Universidade de São Paulo
Instituto de Astronomia, Geofísica e Ciências Atmosféricas
Departamento de Astronomia

Kayleigh Meneghini

Construção de um Catálogo Ambiental para o S-PLUS

São Paulo
2020

Kayleigh Meneghini

Construção de um Catálogo Ambiental para o S-PLUS

Trabalho de Conclusão de Curso apresentado ao Departamento de Astronomia do Instituto de Astronomia, Geofísica e Ciências Atmosféricas da Universidade de São Paulo como requisito parcial para a obtenção do título de Bacharel em Astronomia.

Vertente: Pesquisa Básica
Orientador: Prof. Dr. Laerte Sodré Jr.

São Paulo
2020

à minha família e aos que ficaram da ponte pra lá.

Agradecimentos

Agradeço à minha família que, mesmo tendo muito pouco, me deu o que há de mais valioso. À minha vó, Dona Rosa, razão de eu ter chegado até aqui, por ter me dado o que ninguém poderá tirar de mim. À minha mãe, Rosemary, a melhor pessoa do mundo, meu maior exemplo de humildade, bondade e amor. Ao meu pai, Marino, por me inspirar, por ser de ferro e me ensinar a ser também.

Ao meu orientador, Prof. Laerte, por toda paciência, compreensão e apoio; por me ouvir, entender e aceitar o melhor que pude dar. Serei sempre grata por o ter encontrado.

À minha assistente social, Rosangela Dearo, que me acolheu enquanto eu nadava contra a correnteza, me colocando em um barco e não deixando eu me afogar.

Ao Raul, por muitas vezes me emprestar os seus olhos para que eu me enxergasse capaz; por todo companheirismo, amor e amizade. E pela GATA.

Aos meus amigos do IAG e monitores do OAM, resistência ou cachimbinho, impostores e tripulantes, por todas as noites com as estrelas. Obrigada por fazerem essa graduação valer a pena, vamos comemorar com Xandô.

Ao professor Rama, por ter confiado em mim e me dado a oportunidade de participar desse projeto que levarei para sempre em meu coração.

Aos meus amigos cruspianos, guerreiros de fé, pela vivência, pelos momentos de luta e por todas as risadas.

Este trabalho foi escrito em LaTeX com a classe IAGTESE, para teses e dissertações do IAG.

"*Tudo, tudo, tudo vai, tudo é fase irmão*
Logo mais vamo arrebentar no mundão"

Racionais Mc's

Resumo

Esse trabalho teve como objetivo quantificar o ambiente galáctico em diferentes escalas. Para isso, foi desenvolvido um código para a construção de um catálogo de estimadores ambientais para uma amostra de objetos classificados como galáxias do DR-2 do S-PLUS. Selecionado objetos com magnitudes Petrosianas inferiores a 21 e *redshifts* (z) entre 0.001 e 0.1, foi obtida uma amostra final de 17840 objetos, cuja vizinhança foi definida a partir de um raio de dois desvios padrões em torno do z de cada galáxia. Definida a vizinhança, calculou-se as distâncias físicas cosmológicas a cada vizinho, utilizadas para a construção do catálogo ambiental propostos através de dois métodos distintos: *Nearest Neighbors* (NN) e Aberturas Fixas. Os estimadores calculados pelos método dos NN, utilizando k=1, 7, 10 e 50 vizinhos, apresentam alta variabilidade e distribuições assimétricas, cuja escala de distância varia de milhares de anos-luz a dezenas de Mpc. Em contrapartida, os campos de densidade calculados pelas aberturas fixas de raios variando de 0.25 a 10 Mpc apresentam maior robustez, com distribuição mais simétricas e com valores extremos menos frequentes. No entanto, por sua natureza discreta, o método pode apresentar uma visão míope de ambientes de baixa densidade ou de interação entre pares de galáxias. Observou-se forte correlação de Spearman entre os estimadores calculados para k=50 vizinhos e raio fixo de 2.0 Mpc, escala típica de grupo de galáxias.

Abstract

This work aimed to quantify the galactic environment at different scales. For this, a code was developed to build a catalog of environmental estimators for a sample of objects classified as galaxies of the DR-2 from S-PLUS. Selecting objects with Petrosian magnitudes less than 21 and redshifts (z) between 0.001 and 0.1, a final sample of 17840 objects was obtained, whose neighborhood was defined from a radius of two standard deviations around the z of each galaxy. Once the neighborhood was defined, the cosmological physical distances to each neighbor were calculated and used for the construction of the environmental catalog proposed through two distinct methods: Nearest Neighbors (NN) and Fixed Apertures. The estimators calculated by the NN method show high variability and asymmetric distributions, whose distance scales vary from thousands of light years to tens of Mpc. In contrast, the density fields calculated by the fixed apertures ranging from 0.25 to 10 Mpc are more robust, with more symmetrical distribution and with less frequent extreme values. However, due to its discreet nature, the method can present a myopic view of low density environments or of interaction between pairs of galaxies. A strong Spearman's rank correlation was observed between the estimators calculated for k = 50 neighbors and a fixed radius of 2.0 Mpc, typical scale of a group of galaxies.

Lista de Figuras

2.1 Eficiência total dos filtros do S-PLUS, obtida através da multiplicação das curvas médias de transmissão do filtro, a transmissão atmosférica, a eficiência do CCD e as curvas de refletividade do espelho primário em função do comprimento de onda, as cores representam os 12 diferentes filtros cujos rótulos estão mostrados na legenda à direita da imagem. Fonte: Mendes de Oliveira et al. (2019). 21

2.2 Comparação das distribuições dos *redshifts* fotométricos do DR2 e da amostra selecionada. No eixo x o intervalo de *redshift* e no eixo y sua frequência absoluta. 23

2.3 Comparação das distribuições dos *redshifts* fotométricos e espectroscópicos da amostra selecionada. No eixo x o intervalo de *redshift* e no eixo y sua frequência absoluta 23

2.4 Distribuição das magnitudes Petrosianas. À esquerda, a amostra selecionada para o trabalho com os filtros indicados, à direita a amostra do DR2. . . . 24

4.1 Distribuição dos estimadores até o 90º percentil pelo método dos NN para k=1, 7, 10 e 50. No eixo x constam-se os intervalos de valores para os estimadores e no eixo y sua frequência relativa. 32

4.2 Distribuição dos estimadores pelo método das aberturas fixas, para os raios r = 0.25, 1.0, 2.0, 5.0, 9.0 e 10 Mpc. No eixo x constam-se os intervalos de valores para os estimadores e no eixo y sua frequência relativa 34

4.3 Densidade de *kernel* dos estimadores obtidos para r = 2.0, r = 5.0 e r= 10.0 Mpc, da esquerda para a direita, em função de k = 50. Acima de cada eixo, a distribuição de probabilidades dos valores dos estimadores. 35

Lista de Tabelas

2.1 Descrição dos filtros do S-PLUS, com o nome, comprimento de onda de transmissão efetiva do filtro λ_{eff} e sua largura $\Delta\lambda$. 22

2.2 Comparação entre os *redshifts* fotométricos e espectroscópicos da amostra selecionada. 23

4.1 Estatísticas dos estimadores ambientais calculados pelo método dos NN. . . 32

4.2 Distâncias máxima e mínimo ao k-ésimo vizinho em Mpc. 32

4.3 Estatísticas dos estimadores ambientais calculados pelo método das aberturas fixas. 33

4.4 Valores únicos discretos dos estimadores, número máximo e mínimo de objetos encontrados em cada raio. 34

4.5 Correlação entre os estimadores. 36

Sumário

1. Introdução 19

2. Base de Dados 21

2.1 O Survey 21

2.2 Seleção dos Dados 22

3. Metodologia 25

3.1 *Nearest Neighbors* 25

3.2 Aberturas Fixas 26

3.3 Distâncias em Cosmologia 27

3.3.1 Parâmetros Cosmológicos 27

3.3.2 Distância Comóvel e Distância de Diâmetro Angular 28

4. Resultados e Análise 31

4.1 Estimadores ambientais 31

4.2 Comparação entre os métodos 34

5. Conclusões 37

Referências 39

Capítulo 1

Introdução

"Por que existem diferentes tipos de galáxias?" Eis uma grande questão levantada no século passado, quando os primeiros estudos sobre galáxias se iniciaram.

Características presentes nas populações estelares das galáxias sugeriam que a diferença entre as galáxias elípticas e espirais se tratava de uma questão evolutiva. Em galáxias elípticas observa-se a predominância da população do tipo II, isto é, estrelas mais velhas e de baixa metalicidade, diferente do que se observa nas espirais, onde é observado formação estelar e população do tipo I. Assim, acreditava-se que as galáxias quando jovens seriam espirais e mais tarde evoluiriam a elípticas. No entanto, observa-se que, em ambos tipos de galáxias, as estrelas mais velhas são igualmente velhas, em torno de 11 bilhões de anos. Portanto, essas galáxias se formaram mais ou menos na mesma época e têm a mesma idade. Sendo assim, a pergunta a ser feita é: como essas estruturas se formam e evoluem (de Souza Oliveira Filho e De Fatima Oliveira Saraiva, 2000)?

No Universo primordial, em seu período inflacionário, houve um acúmulo de matéria escura em determinadas regiões que, sob ação da gravidade, colapsou em halos. Segundo o modelo hierárquico, estes evoluíram pela acresção de matéria escura e através de *mergers* com halos satélites. A partir disso, bárions seriam atraídos pelo poço gravitacional desses halos, formando futuramente as galáxias (Etherington e Thomas, 2015).

Na segunda metade do século passado, surgiram as duas teorias principais sobre formação e evolução de galáxias: o modelo monolítico, no qual as galáxias se formaram e evoluíram isoladamente através do colapso de grandes nuvens de gás e o modelo hierárquico, que propõe que as galáxias se formaram e evoluíram através de sucessivos *mergers* entre estruturas menores. Neste último, seriam formados discos primordiais, que formariam galáxias espirais e, no caso de *mergers* sucessivos e frequentes, evoluiriam a galáxias elípticas. Sendo

assim, um fator determinante para evolução das galáxias neste último modelo é o ambiente em que elas se encontram (de Souza Oliveira Filho e De Fatima Oliveira Saraiva, 2000).

Por essa razão, a influência do ambiente nas propriedades das galáxias tem sido muito estudada na literatura. Uma das evidências dessa influência é a relação entre morfologia e densidade: observa-se que galáxias elípticas são encontradas em ambientes mais densos, enquanto as galáxias espirais estão mais frequentemente em ambientes de baixa densidade de galáxias. Podemos também estudar a relação entre o ambiente e propriedades das galáxias como cor, luminosidade, massa, taxa de formação estelar e população estelares. Galáxias encontradas na *"red sequence"* apresentam características tais como serem mais massivas e luminosas, avermelhadas e com baixa taxa de formação estelar. Frequentemente essas galáxias são encontradas em ambientes mais densos. Enquanto as galáxias da *"blue sequence"* apresentam características inversas, como alta taxa de formação estelar, são menos massivas, habitantes de ambientes menos densos (Etherington e Thomas, 2015).

No entanto, o quanto o ambiente afeta as galáxias ainda faz parte do desconhecido, pois alguns resultados se mostram discrepantes a depender da seleção da amostra, segundo Lai et al. (2016) como por exemplo, o limite de magnitude ou dispersão de velocidade escolhido para a amostra. Além disso, a própria definição de ambiente torna esse tipo de estudo não trivial, afinal como definir o ambiente? Podemos considerar o ambiente local, como o associado a uma galáxia central e seus satélites, ou em escalas maiores, como as associadas a grupos e aglomerados. Segundo Muldrew et al. (2012), há dois métodos amplos usados para definir o ambiente galáctico, estes são: *Nearest Neighbors* e Aberturas Fixas.

Esse trabalho tem como objetivo revisitar esses métodos, aplicando-os em uma amostra de dados do *Southern Photometric Local Universe Survey* (S-PLUS), construindo um catálogo de estimadores ambientais para essas galáxias. Esses estimadores poderão ser usados em muitas aplicações, como quantificar a variação das estatísticas de diferentes propriedades galácticas em função do ambiente.

Capítulo 2

Base de Dados

2.1 O Survey

O *S-PLUS* tem como objetivo mapear aproximadamente 9300 graus quadrados do céu do hemisfério sul em doze filtros ópticos compostos por 5 filtros do *Sloan Digital Sky Survey (SDSS)* de banda larga e 7 filtros de banda estreita centrado em várias características proeminentes, ou seja, [OII], Ca H + K, D4000, Hδ, Mgb, Hα e CaT, ilustrados na Figura 2.1. As características detalhadas dos filtros estão descritos na Tabela 2.1.

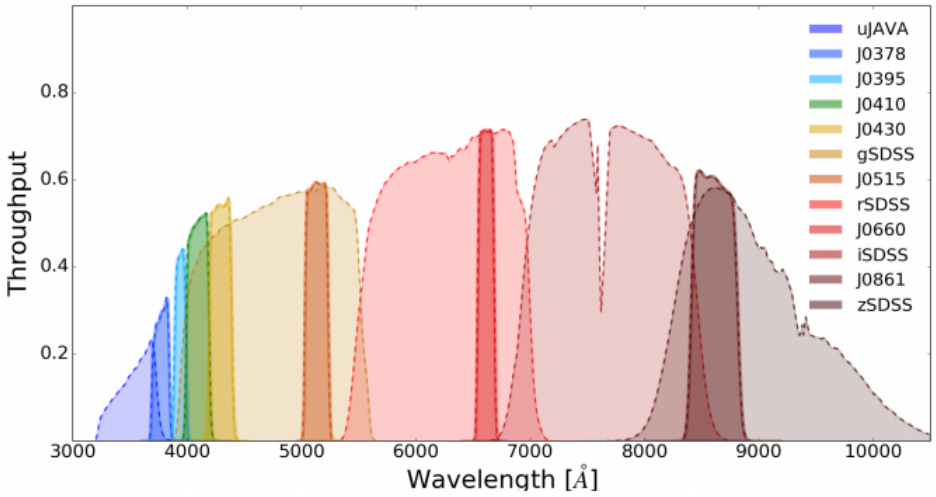

Figura 2.1: Eficiência total dos filtros do S-PLUS, obtida através da multiplicação das curvas médias de transmissão do filtro, a transmissão atmosférica, a eficiência do CCD e as curvas de refletividade do espelho primário em função do comprimento de onda, as cores representam os 12 diferentes filtros cujos rótulos estão mostrados na legenda à direita da imagem. Fonte: Mendes de Oliveira et al. (2019).

Nesse *Survey* são fornecidos *redshifts* fotométricos precisos, com $\delta z/(1+z) = 0.02$ para galáxias com magnitude $r < 19.7$ e $z < 0.4$, produzindo assim um mapa 3D do Universo local sobre um volume de mais de 1 $(Gpc/h)^3$ (Mendes de Oliveira et al., 2019).

Tabela 2.1 - Descrição dos filtros do S-PLUS, com o nome, comprimento de onda de transmissão efetiva do filtro λ_{eff} e sua largura $\Delta\lambda$.

Filtro	λ_{eff} (Å)	$\Delta\lambda$ (Å)	Comentário
uJAVA	3574	330	Javalambre u
J0378	3771	151	[O ii]
J0395	3941	103	Ca H+K
J0410	4094	201	$H\delta$
J0430	4292	200	G-band
gSDSS	4756	1536	SDSS-like g
J0515	5133	207	Mgb Triplet
rSDSS	6260	1462	SDSS-like r
J0660	6614	147	$H\alpha$
iSDSS	7692	1504	SDSS-like i
J0861	8611	408	Ca Triplet
zSDSS	8783	1072	SDSS-like z

Uma etapa extremamente importante de *surveys* fotométricos e fundamental para este trabalho é a classificação de objetos em estrelas e galáxias. No S-PLUS, foi usada a técnica de aprendizado de máquina *Random Forest* utilizando informações fotométricas e morfológicas dos objetos. A performance desse método para objetos com magnitude $r \leq 19$ indica que 97.9% são classificados corretamente.

2.2 Seleção dos Dados

Utilizando o *Data Release 2* (DR2) interno do S-PLUS, filtrando inicialmente objetos classificados no *survey* como galáxias, a seleção dos dados foi feita de modo a se restringir a galáxias de baixos *redshifts* fotométricos, $0.001 \leq z \leq 0.1$, a fim de obter menor dispersão de valores, uma vez que estes *redshifts* tendem a apresentar menor precisão na medida quando comparados aos espectroscópicos (Lai et al., 2016). A resolução do *redshift* é fundamental para os cálculos das distâncias em cosmologia, descritos na Seção 3.3.

Observa-se na Figura 2.2, à esquerda, três picos na distribuição dos *redshifts* fotométricos do DR2, em torno de $z = 0.1$, 0.4 e 0.6, a amostra utilizada neste trabalho está concentrada no primeiro deles, e sua distribuição está ilustrada à direita da imagem.

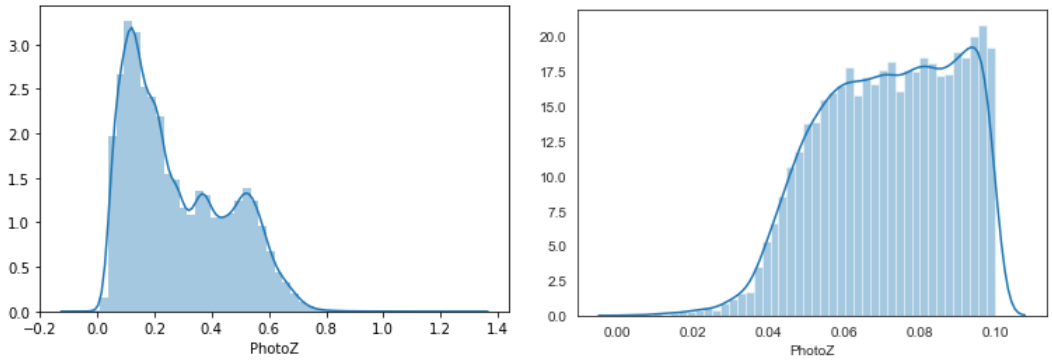

Figura 2.2: Comparação das distribuições dos *redshifts* fotométricos do DR2 e da amostra selecionada. No eixo x o intervalo de *redshift* e no eixo y sua frequência absoluta.

Comparando dentro da amostra os dois diferentes tipos de *redshifts*, observa-se maior consistência, isto é, ausência de valores extremos em relação ao valor médio, e uniformidade nos dados para os *redshifts* fotométricos, trazendo maior confiabilidade para a utilização desses dados. A comparação entre os valores contidos na amostra para cada tipo de *redshift* está descrito na Tabela 2.2 e suas distribuições ilustradas na Figura 2.3.

Tabela 2.2 - Comparação entre os *redshifts* fotométricos e espectroscópicos da amostra selecionada.

	z_photo	z_spec
Valor médio	0.072	0.068
Desvio padrão	0.018	0.043
Valor mínimo	0.0028	-0.010
1° quartil	0.058	0.045
Mediana	0.072	0.064
3° quartil	0.087	0.082
Valor máximo	0.10	1.79

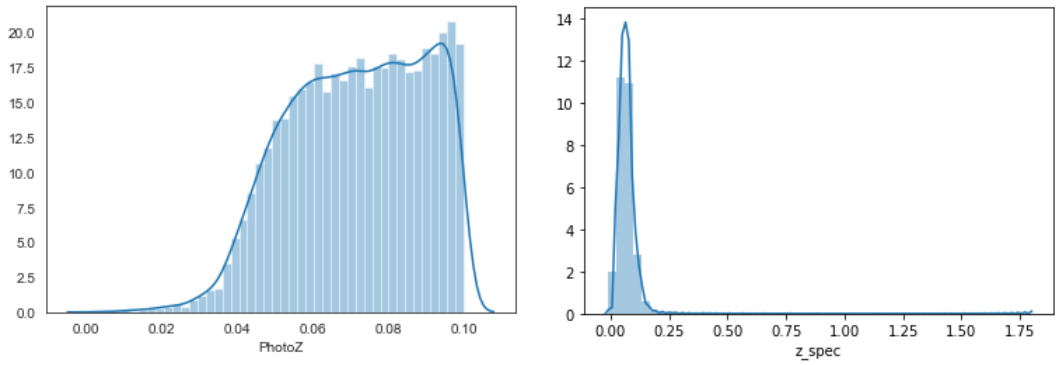

Figura 2.3: Comparação das distribuições dos *redshifts* fotométricos e espectroscópicos da amostra selecionada. No eixo x o intervalo de *redshift* e no eixo y sua frequência absoluta

Selecionamos apenas objetos com magnitude Petrosiana, r_petro, inferior a 21, para evitar que a amostra contenha objetos estelares. Observou-se que 90% dos objetos classificados como galáxias apresentam $r_petro \leq 20.8$. Comparando a distribuição das magnitudes petrosianas na amostra selecionada e na amostra original do DR2, esta última contendo apenas objetos classificados como galáxias, observa-se que a amostra original apresenta objetos pouco luminosos, com valores bem discrepantes ao pico central (Figura 2.4).

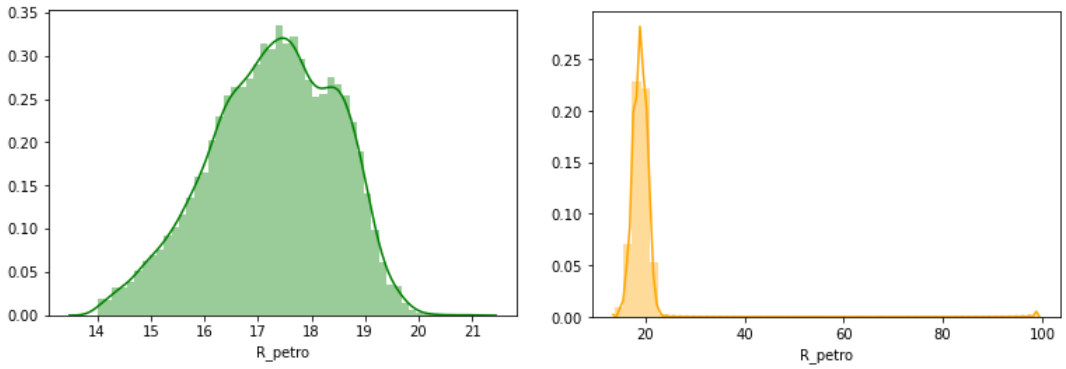

Figura 2.4: Distribuição das magnitudes Petrosianas. À esquerda, a amostra selecionada para o trabalho com os filtros indicados, à direita a amostra do DR2.

Ao final da seleção dos dados, foi obtida uma amostra de 17840 objetos e para cada um calculou-se seus parâmetros ambientais através dos métodos descritos no Capítulo 3.

Capítulo 3

Metodologia

O ambiente galáctico pode ser definido como o campo de densidade onde a galáxia reside (Etherington e Thomas, 2015). Há muitos métodos para medir o ambiente, segundo Muldrew et al. (2012), e a maioria cai em dois amplos grupos: os baseados em encontrar os vizinhos mais próximos (*Nearest Neighbors*) e aqueles que usam aberturas fixas em torno de cada galáxia da amostra.

3.1 Nearest Neighbors

O princípio do *Nearest Neighbors* (NN) é que galáxias com vizinhos mais próximos estão em ambientes mais densos e portanto terão maiores campos de densidade. Esse método define o ambiente galáctico usando uma escala variável, a depender do número de vizinhos de cada galáxia.

Para cada galáxia da amostra escolhe-se valores para k, sendo este o número de vizinhos e calcula-se a distância a cada um deles em torno de um intervalo de z. Assim, nosso estimador ambiental, a densidade superficial de galáxias σ_k, pode ser definido como descrito na Equação 3.1.

$$\sigma_k = \frac{k}{\pi r_k^2}, \tag{3.1}$$

onde k é o número de vizinhos dentro de um raio projetado, r_k, distância ao k-ésimo vizinho. Neste trabalho, usou-se os valores de k iguais a 1, 7, 10 e 50 vizinhos, de modo a tratar de escalas típicas de grupos de galáxias. Além disso, foi utilizado dois desvios padrões de z para selecionar o subgrupo em torno de cada galáxia, tal que os vizinhos de cada objeto estão dentro do intervalo $z_{objeto} - 0.04 < z_{vizinhos} < z_{objeto} + 0.04$.

Para ilustrar o método, tomamos como exemplo o Grupo Local, onde reside a Via Láctea, com cerca de 50 galáxias dentro de um raio aproximado de 1 Mpc (Sparke e Gallagher, 2007). A partir disso, podemos estimar um campo de densidade para a Via Láctea de $\sigma_{50} \approx 16$, utilizando a Equação 3.1, o que corresponderia a um ambiente bastante denso. Entretanto, vale ressaltar que destas 50 galáxias do Grupo Local apenas 3 galáxias espirais (a Via Láctea, M31 e M33) correspondem a 90% da luminosidade total do grupo. Há uma única galáxia elíptica (M32) e o restante são irregulares, anãs esferoidais e elípticas anãs, sendo estas galáxias com baixíssima luminosidade, da ordem de $< 10^7 L_\odot$. Tomando como referência uma amostra contendo apenas as galáxias de maior luminosidade do Grupo Local com $L > 10^8 L_\odot$, o campo de densidade superficial para a Via Láctea seria de $\sigma_7 \approx 3.08$ galáxias por Mpc^2.

Portanto, deve-se levar em consideração em trabalhos semelhantes o número de vizinhos escolhidos, a magnitude limite da amostra e o intervalo de *redshift* do subgrupo dos vizinhos, sendo estes parâmetros escolhidos e ajustados de acordo com o objetivo do estudo e aplicação.

Uma desvantagem desse método é que duas galáxias podem parecer próximas devido a projeção destas no plano de duas coordenadas. Em trabalhos como o de Muldrew et al. (2012) e Lai et al. (2016), foi usado um valor de corte para a velocidade de dispersão das galáxias e o redshift como a "distância real" na linha de visada. Neste trabalho, calculamos a distância entre a galáxia e seu k-ésimo vizinho na superfície esférica e estimamos as distâncias físicas por meio da determinação de distâncias cosmológicas proposta por Hogg (1999), descritas na Seção 3.3.

3.2 Aberturas Fixas

O método de aberturas fixas, ao contrário do anterior, define uma escala para o ambiente, em volta de uma área ou volume fixo para cada galáxia. Assim, quanto mais galáxias dentro dessa área ou volume, mais denso o ambiente.

A densidade superficial de galáxias obtida por esse método é dada por:

$$\sigma_r = \frac{n}{\pi r^2}, \tag{3.2}$$

onde n é o número de galáxias dentro de um raio fixo r ao redor da galáxia alvo.

Neste trabalho, escolheu-se raios projetados de 250 kpc e 1, 2, 5, 8 e 10 Mpc, tratando, portanto, de escalas desde halos individuais até estruturas de larga escala.

A natureza discreta desse método torna difícil distinguir ambientes de baixa densidade, pois podemos encontrar galáxias isoladas dentro de uma abertura fixa centrada nesta (Etherington e Thomas, 2015). Para ilustrar um dos problemas da discretização do método, tomemos como exemplo uma galáxia com um único vizinho em um raio de 1 Mpc, nesse caso a densidade superficial de galáxias seria $\sigma_1 \approx 0.32$. No entanto, esse vizinho pode estar tanto muito próximo à galáxia em questão quanto até 1 Mpc de distância. Portanto, esse método pode passar uma visão míope sobre a interação das galáxias com seus vizinhos.

Assim como no método dos NN, utilizou-se desvios padrões de z para selecionar o subgrupo em torno de cada galáxia, tal que os vizinhos de cada objeto estão dentro do intervalo $z_{objeto} - 0.04 < z_{vizinhos} < z_{objeto} + 0.04$, e para cada um estimou-se a distância às galáxias dentro desse intervalo. A partir disso, calculou-se, para cada objeto, quantas galáxias do subgrupo estão dentro dos raios pré definidos anteriormente.

3.3 Distâncias em Cosmologia

Há muitos métodos de medir e definir distâncias em cosmologia. Sabe-se que com a expansão do Universo, a distância entre objetos está constantemente mudando. Trataremos a seguir das equações para estas medidas propostas por Hogg (1999).

3.3.1 Parâmetros Cosmológicos

A primeira distância tratada que configura a escala do Universo é a distância de Hubble, definida pela Equação 3.3.

$$D_H = \frac{c}{H_0} h^{-1} Mpc, \tag{3.3}$$

onde h é um parâmetro adimensional, H_0 a constante de Hubble e c a velocidade da luz. Neste trabalho adotou-se $h = 1$, $H_0 = 70$ km $s^{-1} Mpc^{-1}$.

Define-se parâmetros de densidade Ω_M, Ω_Λ e Ω_K respectivamente nas Equações 3.4, 3.5 e 3.6

$$\Omega_M = \frac{8\pi G \rho_0}{3H_0^2}, \tag{3.4}$$

$$\Omega_\Lambda = \frac{\Lambda c^2}{3H_0^2}, \tag{3.5}$$

$$\Omega_K = 1 - \Omega_M - \Omega_\Lambda. \tag{3.6}$$

A densidade de massa, ρ, do Universo e o valor da constante cosmológica, Λ, são propriedades dinâmicas do Universo, afetando a evolução temporal da métrica, no entanto em Hogg (1999), são tratados como parâmetros puramente cinemáticos. Os zeros subscritos indicam que essas quantidades devem ser avaliadas na época atual e portanto apresentam uma dependência temporal implícita.

3.3.2 Distância Comóvel e Distância de Diâmetro Angular

A distância comóvel infinitesimal, δD_C, entre dois objetos próximos corresponde a distância entre eles que permaneceria constante no tempo, se ambos objetos se movessem seguindo a Lei de Hubble. Para obter a distância comóvel total na linha de visada, D_C, entre dois objetos em z_1 e z_2, integram-se as contribuições δD_C infinitesimais usando a Equação 3.7

$$D_C = D_H \int_{z_1}^{z_2} \frac{dz'}{\sqrt{\Omega_M(1+z')^3 + \Omega_k(1+z')^2 + \Omega_\Lambda)}}, \tag{3.7}$$

onde Ω_M, Ω_Λ e Ω_K são os parâmetros cosmológicos de densidade definidos anteriormente e D_H a distância de Hubble.

A distância comóvel transversal entre dois objetos a um mesmo *redshift*, mas separados no céu por algum ângulo $\delta\theta$ é dada por $D_M \delta\theta$ e a distância comóvel transversal, D_M, está relacionada com D_C e é obtida a depender do parâmetro de curvatura como descrito na Equação 3.8

$$D_M = \begin{cases} D_H \frac{1}{\sqrt{\Omega_K}} sinh[\sqrt{\Omega_K} D_C/D_H], & \text{Se } \Omega_K > 0, \\ D_C, & \text{Se } \Omega_K = 0, \\ D_H \frac{1}{\sqrt{|\Omega_K|}} sin[\sqrt{|\Omega_K|} D_C/D_H], & \text{Se } \Omega_K < 0. \end{cases} \tag{3.8}$$

Neste trabalho, tomou-se o $\Omega_K = 0$, portanto $D_M = D_C$.

A distância do diâmetro angular (D_A) é definida como a razão entre o tamanho transversal físico de um objeto e seu tamanho angular em radianos e é obtida através da distância comóvel transversal pela Equação 3.9

$$D_A = \frac{D_M}{1+z}, \tag{3.9}$$

A distância de diâmetro angular é muito usada para converter separações angulares de objetos em separações reais das fontes. Para objetos a um mesmo *redshift*, determina-se a distância entre eles multiplicando-se a distância de diâmetro angular por sua distância ortodrômica na esfera celeste. Para o cálculo da distância angular, $\delta\theta$, utilizou-se a fórmula de Haversine descrita na Equação 3.10

$$\delta\theta = 2\arcsin\left(\sqrt{\sin^2\left(\frac{\varphi_2-\varphi_1}{2}\right)+\cos(\varphi_1)\cos(\varphi_2)\sin^2\left(\frac{\lambda_2-\lambda_1}{2}\right)}\right), \tag{3.10}$$

onde φ representa a declinação e λ a ascensão reta, em radianos. Neste trabalho, utilizou-se as rotinas em *Python* desenvolvidas por Wright (2006) para o cálculo das distâncias em cosmologia, especificamente, para a obtenção das distâncias de diâmetro angular e implementou-se um código para o cálculo das distâncias angulares pela fórmula de Haversine. A partir dessas distâncias, aplicou-se o código em uma amostra de 17840 galáxias os métodos descritos nas Seções 3.1 e 3.2 e analisou-se a distribuição dos estimadores, comparando os métodos entre si.

Capítulo 4

Resultados e Análise

Aplicando o código desenvolvido neste trabalho, obtivemos um catálogo ambiental para a amostra, tal que para cada objeto desta está associado valores de campos de densidade, obtidos por dois métodos diferentes.

4.1 Estimadores ambientais

Para o método dos *Nearest Neighbors* observa-se grande dispersão dos valores de cada estimador e entre os estimadores, principalmente para o primeiro vizinho (k=1), conforme descrito na Tabela 4.1.

O valor máximo para o $\sigma_{k=1}$ é bastante discrepante do valor médio e da mediana (50%). Além disso, o desvio padrão do estimador muito alto em relação à média mostra como esse estimador é bastante desbalanceado. O mesmo ocorre para os demais estimadores ($\sigma_{k=7}, \sigma_{k=10}, \sigma_{k=50}$), estes apresentam desvio padrão superiores à média, o que é característico de estimadores não robustos. Na Tabela 4.1, incluiu-se o valor do 90º percentil (90%), que representa o valor máximo do estimador incluindo 90% destes ordenados. Como estes valores para todos os estimadores se mostram bastante distantes do valor máximo, podemos concluir que todos estão sob influência de *outliers*, correspondentes a 10% da amostra. A distribuição dos estimadores para cada k-ésimo vizinho até o 90º percentil está ilustrada da Figura 4.1.

Convertendo os valores extremos dos estimadores em distâncias para analisar as escalas obtidas no método, temos os valores descritos na Tabela 4.2, onde observam-se distâncias da ordem de milhares de anos-luz até 60 Mpc.

Tabela 4.1 - Estatísticas dos estimadores ambientais calculados pelo método dos NN.

Número de vizinhos	1	7	10	50
Valor médio	203.32	5.63	4.98	3.46
Desvio padrão	3711.26	10.14	8.06	4.30
Valor mínimo	0.015	0.0045	0.0032	0.0045
25%	1.90	1.55	1.53	1.42
50%	5.34	2.92	2.81	2.37
75%	20.50	5.66	5.22	4.04
90%	87.89	11.69	10.15	6.54
Valor máximo	381480.75	219.03	164.22	89.70

Tabela 4.2 - Distâncias máxima e mínimo ao k-ésimo vizinho em Mpc.

k-ésimo vizinho	Distância mínima (Mpc)	Distância máxima (Mpc)
1	0.00091	4.65
7	0.10	22.18
10	0.14	31.52
50	0.42	59.62

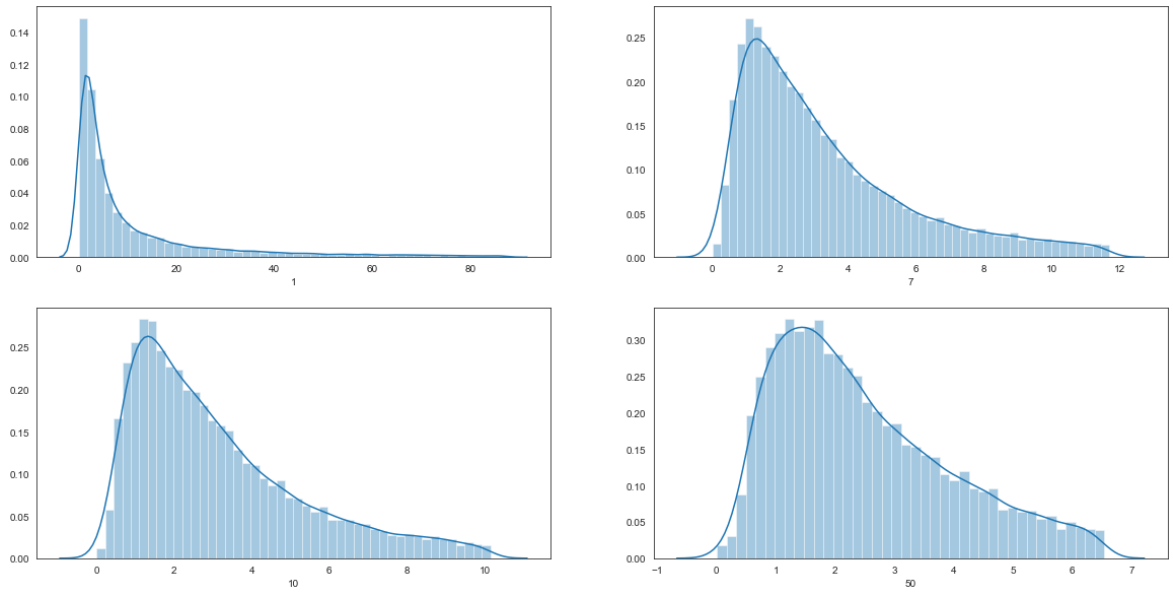

Figura 4.1: Distribuição dos estimadores até o 90º percentil pelo método dos NN para k=1, 7, 10 e 50. No eixo x constam-se os intervalos de valores para os estimadores e no eixo y sua frequência relativa.

O método das aberturas fixas se mostrou mais robusto em comparação com os NN, com distribuição dos estimadores menos dispersa. Na Tabela 4.3, observam-se as estatísticas de cada estimador. A medida que o raio da abertura cresce, a mediana e o valor médio aproximam, a menos de um desvio padrão de diferença. Para os raios menores (de 0.25, 1.0 e 2.0 Mpc) ainda há grande dispersão, representada pelo coeficiente de variação. Como esse método é discreto, temos um número bem definido de valores distintos para os estimadores, como ilustra o primeiro gráfico da Figura 4.2, que apresenta a distribuição dos estimadores para cada raio proposto.

Na Tabela 4.4 estão descritos os valores de densidades únicas para cada raio e os números máximo e mínimo de objetos encontrados dentro deste. Embora o método seja mais robusto, observa-se que, para todos os raios, há casos em que a galáxia têm apenas um vizinhos próximo, o que reforça o problema da discretização tal que há chances de encontrar galáxias sem nenhum vizinho em determinados raios.

Tabela 4.3 - Estatísticas dos estimadores ambientais calculados pelo método das aberturas fixas.

Raios em Mpc	0.25	1.0	2.0	5.0	8.0	10.0
Valor médio	8.59	2.60	1.93	1.42	1.17	1.03
Desvio padrão	6.48	2.58	1.53	0.83	0.61	0.50
Valor mínimo	5.09	0.32	0.078	0.018	0.0050	0.0032
25%	5.09	1.27	0.95	0.80	0.72	0.66
50%	5.09	1.91	1.51	1.26	1.07	0.96
75%	10.19	3.18	2.47	1.83	1.51	1.32
Valor máximo	91.67	28.33	12.33	4.94	3.25	2.57
Coeficiente de variação	75.46%	99.27%	79.20%	58.71%	51.78%	48.16%

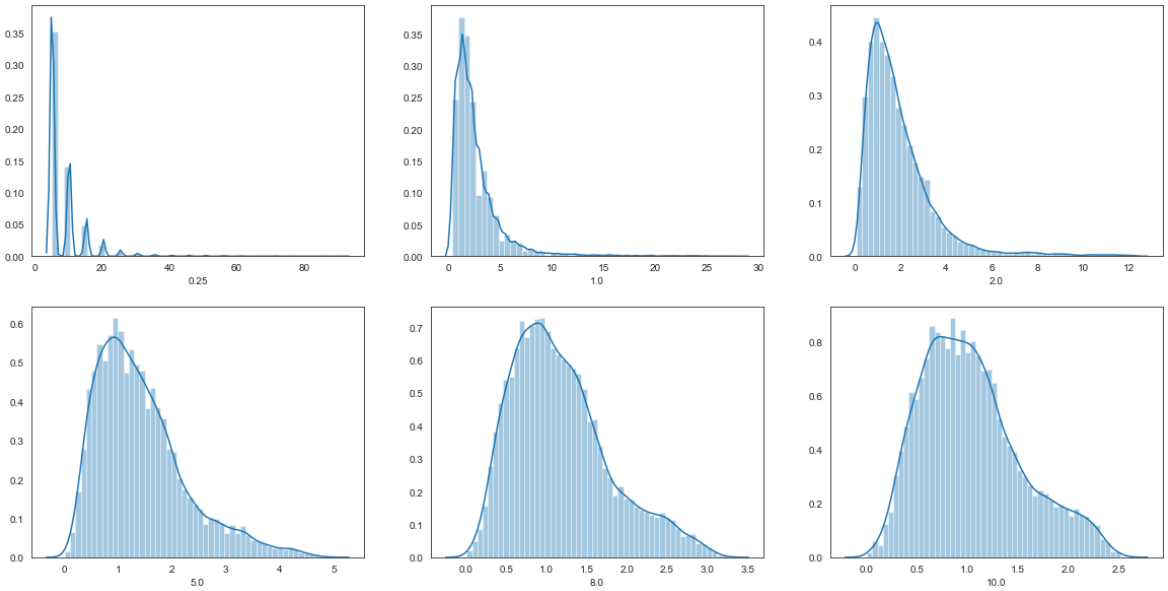

Figura 4.2: Distribuição dos estimadores pelo método das aberturas fixas, para os raios r = 0.25, 1.0, 2.0, 5.0, 9.0 e 10 Mpc. No eixo x constam-se os intervalos de valores para os estimadores e no eixo y sua frequência relativa

Tabela 4.4 - Valores únicos discretos dos estimadores, número máximo e mínimo de objetos encontrados em cada raio.

Raios em Mpc	Densidades únicas	nº de vizinhos máximo	nº de vizinhos mínimo
0.25	18	18	1
1.0	84	89	1
2.0	154	155	1
5.0	371	388	1
8.0	628	653	1
10.0	776	807	1

4.2 Comparação entre os métodos

Nesta seção, fazemos uma análise de correlação entre eles com o intuito de identificar aquele que melhor representa ambos os métodos. Na Tabela 4.5 estão descritos os coeficientes de correlação entre os estimadores de cada método.

As distribuições dos estimadores obtidas são assimétricas para todos os casos, por isso utilizamos o coeficiente de correlação de Spearman (Eq. 4.1) que não pressupõe distribuição normal para os dados e avalia também relações monótonas entre variáveis, sejam estas

lineares ou de outro tipo.

Seja o coeficiente de correlação de Spearman, r_s, dado por

$$r_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}, \tag{4.1}$$

onde d_i é a diferença entre dois postos de cada observação e n o número de observações. O coeficiente r_s é compreendido no intervalo $-1 < r_s < 1$, tal que correlações próximas de -1 e 1 são altas e em torno de 0 indicam nenhuma correlação.

No método dos NN, para os k's selecionados neste trabalho, estamos tratando de escalas típicas de grupos de galáxias que tipicamente apresentam diâmetros de até 2 Mpc. A alta correlação entre os estimadores k = 50 e r = 2.0 Mpc mostra que, para essas escalas, ambos estimadores têm o mesmo padrão, de modo que pode-se escolher um deles para o estudo das propriedades galácticas de acordo com o objetivo almejado sem muita perda de informação. A correlação do estimador k=50 vai decrescendo conforme o raio da abertura aumenta. Isso porque passamos a tratar de escalas típicos de aglomerados, que contém geralmente de 50 à ordem de milhar de galáxias (Figura 4.3).

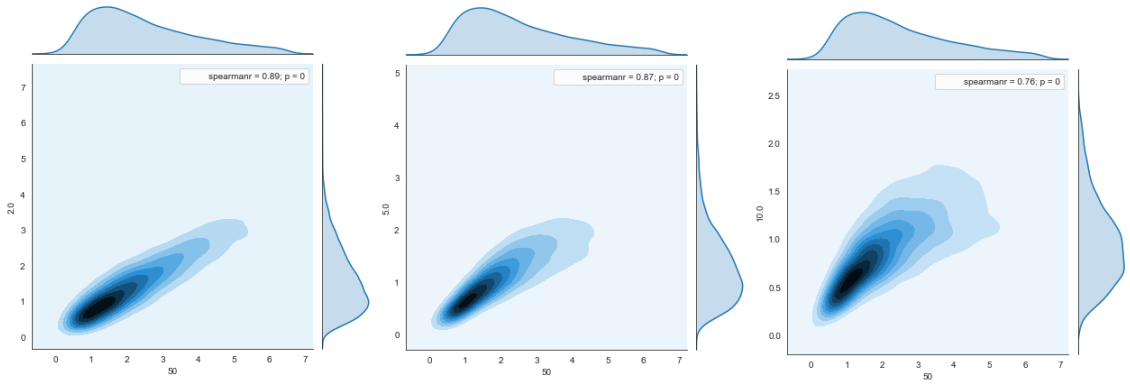

Figura 4.3: Densidade de *kernel* dos estimadores obtidos para r = 2.0, r = 5.0 e r= 10.0 Mpc, da esquerda para a direita, em função de k = 50. Acima de cada eixo, a distribuição de probabilidades dos valores dos estimadores.

Para k=1 e r = 0.25 Mpc, observa-se um coeficiente de correlação $r_s = 0.71$. Embora seja esperado que o primeiro vizinho mais próximo seja encontrado dentro desse raio, há valores de distâncias até 4.65 Mpc para o primeiro vizinho, como descrito na Tabela 4.2. Desse modo, para escalas inferiores a 2.0 Mpc, indica-se o uso dos estimadores calculados por ambos os métodos, para uma mensuração mais precisa do ambiente.

Tabela 4.5 - Correlação entre os estimadores.

	r = 0.25	r = 1.0	r = 2.0	r = 5.0	r = 8.0	r = 10.0
k = 1	0.717	0.529	0.448	0.373	0.345	0.331
k = 7	0.494	0.870	0.796	0.664	0.609	0.586
k = 10	0.466	0.874	0.839	0.703	0.647	0.622
k = 50	0.367	0.751	0.906	0.869	0.802	0.768

Capítulo 5

Conclusões

Como visto anteriormente, para a construção de um catálogo ambiental é necessário levar em considerações dois principais fatores:

(i) Os estimadores de ambiente devem variar de acordo com os parâmetros da amostra selecionada, de modo que um passo seguinte do trabalho em questão pode ser estudar como a variação de parâmetros como a magnitude limite, número de vizinhos e o *redshift* afetam os campos de densidades calculados em ambos os métodos revisados.

(ii) Os métodos possuem diferenças intrínsecas e, portanto, devem se escolhidos de acordo com o objetivo do estudo e tendo em vista a possibilidade do uso conjunto desses estimadores. O método nos NN utiliza uma escala variável contínua de distância, de milhares de anos-luz a dezenas de Mpc, e apresentam alta dispersão de valores, de modo que a distribuição de cada estimador apresenta assimetria positiva, com uma longa cauda de valores extremos. Desse modo, para estudar sua influência em propriedades de galáxias, deve se considerar uma normalização ou definição de faixas de valores de escala. Já o método das aberturas fixas apresentam resultados mais robustos, com distribuições mais simétricas e valores bem definidos, no entanto, por se tratar de um método discreto, apresenta dificuldades em mensurar efeitos de ambientes pouco densos, uma vez que é possível encontrar galáxias isoladas nos raios escolhidos. Além disso, ocorrem muitos casos de galáxias com apenas um vizinho no raio proposto, deixando uma visão míope a respeito da interação entre estes, o que pode ser complementado com os estimadores dos NN.

Por fim, foi estudada a correlação entre os estimadores calculados pelos dois métodos e encontrou-se alto coeficiente de Spearman para os valores k=50 e r=2.0 Mpc, escalas típicas de grupos de galáxias, o que significa que, para essas escalas, ambos estimadores apresentam valores com mesmo padrão, de modo que pode-se escolher um deles para o

estudo das propriedades galácticas de acordo com o objetivo almejado sem muita perda de informação. Já para estimadores de menores escalas, como k=1 e r = 1.0 Mpc, a correlação de mostra mais baixa, de modo que análise conjunta deve ser mais vantajosa, por apresentar visões complementares particulares de cada método.

Referências Bibliográficas

de Souza Oliveira Filho K., De Fatima Oliveira Saraiva M., Astronomia e astrofisica. Ed. Universidade/UFRGS, 2000

Etherington J., Thomas D., Measuring galaxy environments in large-scale photometric surveys, Monthly Notices of the Royal Astronomical Society, 2015, vol. 451, p. 660

Hogg D. W., , 1999 Distance measures in cosmology

Lai C.-C., Lin L., Jian H.-Y., Chiueh T.-H., Merson A., Baugh C. M., Foucaud S., Chen C.-W., Chen W.-P., CAN WE DETECT THE COLOR–DENSITY RELATION WITH PHOTOMETRIC REDSHIFTS?, The Astrophysical Journal, 2016, vol. 825, p. 40

Mendes de Oliveira C., Ribeiro T., Schoenell W., Kanaan A., Overzier R. A., Molino A., Sampedro L., Coelho P., Barbosa C. E., Cortesi A., et al. The Southern Photometric Local Universe Survey (S-PLUS): improved SEDs, morphologies, and redshifts with 12 optical filters, Monthly Notices of the Royal Astronomical Society, 2019, vol. 489, p. 241–267

Muldrew S. I., Croton D. J., Skibba R. A., Pearce F. R., Ann H. B., Baldry I. K., Brough S., Choi Y.-Y., Conselice C. J., Cowan N. B., Gallazzi A., Gray M. E., Grützbauch R., Li I.-H., Park C., Pilipenko S. V., Podgorzec B. J., Robotham A. S. G., Wilman D. J., Yang X., Zhang Y., Zibetti S., Measures of galaxy environment – I. What is 'environment'?, Monthly Notices of the Royal Astronomical Society, 2012, vol. 419, p. 2670

Sparke L. S., Gallagher III J. S., Galaxies in the Universe: An Introduction 2 edn. Cambridge University Press, 2007

Wright E., A Cosmology Calculator for the World Wide Web, Publications of the Astronomical Society of the Pacific, 2006, vol. 118, p. 1711–1715