy sample based on the SDSS DR9 and selected randomly with different redshift and luminosity from the total sample for the following visual classification. The second step is a training of the classifier. With this aim, we can use the absolute magnitudes M_u, M_g, M_r, M_i, and M_z, all the kinds of color indices M_u–M_r, M_g–M_i, M_u–M_g, M_r–M_z, and inverse concentration indexes $R50/R90$ to the center in each photometric band.

Using our own code in Scikit Learn Python[3] to predict correctly the galaxy morphology (late and early types) we verified several machine learning methods

[2] See, for example, Chapter 12 in this book.
[3] https://scikit-learn.org/.

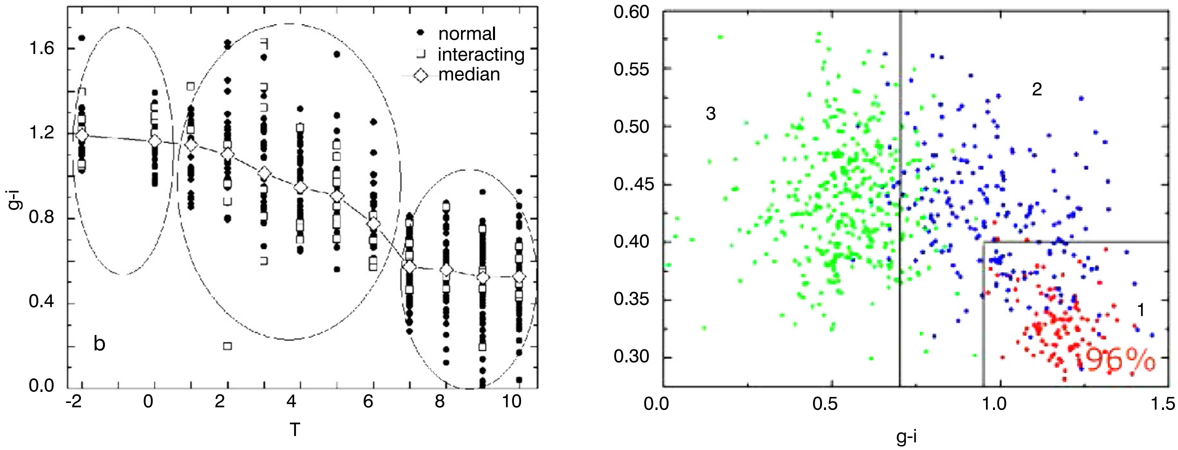

FIG. 16.2 (Left) The dependence of the morphological types T on the color indices g–i for 730 galaxies from the SDSS DR5. (Right) The inverse concentration index R50/R90 as functions of color indices for these galaxies; the red circles correspond to early types (−2 to 0), the blue circles to spirals (1–6), and the green circles to late type spiral and irregular galaxies (see, also, Table 16.1). The lines define regions into which a maximum (more than 90%) number of galaxies of morphological types (−2 to 0), (1 to 6), and (7 to 10), respectively, fall (or with a minimum number of the missclassified morphological types.

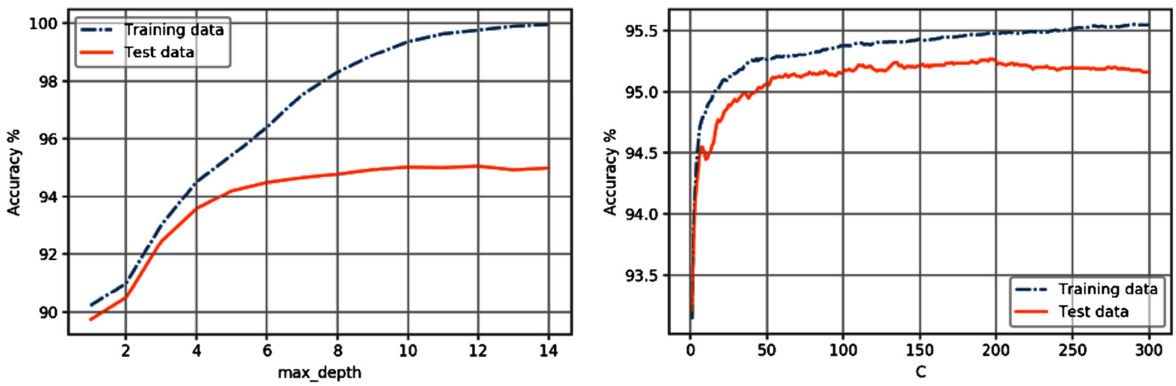

FIG. 16.3 Dependence of prediction accuracy for different machine learning methods of the automated morphological classification with training SDSS galaxy sample: (left) for the random forest classifier on the parameter “max depth”; (right) for the support vector machine classifier on the “C” parameter.

for binary morphological classification of the SDSS galaxies. With this aim we used the sample of 60,561 galaxies from the SDSS DR9 survey with a redshift of $0.02 < z < 0.06$ and absolute magnitudes of $-24^m < M_r < -19.4^m$. Among the machine learning methods were the following: naive Bayes, random forest, support vector machines, logistic regression, and the k-nearest neighbor algorithm. Prediction accuracy was evaluated for each of these methods for the training galaxy sample and reached the following values (all the above-mentioned classifiers include the k-fold cross-validation method):

naive Bayes classifier: 0.89 (E – 0.92, L – 0.82) $\pm$ 0.01;

k-nearest neighbors classifier: 0.945 (E – 0.9389, L – 0.958) $\pm$ 0.006;

logistic regression classifier: 0.949 (E – 0.968, L – 0.911) $\pm$ 0.006;

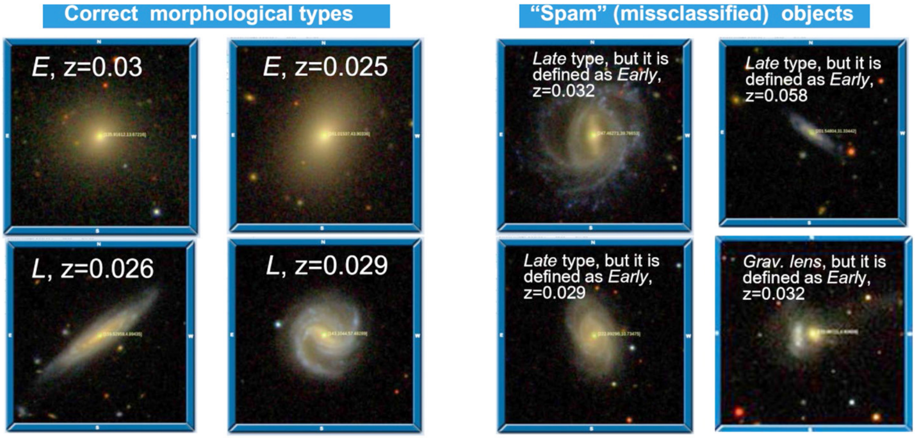

FIG. 16.4 Images of SDSS galaxies. Left: With correctly classified morphology. Top: Early type. Bottom: Late type. Right: with the misclassified morphology. Top and left bottom: Late types, which are classified as early types. Right bottom: gravitational lens classified as early type galaxy.

random forest classifier: 0.955 (E – 0.967, L – 0.928) ± 0.003;

support vector machine classifier: 0.964 (E – 0.961, L – 0.969) ± 0.006.

It turned out (Fig. 16.3) that the methods of random forest and support vector machine provide the highest accuracy (Vasylenko et al., 2019). Examples of images of galaxies with a correct classification on the early and late types are given in Fig. 16.4 (left panel).

The problem points arise when we have cases of the face-on and edge-on galaxies (Fig. 16.4, right panel). Most of these galaxies are misclassified as elliptical galaxies (early type). The good thing is that this approach allow us to recover gravitational lenses (point-like sources, arcs) and most of such misclassifications are also among elliptical galaxies. So, we have overestimated the number of elliptical and underestimated the number of spiral galaxies (about 10%).

But this problem can be solved when we form training samples through several steps (pretraining, fine-tuning, and classification). The step of fine-tuning should include the limitations on the axes-ratio for elliptical galaxies and additional photometry parameters for the face-on galaxies, as well as trainings with images and spectral features of galaxies which requires a specific algorithm with deep learning methods.

The distribution of the SDSS galaxies at $0.02 < z < 0.1$ with the automated morphological binary classification (early and late types) is given in Fig. 16.5.

Last remarks. The machine learning methods are indispensable assistants in solving morphological classification since their first application to tackle this problem with the ANN algorithm (Storrie-Lombardi et al., 1992). They are also effective for reconstruction of the Zone of Avoidance, distance modulus for local galaxies, gravitational lenses search, and other important tasks. The race in accuracy of machine learning methods leads to the search for the most effective among them and to the selection of the most reliable galaxy parameters (photometry, spectra, images), which can be used to determine galaxy morphology. Note that the diversity of the morphological types (Hubble stage, optical range), which we discussed in this subsection, is observed at redshifts $z < 3$; at larger redshifts the other approaches and algorithms should be applied. Also, another frontier in classification problems is approaching when we consider galaxies in the ultraviolet, infrared, or radio ranges (see, for example, Buta et al., 2010; Bell and Salim, 2011; Banfield et al., 2015; Smith and Donohoe, 2019), when their parameters and images should be complemented or cross-matched with optical counterparts.

16.3 ZONE OF AVOIDANCE OF THE MILKY WAY

The data incompleteness in dependence on the wavelength at which galaxies are sampled says that there are important problems in the sky area obscured by our galaxy. This sky area is the so-called Zone of Avoidance (ZoA) of the Milky Way (see, for example, Fig. 16.5, where several large-scale structures as the SDSS Great Wall, SDSS voids (Mao et al., 2016), CfA2 Great Wall, Great Attractor, and the Zone of Avoidance are pointed

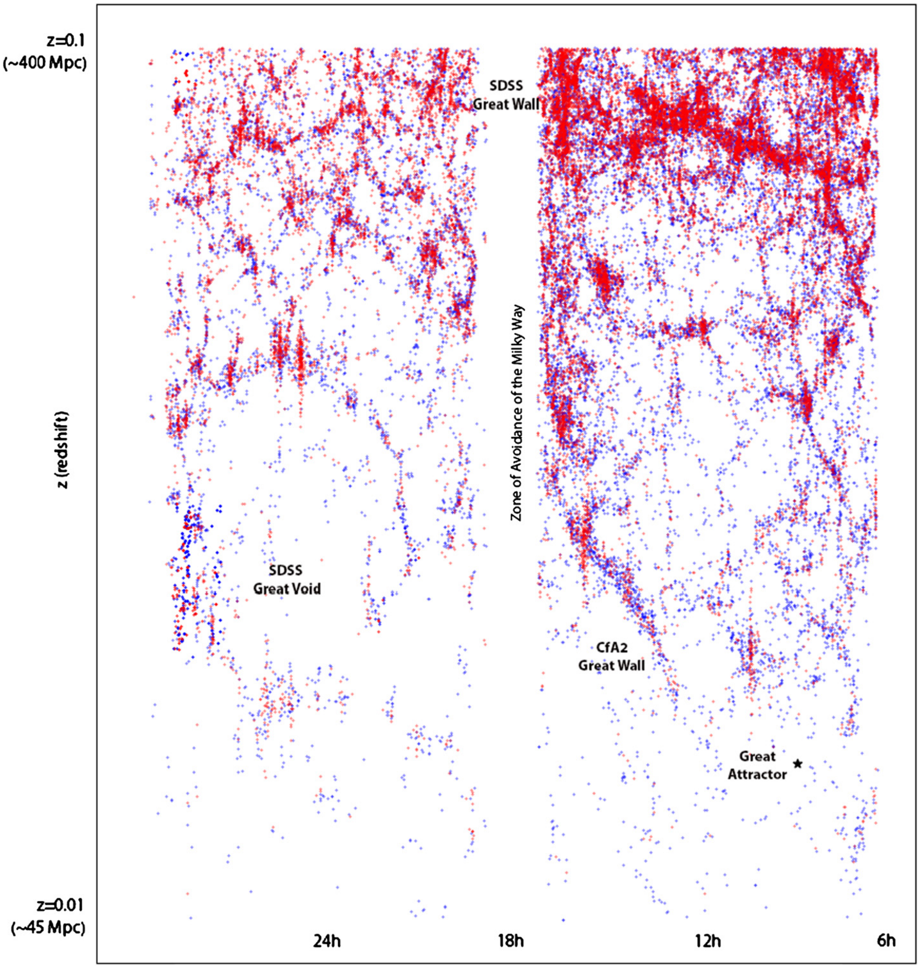

FIG. 16.5 Distribution of 48,651 galaxies from the SDSS DR9 at $0.02 < z < 0.1$ ($\delta = \pm 5°$) with determined morphological classification. Red dots, elliptical galaxies and lenticulars ($E - S0a$, N = 24916); blue dots, spirals, late spirals, and irregulars ($Sa - Irr$, N = 23735). Several large-scale structures (SDSS Great Wall, SDSS voids, CfA2 Great Wall, Great Attractor) and the Zone of Avoidance of the Milky Way are pointed out. At distances of more than 200 Mpc, the Universe becomes "gradually" homogeneous and isotropic.

out). "Why is it of interest to know the galaxy distribution behind the Milky Way, and why is it not sufficient to study galaxies and their large-scale distribution away from the foreground "pollution" of the Milky Way? To understand the dynamics in the nearby Universe and answer the question whether the dipole in the Cosmic Microwave Background (CMB) and other velocity flow fields (e.g. towards the Great Attractor) can be fully explained by the clumpy galaxy/mass distribution, whole-sky coverage is essential" (cited by Kraan-Korteweg and Lahav, 2000).

Brief history. The English astronomer Proctor (1878) firstly noted the Zone of Avoidance of the Milky Way as a Zone of Few Nebulae. Later, in papers by Stratonoff (1900), Easton (1904), Sanford (1917), Charlier (1922) and other authors, who used mostly isopleths as a cosmographic method (contour maps "number of galaxies per the sky area"), the presence of this zone in the distribution of galaxies became obvious. A first definition of the Zone of Avoidance was proposed by Shapley (1961) as the region delimited by "the isopleth of five galaxies per square degree from the Lick and Harvard surveys" (compared to a mean of 54 galaxies/square degree found in unobscured regions by Shane and Wirtanen, 1967).

Due to the incomplete sampling in the area of absorption, on the basis of which the velocity field is constructed, we cannot determine its homogeneity, which gives an error in the definite direction of motion of our Galaxy by this method. We can assume that there are a significant number of galaxies in this zone (Fig. 16.5) based on discrepancy between the vectors of movement of galaxies of the Local Group relative to the coordinate system associated with the cosmic microwave background (CMB) radiation. The Zone of Avoidance is also heterogeneous because the Solar System is not located in the center of our Galaxy.

Due to the small number of known objects, decreasing the brightness of the extragalactic objects when we approach the galactic equator, increasing the concentration of stars on the line of sight, which results in increasing the overlap of the extragalactic object with the star, the extragalactic astronomers usually avoid this area (Kraan-Korteweg and Lahav, 2000).

The problem can be solved by either direct or indirect techniques. Under direct methods we understand the observations of whole-sky surveys in different spectral ranges near the galactic equator ($b \in [-20°, +20°]$). Indirect methods consist in applying the mathematical simulation and data mining methods to fill the Zone of Avoidance as well as to determine the gravitational potentials of the nearest galaxies in order to predict the positions of galaxies and galaxy systems in the area of Milky Way absorption. Great attention is also focused on the machine learning technique.

Multiwavelength observations. Direct methods. Since the 1970s the Zone of Avoidance has decreased significantly due to studies in the infrared and radio spectral ranges (due to the decrease in the amount of light absorption with increasing wavelength, the Zone of Avoidance becomes more transparent in these spectral ranges).

First of all, on 29 September 1967, Italian astronomer P. Maffei discovered the elliptical galaxy Maffei 1 together with the spiral galaxy Maffei 2 in the Zone of Avoidance. He used a hypersensitized I-N photographic plate for the infrared range and exposed it with the Schmidt telescope at Asiago Observatory (see the paper by Maffei, 2003 for a review of his own works). Maffei 1 is located 0.55° from the galactic plane in the middle of the Zone of Avoidance ($\alpha = 02^h\ 36^m\ 35.4^s$, $\delta = +59°\ 39'\ 19''$, $m = 11.14 \pm 0.06$ in the V-band). Maffei 1 would be one of the largest and brightest elliptical galaxies in the sky (about 3/4 the size of the full moon) if there were no 4.7^m of extinction (a factor of about 1/70) in the visible range (Fig. 16.6).

FIG. 16.6 Galaxies Maffei 1 (down right) and Maffei 2 (top left), discovered by P. Maffei in the Zone of Avoidance.

Maffei's discovery promoted a lively discussion in those times about possible membership of these galaxies to the Local Group. In 1970 Spinrad suggested that Maffei 1 is a nearby heavily obscured giant elliptical galaxy and estimated the distance to Maffei 1 as 1 Mpc (Local Group member?). In 1983 this estimate was revised up to $2.1^{+1.3}_{-0.8}$ Mpc by Buta and McCall (Maffei 1 is outside the Local Group!). In 2001, Davidge and van den Bergh used adaptive optics to observe the brightest AGB stars in Maffei 1 and concluded that the distance is $4.4^{+0.6}_{-0.5}$ Mpc. A latest determination of the distance to Maffei 1 is 2.85 ± 0.36 Mpc, which is based on the recalibrated luminosity/velocity dispersion relation for E-galaxies and the updated extinction. It proves that Maffei 1 is a key member of a nearby galaxy group named Maffei Group, where among other members are the giant spiral galaxies IC342 and Maffei 2. Maffei 1 has also a small satellite spiral galaxy Dwingeloo 1 as well as a number of dwarf satellites like MB1. The IC 342/Maffei Group is one of the closest galaxy groups to the Milky Way (Huchtmeier et al., 1995; Karachentsev et al., 2003). The larger (≥ 3 Mpc) distances reported in

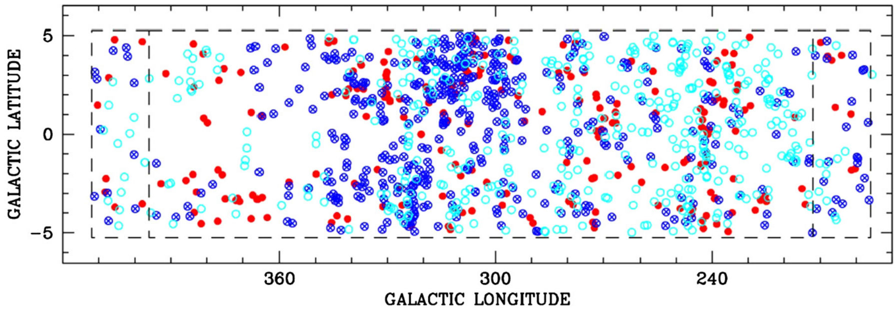

FIG. 16.7 Distribution in galactic coordinates of the 1036 galaxies detected in the deep HI ZOA survey. Open circles, $V_{hel} < 3500$; circled crosses, $3500 < V_{hel} < 6500$; filled circles, $V_{hel} > 9500$ km/s (Kraan-Korteweg et al., 2003, Open Astronomy).

the past 20 years would imply that Maffei 1 has never been close enough to the Local Group to significantly influence its dynamics.

The current notion of the Zone of Avoidance has changed in the 1990s and was connected with exploration of the infrared satellite IRAS and the releases of 2MASS survey as well as with several projects in the radio range. If it was previously believed that this area closes an observer about 20% of the spatial distribution of galaxies in the optical range, which leads to an incomplete catalogue of galaxies near the Galactic Plane, then this value is now about 10%. Completeness of Zone of Avoidance galaxy catalogues as a function of the foreground extinction is as follows: optical Zone of Avoidance surveys are complete to an apparent diameter of $D = 14''$, where the diameters correspond to an isophote of 24.5 mag/arcsec^2 for extinction levels less than $A_B = 3.0^m$.

Because of the transparency of the galaxy to the 21 cm radiation of neutral hydrogen, systematic HI-surveys are particularly powerful in mapping large-scale structures in this part of the sky. The redshifted 21 cm emission of HI-rich galaxies are readily detectable at lowest latitudes and highest extinction levels and the signal will furthermore provide immediate redshift and rotational velocity information. Observations of the neutral hydrogen (21 cm) in the frame of the DOGS project revealed the Dwingeloo 1 (Kraan-Korteweg et al., 1994) and Dwingeloo 2 (Burton et al., 1996) galaxies in this zone (see, for example, Huchtmeier et al., 1995; Buta and McCall, 1999; Karachentsev, 2005 on the estimates of their kinematic and dynamic parameters). Supplementary to these surveys, the Parkes Multibeam HI ZOA Survey as a systematic deep blind HI survey of the Southern Milky Way was begun in 1997 with the Multibeam receiver at the 64 m Parkes telescope. Surveys were centered on the Southern Galactic Plane: $196° \leq l \leq 52°$, $|b| \leq 5°$. The coverage in redshift space was $-1200 < V_{hel} < 12700$ km/s (see, for example, Saurer et al., 1997). Distribution of the 1036 galaxies in galactic coordinates detected in the deep HI ZOA survey is shown in Fig. 16.7.

It should be noted that the absence of a signal does not always indicate the absence of a galaxy, but may be associated with a low HI content (Lahav et al., 1998). This method is slow and requires a lot of time, but the conjunction of HI surveys and 2MASS will greatly increase the current census of galaxies hidden behind the Milky Way. In 2000, Jarrett et al. (2000) reported on the detection of newly discovered sources from 2MASS. There were also identification results of the HI spectra of galaxies which were observed by the IRAS (Lu et al., 1990).

The Milky Way is transparent to the hard X-ray emission above a few keV, and because the rich clusters are strong X-ray emitters. Since the X-ray luminosity is roughly proportional to the cluster mass as $L_X \propto M^{3/2}$ or M^2, depending on the still uncertain scaling law between the X-ray luminosity and temperature (see, for example, Babyk and Vavilova, 2012; Babyk and Vavilova, 2013; Babyk and Vavilova, 2014 and references therein), the massive clusters hidden by the Milky Way should be gravitationally stable through their X-ray emission (Kocevski et al., 2004; Ebeling et al., 2001). The clusters are primarily composed of early-type galaxies, which are not recovered by infrared galaxy surveys or by systematic HI surveys; that is why the method is particularly interesting (see Fig. 16.8).

The inhomogeneous distributed mass of matter in the Zone of Avoidance surrounding the Local Group may cause the unbalanced gravity toward the Local Group in one direction. Despite the fact that the result-

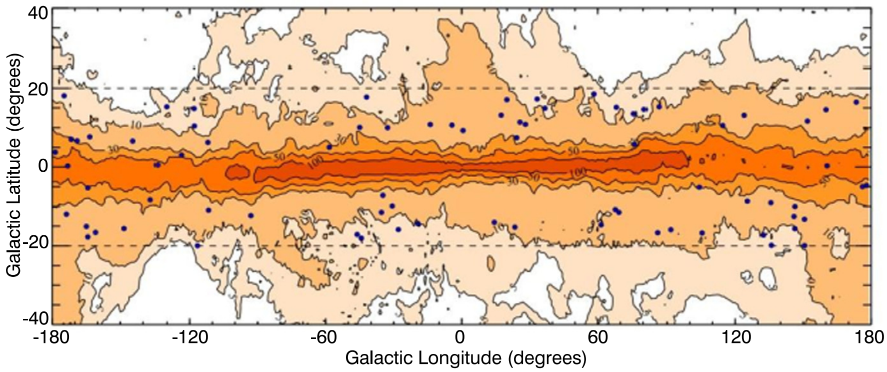

FIG. 16.8 Distribution in galactic coordinates of the 76 by Ebeling et al. (2002) so far spectroscopically confirmed X-ray clusters (solid dots) of which 80% were previously unknown. Superimposed are galactic HI column densities in units of 10^{20} cm^{-2} (Dickey and Lockman, 1990). Note that the region of relatively high absorption ($N_{HI} > 5 \times 10^{21}$ cm^{-2}) actually is very narrow and that clusters could be identified to very low latitudes (Kraan-Korteweg and Lahav, 2000, Open Astronomy).

ing vector of velocity of the Local Group lies within 20° of the observed cosmic background dipole, the calculations remain ambiguous (Karachentsev et al., 2013; Kashibadze et al., 2018), partly because galaxies in the Zone of Avoidance are not taken into account (Vavilova, 2000; Erdoğdu and Lahav, 2009).

A dipole known by CBM studies is the asymmetry of the radiation temperature. It is the heating of 0.1% of CMB radiation in comparison with the average in one direction and in the same cooling in the opposite direction. The COBE (1989–1990) studies indicated that the Milky Way and the Local Group are moving at a velocity $\sim V_p = 627$ km/s to ($l = 276°$, $b = 30°$), towards the Hydra constellation (Kogut et al., 1993). This motion determines the distribution of matter M_i in the Local Group and the cosmological parameter Ω_0 (Giovanelli and Haynes, 1989): $\vec{V}_p \propto \frac{\Omega_0^{0.6}}{b} \sum_i \frac{M_i}{r_i^2} r_i$. Filling the zone $|b| \leq 20°$ by galaxies changes the direction of movement measured in the volume of 6000 km/s by 31° (Kolatt and Dekel, 1997; Vasylenko and Kudrya, 2017). Nearby unknown galaxies in the Zone of Avoidance can make a larger contribution to the definition of a vector of collective velocity than whole clusters over long distances: $\vec{V}_p \propto \sum_i 10^{-0.4m} r_i$.

This discrepancy between the direction on the dipole and the expected velocity vector made it necessary to introduce the concept of "attractors" (the Great Attractor at a distance of about 60 Mpc; see also Fig. 16.5). Perseus-Pisces and the Great Attractor overdensity lie at similar distances on opposite sides of the Local Group and are partially obscured by the Zone of Avoidance. The Zone of Avoidance is fully incomplete at low galactic latitudes in the larger Galactic Bulge area ($l \approx 0° \pm 90°$). Even if the obscured galaxies can be identified, it is very difficult to determine their redshifts because of the higher extinction levels. Since the method involves uniform filling of the sky by the galaxies of the field, and chaotic filling them with nonreal objects leads to the formation of nonexistent fields, attempts to solve the problem of the incompatibility of the vector apex motion of the Local Group determined by the CMB and the velocity field did not give a positive result.

So, the multiwavelength surveys of the Zone of Avoidance in the last decades were aimed at addressing such key problems as the cosmological questions about the dynamics of the Local Group, the possible existence of nearby hidden massive galaxies, the dipole determinations based on luminous galaxies, the continuity and size of nearby superclusters, and the mapping of cosmic flow fields (a very comprehensive review is given by Kraan-Korteweg and Lahav, 2000).

Machine learning. Indirect methods. The solution of these problems is possible also by indirect methods, which include the methods of signal processing applied to obscured and incomplete data; indirect estimates of averaged variables; the mask inversion using Wiener filtering in spherical harmonic analysis; reconstruction of the projected galaxy distribution in infrared, radio, and X-ray spectral ranges; two-dimensional Wiener reconstruction to three dimensions; methods of Voronoi mosaic, cluster, and fractal analysis; and machine learning techniques.

The last successful results of analysis of the spatial distribution of galaxies and their systems in the areas surrounding the Milky Way Zone of Avoidance based on the 2MASS Tully–Fisher Survey and the HI observational surveys are presented in works by Said et al. (2014, 2016b, 2016a), where the optimized Tully–Fisher relation for measured distances and peculiar velocities is developed for dust-obscured galaxies. But it remains a complex and unresolved problem, as well as the estimation of the "invisible" content of the spatial galaxy distribution

The problem of Zone of Avoidance reconstruction is related to dealing with gaps in the spectroscopic observations to restore homogeneous sky coverage. Classical three-dimensional reconstruction of the extragalactic objects behind the Milky Way to preserve the coherence of the large scale structure was triggered by the search of the Great Attractor in the 1990s (Kraan-Korteweg, 2005). And reconstruction of missing information could be oriented towards the observations of galaxies and their systems that surround the Zone of Avoidance (Courtois et al., 2012; Sorce et al., 2017).

The existence of unobserved zones in scale comparable to the size of investigated zones can have a serious impact on the study of galaxy properties and local environments. In this case, the local and deterministic recovery of the missing data is needed (Cucciati et al., 2006). For small-scale reconstruction techniques such as the following are common: direct cloning (Elyiv, 2006), wavelet analysis (Vavilova, 1997), cluster analysis (Gregul et al., 1991; Vavilova and Melnyk, 2005), randomized cloning of objects into unobserved areas or application of Wiener filtering (Lahav et al., 1994; Branchini et al., 1999), Voronoi tessellation (Melnyk et al., 2006; Elyiv et al., 2009; Dobrycheva et al., 2014). Cucciati et al. (2014) proposed two algorithms that use photometric redshift of target objects and assign redshifts based on the spectroscopic redshifts of the nearest galaxies. A Wiener filter applied in this work was very efficient also to reconstruct the continuous density field instead of individual galaxy positions. These methods can clearly separate underdense from overdense regions on scales of 5 h^{-1} Mpc at moderate redshifts $0.5 < z < 1.1$, which is important for studies of cosmic variance and rare population galaxy systems.

There are limits of optical observations of extended objects due to random and systematic noise from detector, the telescope system, and the sky background. Schawinski et al. (2017) estimated a possibility to recover artificially degraded images with a high noise using state-of-the-art methods of machine learning, namely, deep learning – generative adversarial networks (GANs). It works better than simple deconvolution.

Generative adversarial neural networks as the type of unsupervised machine learning algorithms were first invented by Goodfellow et al. (2014). The main idea of these classes of algorithms are two neural networks contesting with each other. First, a neural network called "generative" (typically a deconvolutional one) generates candidate images and a second neural network (a convolutional discriminative one) evaluates them. The generative network trains to transfer from a space of features to a particular data distribution. At the same time the discriminative network discriminates between the produced candidates and real examples. Schawinski et al. (2017) applied the GAN to 4550 galaxies from the SDSS DR12. The authors have proved that this method can reliably recover features in images of galaxies and can go well beyond the limitation of deconvolutions. As the training sample they used image pairs: one original image of a galaxy and the same image artificially degraded (convolved with PSF). In general, the GAN learns how to recover the degraded image by minimizing the difference between the recovered and true images. With this purpose, the authors used a second neural network, whose aim was to distinguish the synthetic recovered image from the true image. These two neural networks are trained simultaneously. Therefore, by training on higher-quality images, the GAN method can learn how to recover information from the lower-quality data by building priors. Such approach has a potential for recovering partially damaged images with gaps and dead CCD chips. The algorithm of reconstruction of three-dimensional structures behind the Zone of Avoidance with the modified GAN method is presented in Fig. 16.9 and described by Vavilova et al. (2018).

We have just one unique sample of galaxies, i.e., just one set for training, which is a principal problem. In the approach described above we cannot use a set of many images for training. One solution could be to prepare the mock catalogues from numerical simulations, which reproduce a target sample. In this case we may generate as many pairs as possible – a real survey and a survey with Zone of Avoidance. Additionally, the position of the Zone of Avoidance could be randomized over the survey field. A goal of generative ANN will be to generate galaxy distributions and their properties in the Zone of Avoidance from a latent space of features. At the same time, a discriminative network will compare the obtained survey with the real one and evaluate how realistic it is. The generative network produces better surveys with iteration, while the discriminative

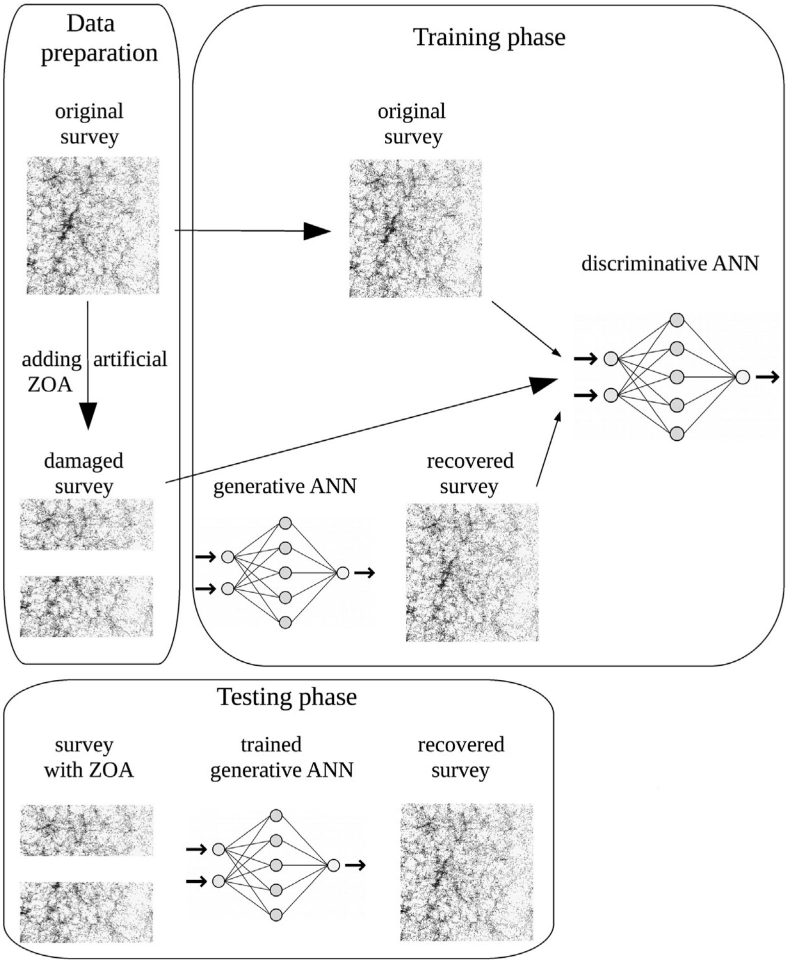

FIG. 16.9 Scheme of the data preparation, the training and testing phases for the Zone of Avoidance recovering by the GAN method. The input is a set of mock surveys from which the artificial Zone of Avoidance was generated to train the GAN. A generative ANN is used to recover surveys in the at the testing phase (Vavilova et al., 2018).

one becomes more experienced at labeling the synthetic ones. In such a way the system learns the sophisticated loss functions automatically without its predefinition.

To apply the algorithm, we should prepare a sample of galaxies surrounding the Zone of Avoidance, which is complete by stellar magnitudes. To get a three-dimensional spatial distribution of galaxies in this sample, we must obtain their photometric redshifts and to divide this sample on the slices by coordinates, taking into account the cosmological parameters. Each of these slices should contain a real distribution and the damaged image (part of the Zone of Avoidance region), which will require darning. The preliminary step how the algorithm works and restores a galaxy distribution should be conducted and tested with subsamples of real galaxies selected from the nondamaged regions.

16.4 FLUX VARIABILITY OF THE BLAZAR 3C 454.3

Another good example for illustration of the data mining from multiwavelength astronomical databases is related to the active galactic nuclei (AGNs). We explain it with the blazar 3C 454.3 (see, all-sky view taken with Fermi/LAT in Fig. 5.16), which is one of the brightest AGNs at all frequencies.[4]

This blazar is located in direction of Alpha Pegasi (Markab) at the distance of 7.7 Gly (redshift $z = 0.859001 \pm 0.000170$); right ascension is $\alpha = 22^h\ 53^m\ 57.7^s$, declination $\delta = +16°\ 08'\ 53.6''$. It has a strong flux variability at all wavelengths from gamma ray to radio. The spectral energy distribution of 3C 454.3 displays the two peaks typical for AGNs, one in the infrared and optical, and the other in the X-ray and gamma ray. The spectral characteristics of these peaks are determined by two radiation mechanisms – synchrotron radiation by relativistic electrons and inverse-Compton scattering of "soft" photons on relativistic electrons.

Observations have also established that a single radiation mechanism operates from the radio to the optical spectrum. This was first confirmed directly when correlations were found between flux variations at different frequencies during the development of a major flare in 2005–2006 (Volvach et al., 2007). Variations in the flux of 3C 454.3 were observed on scales from days to a year, which were repeated in the optical and radio spectra. It was shown that both the duration of the flare (about a year) and individual features of the flare were the same in these two frequency ranges. This is possible only if a single mechanism is generating the radiation in these different ranges. Thus, it was established that both radio and optical emission is produced by the jet. The delay between the flares in the optical and millimeter ranges was about ten months, with about the same delay observed for centimeter wavelengths. The frequency dependence of the delays and the intervals between flares can be used to predict future flares in this object in various frequency ranges.

For example, the three flares in the blazar 3C 454.3 were observed during 2005–2010 and allowed to determine their locations in the jet from gamma ray to radio range and to estimate a size of the Stromgren zone for sources of ionization associated with a binary supermassive black hole in the central region (Volvach et al., 2011).

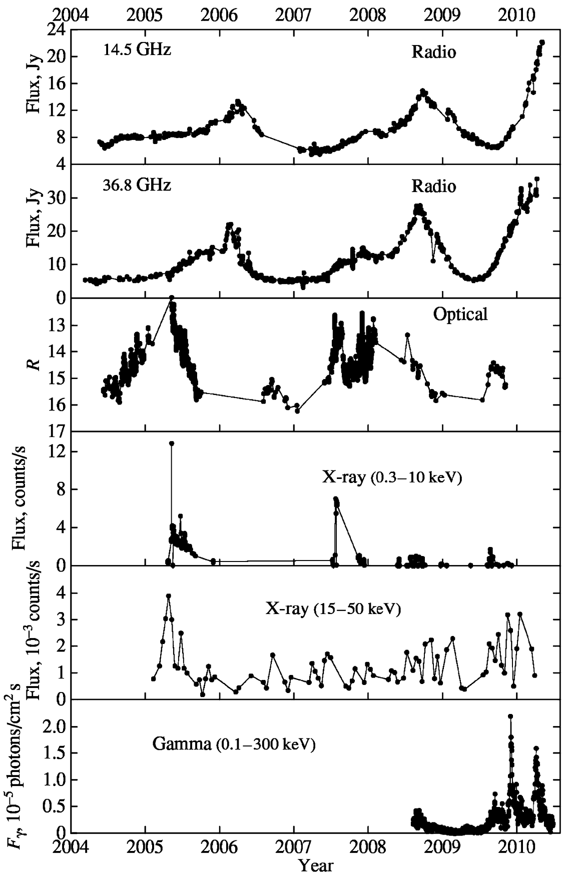

FIG. 16.10 Light curves for the blazar 3C 454.3 at various frequencies from radio to gamma ray ranges, obtained for the observational period of 2004–2010 (Volvach et al., 2011).

Light curves for the blazar 3C 454.3 in spectral ranges from radio to gamma ray obtained for the observational period of 2004–2010 years are presented in Fig. 16.10.

- Radio observations were conducted at 22.2 and 36.8 GHz with the 22 m radio telescope of the Crimean Astrophysical Observatory using modulation radiometers (Efanov et al., 1979). Observations at 4.8, 8, and 14.5 GHz were carried out on the 26 m telescope of the University of Michigan Radio Astronomy Observatory (Aller et al., 2001). Observations at 37 GHz were obtained using the 14 m telescope of the Metsahovi Radio Observatory of Aalto University. The data for radio range obtained with different telescopes are in good agreement and sup-

[4] Blazar is an active galaxy nucleus, which is located at the center of the elliptical galaxy and has a relativistic jet oriented close to the line of sight with the observer. The name *blazar* was coined in 1978 by E. Spiegel to denote the combination of BL Lac objects and of optically violently variable quasars. Being one of the powerful sources of emission, blazars are characterized by high polarization and very rapid fluctuations in brightness. Among well-known blazars are BL Lacertae, 3C 454.3, 3C 273, PKS 2155-304, S5 0014+81 (TeV Blazar with the most supermassive black hole, $10^9\ \odot$), and others. The unique blazar TXS 0506+056, which is a source of high-energy neutrinos, was discovered in the frame of the IceCube project (icecube.wisc.edu) in July, 2018 (see, for example, Overbye Dennis, "It Came From a Black Hole, and Landed in Antarctica – For the first time, astronomers followed cosmic neutrinos into the fire-spitting heart of a supermassive blazar," 12 July 2018, The New York Times).

plement each other during the long-term monitoring period.

- Optical data were obtained from the WEBT archive at the Osservatorio Astronomico di Torino of the Istituto Nazionale Di Astrofisica as a part of the WEBT program (Whole Earth Blazar Telescope) and were supplemented with observational data from the 70 cm telescope of the Crimean Astrophysical Observatory (Sergeev et al., 2005).
- The gamma ray observational data of 3C 454.3 from 24 April 2005 through 18 September 2010 were provided from the Swift spacecraft Burst Alert Telescope (BAT), which operates at 15–195 keV, and from the X-ray Telescope (XRT), which operates at 0.3–10 keV. These data are available through the HEASARC web site.[5] The energy range for the 3C 454.3 light curve obtained during the Swift/BAT transient monitoring program is 15–50 keV (Fig. 16.10). X-ray observations at 2–10 keV range were obtained with the Rossi X-ray Timing Explorer (RXTE). To exclude short-period variability and reduce uncertainty in the measured fluxes these light curves could be averaged over time intervals of one day, which corresponds to 15 orbital periods of the RXTE satellite (Chesnok et al., 2009). The data from the Fermi gamma ray telescope (the main instrument is the Large Area Telescope) were used to calculate light curves from 100 MeV to 300 GeV (lower panel of Fig. 16.10) of 3C 454.3, which is identified with the gamma ray source 1FGL 2253.9+1608 (Abdo et al., 2010).

Using the Shuster method one can conduct a harmonic analysis of the flux variations of 3C 454.3 from radio to gamma ray ranges and derive a unified law for the frequency-dependent delays of the flares. The double character of the flares in the period 2005–2010 may indicate the passage of a companion of the central supermassive black hole through the accretion disk at the pericenter, with the disk oriented at some angle to the orbit of the companion (Vol'Vach et al., 2011).

The American Association of Variable Star Observers installed a "Light Curve Generator for 3C 454.3," which is available through http://www.aavso.org/, where everybody can find periods of its outbursts, brightening to a peak apparent magnitude of 13.4 in June 2014. Using SIMBAD entry for "3C 454.3" one can access to all the available multiwavelength data on this blazar.

Instead of conclusion to the chapter. Summarizing what is written above, we emphasize that the current multiwavelength databases help us not only to study deeply the known phenomena in the universe, but also to discover new features that have not been seen before. Current data in various archives save a lot of time needed for downloading, analyzing, collecting, and describing soft observational data. Astronomical data are very heterogeneous, allowing everyone to find in this variety a solution to the new puzzles of the Universe.

[5] http://heasarc.gsfc.nasa.gov/docs/swift/results/transients/weak/.

REFERENCES

Abdo, A.A., et al., 2010. Fermi Large Area Telescope first source catalog. The Astrophysical Journal. Supplement Series 188 (2), 405–436. https://doi.org/10.1088/0067-0049/188/2/405. arXiv:1002.2280 [astro-ph.HE].

Aller, M.F., Aller, H.D., Hughes, P.A., 2001. The longterm centimeter-band total flux and linear polarization properties of the Pearson-Readhead Survey Sources. In: American Astronomical Society Meeting Abstracts, vol. 199. American Astronomical Society Meeting Abstracts, p. 138.15.

Andrae, R., Melchior, P., Bartelmann, M., 2010. Soft clustering analysis of galaxy morphologies: a worked example with SDSS. Astronomy & Astrophysics 522, A21. https://doi.org/10.1051/0004-6361/201014169. arXiv:1002.0676 [astro-ph.CO], 2010.

Babyk, Iu.V., Vavilova, I.B., 2012. The distribution of baryon matter in the nearby X-ray galaxy clusters. Odessa Astronomical Publications 25, 119.

Babyk, Iu.V., Vavilova, I.B., 2013. Comparison of optical and X-ray mass estimates of the Chandra galaxy clusters at $z < 0.1$. Odessa Astronomical Publications 26, 175.

Babyk, I., Vavilova, I., 2014. The Chandra X-ray galaxy clusters at $z < 1.4$: constraints on the evolution of L_X–T–M_g relations. Astrophysics and Space Science 349 (1), 415–421. https://doi.org/10.1007/s10509-013-1630-z.

Ball, N.M., Brunner, R.J., 2010. Data mining and machine learning in astronomy. International Journal of Modern Physics D 19 (7), 1049–1106. https://doi.org/10.1142/S0218271810017160. arXiv:0906.2173 [astro-ph.IM].

Ball, N.M., Loveday, J., et al., 2004. Galaxy types in the Sloan Digital Sky Survey using supervised artificial neural networks. Monthly Notices of the Royal Astronomical Society 348 (3), 1038–1046. https://doi.org/10.1111/j.1365-2966.2004.07429.x. arXiv:astro-ph/0306390 [astro-ph].

Banerji, Manda, et al., 2010. Galaxy Zoo: reproducing galaxy morphologies via machine learning. Monthly Notices of the Royal Astronomical Society 406 (1), 342–353. https://doi.org/10.1111/j.1365-2966.2010.16713.x. arXiv:0908.2033 [astro-ph.CO].

Banfield, J.K., et al., 2015. Radio Galaxy Zoo: host galaxies and radio morphologies derived from visual inspection. Monthly Notices of the Royal Astronomical Society 453 (3), 2326–2340. https://doi.org/10.1093/mnras/stv1688. arXiv:1507.07272 [astro-ph.GA].

Bell, Keaton, Salim, S., 2011. Automatic morphological classification of galaxies in SDSS/GALEX based on catalog photometry. In: American Astronomical Society Meeting Abstracts, vol. 217. American Astronomical Society Meeting Abstracts, p. 335.33.

Branchini, E., et al., 1999. A non-parametric model for the cosmic velocity field. Monthly Notices of the Royal Astronomical Society 308 (1), 1–28. https://doi.org/10.1046/j.1365-8711.1999.02514.x. arXiv:astro-ph/9901366 [astro-ph].

Brunner, R.J., Djorgovski, S.G., Prince, T.A., Szalay, A.S., 2002. Massive datasets in astronomy. In: Abello, James, Pardalos , Panos M., Resende, Mauricio G.C. (Eds.), Handbook of Massive Data Sets. In: Massive Computing, vol. 4. Springer, Boston, MA. ISBN 978-1-4613-4882-5.

Burton, W.B., et al., 1996. Neutral hydrogen in the nearby galaxies Dwingeloo 1 and Dwingeloo 2. Astronomy & Astrophysics 309, 687–701. arXiv:astro-ph/9511020 [astro-ph].

Buta, Ronald J., McCall, Marshall L., 1999. The IC 342/Maffei Group revealed. The Astrophysical Journal. Supplement Series 124 (1), 33–93. https://doi.org/10.1086/313255.

Buta, Ronald J., Sheth, Kartik, et al., 2010. Mid-infrared galaxy morphology from the Spitzer Survey of Stellar Structure in Galaxies (S^4G): the imprint of the De Vaucouleurs revised Hubble–Sandage classification system at 3.6 μm. The Astrophysical Journal. Supplement Series 190 (1), 147–165. https://doi.org/10.1088/0067-0049/190/1/147. arXiv:1008.0805 [astro-ph.CO].

Charlier, C., 1922. How an infinite world may be built up. Arkiv för Matematik, Astronomi och Fysik 16 (22), 1–34.

Chesnok, N.G., Sergeev, S.G., Vavilova, I.B., 2009. Optical and X-ray variability of Seyfert galaxies NGC 5548, NGC 7469, NGC 3227, NGC 4051, NGC 4151, Mrk 509, Mrk 79, and Akn 564 and quasar 1E 0754. Kinematics and Physics of Celestial Bodies 25 (2), 107–113. https://doi.org/10.3103/S0884591309020068.

Conselice, Christopher J., 2014. The evolution of galaxy structure over cosmic time. Annual Review of Astronomy and Astrophysics 52, 291–337. https://doi.org/10.1146/annurev-astro-081913-040037. arXiv:1403.2783 [astro-ph.GA].

Courtois, Hélène M., et al., 2012. Three-dimensional velocity and density reconstructions of the local universe with Cosmicflows-1. The Astrophysical Journal 744 (1), 43. https://doi.org/10.1088/0004-637X/744/1/43. arXiv:1109.3856 [astro-ph.CO].

Cucciati, O., Granett, B.R., et al., 2014. The VIMOS Public Extragalactic Redshift Survey (VIPERS). Never mind the gaps: comparing techniques to restore homogeneous sky coverage. Astronomy & Astrophysics 565, A67. https://doi.org/10.1051/0004-6361/201423409. arXiv:1401.3745 [astro-ph.CO].

Cucciati, O., Iovino, A., et al., 2006. The VIMOS VLT Deep Survey: the buildup of the colour-density relation. Astronomy & Astrophysics 458 (1), 39–52. https://doi.org/10.1051/0004-6361:20065161. arXiv:astro-ph/0603202 [astro-ph].

de la Calleja, J., Fuentes, O., 2004. Machine learning and image analysis for morphological galaxy classification. Monthly Notices of the Royal Astronomical Society 349 (1), 87–93. https://doi.org/10.1111/j.1365-2966.2004.07442.x.

Dickey, John M., Lockman, Felix J., 1990. H I in the galaxy. Annual Review of Astronomy and Astrophysics 28, 215–261. https://doi.org/10.1146/annurev.aa.28.090190.001243.

Dobrycheva, D.V., Melnyk, O.V., et al., 2014. Environmental properties of galaxies at $z < 0.1$ from the SDSS via the Voronoi tessellation. Odessa Astronomical Publications 27, 26.

Dobrycheva, D.V., Vavilova, I.B., et al., 2017. Machine learning technique for morphological classification of galaxies at $z < 0.1$ from the SDSS. arXiv:1712.08955 [astro-ph.GA].

Easton, C., 1904. La distribution des nebeleuses et leurs relation avec le systeme galactique. Astronomische Nachrichten 166, 131.

Ebeling, H., et al., 2001. Discovery of a very X-ray luminous galaxy cluster at $z = 0.89$ in the Wide Angle ROSAT Pointed survey. The Astrophysical Journal Letters 548 (1), L23–L27. https://doi.org/10.1086/318915. arXiv:astro-ph/0012175.

Ebeling, Harald, Mullis, Christopher R., Brent Tully, R., 2002. A systematic X-ray search for clusters of galaxies behind the Milky Way. The Astrophysical Journal 580 (1), 774–788. https://doi.org/10.1086/343790.

Efanov, V.A., Moiseev, I.G., Nesterov, N.S., 1979. A 1.35 CM wavelength survey of extragalactic radio sources. Izvestiya Ordena Trudovogo Krasnogo Znameni Krymskoj Astrofizicheskoj Observatorii 60, 3–13.

Elyiv, Andrii, 2006. UHECRs deflections in the IRAS PSCz catalogue based models of extragalactic magnetic field. ArXiv e-prints, arXiv:astro-ph/0611696 [astro-ph].

Elyiv, A., Melnyk, O., Vavilova, I., 2009. High-order 3D Voronoi tessellation for identifying isolated galaxies, pairs and triplets. Monthly Notices of the Royal Astronomical Society 394 (1), 1409–1418. https://doi.org/10.1111/j.1365-2966.2008.14150.x. arXiv:0810.5100 [astro-ph].

Erdoğdu, Pirin, Lahav, Ofer, 2009. Is the misalignment of the Local Group velocity and the dipole generated by the 2MASS Redshift Survey typical in Λ cold dark matter and the halo model of galaxies? Physical Review D 80 (4), 043005. https://doi.org/10.1103/PhysRevD.80.043005. arXiv:0906.3111 [astro-ph.CO].

Feigelson, E.D., Babu, G.J., 2006. Statistical Challenges in Astronomy. Springer. ISBN 978-0-387-21529-7.

Giovanelli, Riccardo, Haynes, Martha P., 1989. A 21-cm survey of the Pisces-Perseus supercluster. IV. Addenda to the declination zone 21.5 degrees to 33.5 degrees. The Astronomical Journal 97, 633. https://doi.org/10.1086/115010.

Goodfellow, Ian J., et al., 2014. Generative adversarial networks. ArXiv e-prints, arXiv:1406.2661 [stat.ML].

Gregul, A.Ia., Mandzhos, A.V., Vavilova, I.B., 1991. The existence of the structural anisotropy of the Jagiellonian field of the galaxies. Astrophysics and Space Science 185 (2), 223–235. https://doi.org/10.1007/BF00643190.

Huchtmeier, W.K., et al., 1995. Two new possible members of the IC342-Maffei1/2 group of galaxies. Astronomy & Astrophysics 293, L33–L36.

Ivezić, Željko, Connolly, Andrew J., VanderPlas, Jacob T., Gray, Alex, 2014. Statistics, Data Mining, and Machine Learning in Astronomy: A Practical Python Guide for the Analysis of Survey Data. Prinseton University Press. ISBN 9780691151687.

Jain, Archa, Gauthier, Alexandre, Noordeh, Emil, 2016. Galaxy morphology classification. In: CS 229 Machine Learning Final Projects. Stanford University, 6 p.

Jarrett, T.H., et al., 2000. 2MASS extended source catalog: overview and algorithms. The Astronomical Journal 119 (5), 2498–2531. https://doi.org/10.1086/301330. arXiv:astro-ph/0004318.

Karachentsev, I.D., 2005. The Local Group and other neighboring galaxy groups. The Astronomical Journal 129 (1), 178–188. https://doi.org/10.1086/426368. arXiv:astro-ph/0410065.

Karachentsev, I.D., et al., 2003. Distances to nearby galaxies around IC 342. Astronomy & Astrophysics 408, 111–118. https://doi.org/10.1051/0004-6361:20030912.

Karachentsev, Igor D., Makarov, Dmitry I., Kaisina, Elena I., 2013. Updated nearby galaxy catalog. The Astronomical Journal 145 (4), 101. https://doi.org/10.1088/0004-6256/145/4/101. arXiv:1303.5328 [astro-ph.CO].

Kashibadze, O.G., Karachentsev, I.D., Karachentseva, V.E., 2018. Surveying the local supercluster plane. Astrophysical Bulletin 73 (2), 124–141. https://doi.org/10.1134/S1990341318020025.

Kasivajhula, Siddhartha, Raghavan, Naren, Shah, Hemal, 2007. Morphological galaxy classification using machine learning. In: CS 229 Machine Learning Final Projects. Stanford University, 5 p.

Kates-Harbeck, Julian, 2012. Galaxy image processing and morphological classification using machine learning. In: APS April Meeting Abstracts, vol. 2012, E1.075.

Keel, W.C., 2007. The Road to Galaxy Formation. Springer-Verlag Berlin Heidelberg. ISBN 978-3-540-72534-3.

Kocevski, D.D., Ebeling, H., Mullis, C.R., 2004. Clusters in the Zone of Avoidance. In: Mulchaey, J.S., Dressler, A., Oemler, A. (Eds.), Clusters of Galaxies: Probes of Cosmological Structure and Galaxy Evolution, p. 26. arXiv:astro-ph/0304453 [astro-ph].

Kogut, A., et al., 1993. Dipole anisotropy in the COBE differential microwave radiometers first-year sky maps. The Astrophysical Journal 419, 1. https://doi.org/10.1086/173453. arXiv:astro-ph/9312056.

Kolatt, Tsafrir, Dekel, Avishai, 1997. Large-scale power spectrum from peculiar velocities. The Astrophysical Journal 479 (2), 592–605. https://doi.org/10.1086/303894.

Kraan-Korteweg, Renée C., 2005. Cosmological structures behind the Milky Way. Reviews in Modern Astronomy 18, 48–75. https://doi.org/10.1002/3527608966.ch3. arXiv:astro-ph/0502217 [astro-ph].

Kraan-Korteweg, R.C., Loan, A.J., et al., 1994. Discovery of a nearby spiral galaxy behind the Milky Way. Nature 372 (6501), 77–79. https://doi.org/10.1038/372077a0.

Kraan-Korteweg, R.C., Staveley-Smith, L., et al., 2003. The Universe behind the Southern Milky Way. ArXiv e-prints, arXiv:astro-ph/0311129 [astro-ph].

Kraan-Korteweg, Renée C., Lahav, Ofer, 2000. The Universe behind the Milky Way. The Astronomy and Astrophysics Review 10 (3), 211–261. https://doi.org/10.1007/s001590000011. arXiv:astro-ph/0005501 [astro-ph].

Kuminski, Evan, Shamir, Lior, 2016. A computer-generated visual morphology catalog of ~3,000,000 SDSS galaxies. The Astrophysical Journal. Supplement Series 223 (2), 20. https://doi.org/10.3847/0067-0049/223/2/20. arXiv:1602.06854 [astro-ph.GA].

Lahav, Ofer, et al., 1998. Galaxy candidates in the Zone of Avoidance. Monthly Notices of the Royal Astronomical Society 299 (1), 24–30. https://doi.org/10.1046/j.1365-8711.1998.01686.x. arXiv:astro-ph/9707345 [astro-ph].

Lahav, O., Fisher, K.B., et al., 1994. Wiener reconstruction of all-sky galaxy surveys in spherical harmonics. The Astrophysical Journal Letters 423, L93. https://doi.org/10.1086/187244. arXiv:astro-ph/9311059 [astro-ph].

Lahav, O., Naim, A., et al., 1996. Neural computation as a tool for galaxy classification: methods and examples. Monthly Notices of the Royal Astronomical Society 283, 207. https://doi.org/10.1093/mnras/283.1.207. arXiv:astro-ph/9508012.

Lu, N.Y., et al., 1990. Identifying galaxies in the Zone of Avoidance. The Astrophysical Journal 357, 388. https://doi.org/10.1086/168929.

Maffei, Paolo, 2003. My researches at the infrared doors. Memorie Della Società Astronomica Italiana 74, 19.

Mao, Q., et al., 2016. A cosmic void catalog of SDSS DR12 BOSS galaxies. The Astrophysical Journal 835. https://doi.org/10.3847/1538-4357/835/2/161.

Melnyk, O.V., Dobrycheva, D.V., Vavilova, I.B., 2012. Morphology and color indices of galaxies in pairs: criteria for the classification of galaxies. Astrophysics 55 (3), 293–305. https://doi.org/10.1007/s10511-012-9236-7.

Melnyk, O.V., Elyiv, A.A., Vavilova, I.B., 2006. The structure of the Local Supercluster of galaxies detected by three-dimensional Voronoi's tessellation method. Kinematika i Fizika Nebesnykh Tel 22 (4), 283–296. arXiv:0712.1297 [astro-ph].

Murrugarra, J.H., Hirata, N.S.T., 2017. Galaxy image classification, pp. 01–04. URL: http://sibgrapi2017.ic.uff.br/e-proceedings/assets/papers/WIP/WIP8.pdf.

Naim, A., et al., 1995. A comparative study of morphological classifications of APM galaxies. Monthly Notices of the Royal Astronomical Society 274 (4), 1107–1125. https://doi.org/10.1093/mnras/274.4.1107. arXiv:astro-ph/9502078.

Nair, Preethi B., Abraham, Roberto G., 2010. A catalog of detailed visual morphological classifications for 14,034 galaxies in the Sloan Digital Sky Survey. The Astrophysical Journal. Supplement Series 186 (2), 427–456. https://doi.org/10.1088/0067-0049/186/2/427. arXiv:1001.2401 [astro-ph.CO].

Proctor, R., 1878. The Universe of Stars. Longmans, Green and Co., London, 41.

Said, K., Kraan-Korteweg, R.C., Jarrett, T.H., 2014. Galaxy peculiar velocities in the Zone of Avoidance. ArXiv e-prints, arXiv:1410.2992 [astro-ph.GA].

Said, Khaled, Kraan-Korteweg, Renée C., Jarrett, T.H., et al., 2016a. NIR Tully–Fisher in the Zone of Avoidance - III. Deep NIR catalogue of the HIZOA galaxies.

Monthly Notices of the Royal Astronomical Society 462 (3), 3386–3400. https://doi.org/10.1093/mnras/stw1887. arXiv:1607.08596 [astro-ph.GA].

Said, Khaled, Kraan-Korteweg, Renée C., Staveley-Smith, Lister, et al., 2016b. NIR Tully–Fisher in the Zone of Avoidance - II. 21 cm H I-line spectra of southern ZOA galaxies. Monthly Notices of the Royal Astronomical Society 457 (3), 2366–2376. https://doi.org/10.1093/mnras/stw105. arXiv:1601.07162 [astro-ph.GA].

Sanford, R.F., 1917. On some relation of the spiral nebulae to the Milky Way. Lick Observatory Bulletin 297, Part IX, 80.

Saurer, W., Seeberger, R., Weinberger, R., 1997. Penetrating the "zone of avoidance": IV. An optical survey for hidden galaxies in the region $130° \leq l \leq 180°$, $-5° \leq b \leq +5°$. Astronomy & Astrophysics. Supplement Series 126, 247–250. https://doi.org/10.1051/aas:1997385.

Schawinski, Kevin, et al., 2017. Generative adversarial networks recover features in astrophysical images of galaxies beyond the deconvolution limit. Monthly Notices of the Royal Astronomical Society 467 (1), L110–L114. https://doi.org/10.1093/mnrasl/slx008. arXiv:1702.00403 [astro-ph.IM].

Sergeev, S.G., et al., 2005. Lag-luminosity relationship for interband lags between variations in B, V, R, and I bands in active galactic nuclei. The Astrophysical Journal 622 (1), 129–135. https://doi.org/10.1086/427820.

Shane, C.D., Wirtanen, C.A., 1967. The distribution of galaxies. Publications of the Lick Observatory XXII, Pt. I.

Shapley, H., 1961. Galaxies. Harvard Univ. Press, Cambridge, 159.

Smith, Michael D., Donohoe, Justin, 2019. The morphological classification of distant radio galaxies explored with three-dimensional simulations. Monthly Notices of the Royal Astronomical Society 490 (1), 1363–1382. https://doi.org/10.1093/mnras/stz2525. arXiv:1909.03905 [astro-ph.GA].

Sorce, Jenny G., et al., 2017. Predicting structures in the Zone of Avoidance. Monthly Notices of the Royal Astronomical Society 471 (3), 3087–3097. https://doi.org/10.1093/mnras/stx1800. arXiv:1707.04267 [astro-ph.CO].

Stratonoff, W., 1900. Etudes sur la structure de l'univers. Publications de l'Observatoire astronomique et physique de Tachkent. 2, 136P.

Storrie-Lombardi, M.C., et al., 1992. Morphological classification of galaxies by artificial neural networks. Monthly Notices of the Royal Astronomical Society 259, 8P. https://doi.org/10.1093/mnras/259.1.8P.

Vasylenko, M.Yu., et al., 2019. Verification of machine learning methods for binary morphological classification of galaxies from SDSS. Odessa Astronomical Publications 32, 46–51. https://doi.org/10.18524/1810-4215.2019.32.182538.

Vasylenko, M.Yu., Kudrya, Yu.N., 2017. Dipole bulk velocity based on new data sample of galaxies from the catalogue 2MFGC. Advances in Astronomy and Space Physics 7, 6–11. https://doi.org/10.17721/2227-1481.7.6-11.

Vavilova, I.B., 1997. Cluster and wavelet analysis for detachment of the structure of galaxy cluster: comparison. Data Analysis in Astronomy, 297–302.

Vavilova, I.B., 2000. Wavelet analysis as approach to recognize abundance zone in galaxy distribution. Kinematika i Fizika Nebesnykh Tel Supplement 3, 155.

Vavilova, I.B., Elyiv, A.A., Vasylenko, M.Yu., 2018. Behind the Zone of Avoidance of the Milky Way: what can we restore by direct and indirect methods? Russian Radio Physics and Radio Astronomy 23 (4), 244–257. https://doi.org/10.15407/rpra23.04.244.

Vavilova, Iryna, Melnyk, Olga, 2005. Voronoi tessellation for galaxy distribution. In: Mathematics and Its Applications. Proceedings of the Institute of Mathematics of the NAS of Ukraine, vol. 55, pp. 203–212.

Volvach, A.E., Larionov, M.G., et al., 2011. Flare activity of the blazar 3C 454.3 from gamma to radio wavelengths in 2004–2010. Kosmichna Nauka i Tekhnologiya 17 (2), 68–76. https://doi.org/10.15407/knit2011.02.068.

Volvach, A.E., Volvach, L.N., et al., 2007. The variability of a 3C 454.3 blazar over a 40-year period. Astronomy Reports 51 (6), 450–459. https://doi.org/10.1134/S1063772907060030.

Vol'Vach, A.E., et al., 2011. Multi-frequency studies of the non-stationary radiation of the blazar 3C 454.3. Astronomy Reports 55 (7), 608–615. https://doi.org/10.1134/S1063772911070092.

Wahaono, M.A.R., Azhari, S.N., 2018. Classification of galaxy morphological image based on convolutional neural network. International Journal of Advanced Research in Science, Engineering and Technology 5 (6), 6066–6073.

Willett, Kyle W., et al., 2013. Galaxy Zoo 2: detailed morphological classifications for 304 122 galaxies from the Sloan Digital Sky Survey. Monthly Notices of the Royal Astronomical Society 435 (1), 2835–2860. https://doi.org/10.1093/mnras/stt1458. arXiv:1308.3496 [astro-ph.CO].

York, Donald G., et al., 2000. The Sloan Digital Sky Survey: technical summary. The Astronomical Journal 120 (3), 1579–1587. https://doi.org/10.1086/301513. arXiv:astro-ph/0006396.