

# Makine Öğrenmesi ile Kuazarların Rüzgar Değişimlerinin Tahmini

Seyit Hökelek<sup>1</sup>  , N. Filiz Ak<sup>2</sup> 

<sup>1</sup> Erciyes Üniversitesi, Fen Bilimleri Enstitüsü, Astronomi ve Uzay Bilimleri Ana bilim Dalı, 38039, Kayseri, Türkiye

<sup>2</sup> Erciyes Üniversitesi, Fen Fakültesi Astronomi ve Uzay Bilimleri Bölümü, 38039, Kayseri, Türkiye

Accepted: December 19, 2022. Revised: December 4, 2022. Received: November 13, 2022.

## Özet

Kuazarların %20'lik kısmında, merkezdeki süper kütleli karadeliğin etrafını çevreleyen yığılma diskinden dışa doğru madde akımlarının göstergesi olan Geniş Soğurma Çizgileri (GSÇ) gözlenmektedir. GSÇ şiddetlerinin zamana bağlı değişimi, rüzgar oluşum ve değişimine etki eden mekanizmaların anlaşılması için önemlidir. Bu çalışmada, optik/morötesi tayfında rüzgar yapısı olduğu belirlenmiş kuazarların ölçülen geniş soğurma çizgisi parametreleri makine öğrenmesi metotlarıyla sınıflandırılarak modellenmiş ve en iyi doğruluk değerini Linear Discriminant yönteminin verdiği görülmüştür. Oluşturulan en duyarlı modelin bile, rüzgarın değişim gösterip gösteremeyeceğini tahmin etmek için yeterli olmadığı görülmüştür. Bu sonuç, değişime sebep olan ana mekanizmaların rüzgarla taşınan soğurucu gazın iç özellikleri ile önemli bir bağlantısı olmadığını işaret etmektedir.

## Abstract

Approximately 20% of the quasars exhibit the Broad Absorption Lines (BALs) as a tracer of the outflows from the accretion disk of the central supermassive black holes. The time dependent variability of the BALs carries important information to understand the formation and evolution of the quasar winds. In this study, using a sample of BALs that have at least two epoch spectra, we develop a model to predict the variability of the winds based on the BAL parameters measured from their first epoch observation. We compare a few machine learning methods from which the Linear Discriminant method results the best accuracy. Even the most accurate method produces insufficient models to predict the BAL variability from the BAL parameters measured from the first epoch. This suggests that the internal properties of the absorbing gas carried by the outflows have, if any, a weak connection with the variations of BALs.

**Anahtar Kelimeler:** Galaxeis: active, Galaxies: quasars: absorption lines, methods: data analysis

## 1 Giriş

En yüksek ışınım gücüne sahip aktif galaksi çekirdeği sınıfı olan kuazarların yaklaşık %20 kadarlık kısmı güçlü rüzgarlara sahiptir. Kuazar çekirdek bölgesinin önemli bir bileşeni olan rüzgarların, merkezi süper kütleli karadeliğin üzerine madde akışını düzenlemede etkili olduğu düşünülmektedir. Bu düşünce yığılma diskindeki maddenin, rüzgarlar tarafından tetiklenen açıl momentum kaybıyla daha iç kısımlara hareket etmesine dayanmaktadır. Aynı zamanda rüzgarlar merkezi bölgedeki sıcak gazı konak galaksinin içine doğru taşıyor (Gibson ve diğ. 2009; Allen ve diğ. 2011; Borguet ve diğ. 2013; Filiz Ak ve diğ. 2013).

Kuazarlarda görülen bu güçlü rüzgarlar optik tayflarda genişliği  $2000 \text{ km s}^{-1}$  üzerinde olan geniş soğurma çizgileri (GSÇ) oluştururlar. Çok sayıda gözlemlerin incelenmesiyle yapılan çalışmalar kuazar tayflarındaki GSÇ'lerin kısa, orta ve uzun zaman ölçeklerinde değişimler gösterebildiğine işaret etmektedir (Capellupo ve diğ. 2011; Filiz Ak ve diğ. 2013; Rogerson ve diğ. 2018).

Sloan Digital Sky Survey (SDSS) kapsamında 2002 yılından bu yana toplam 750414 kuazar tayfsal ve fotometrik gözlemlerle doğrulanarak kataloglanmıştır (Lyke ve diğ.

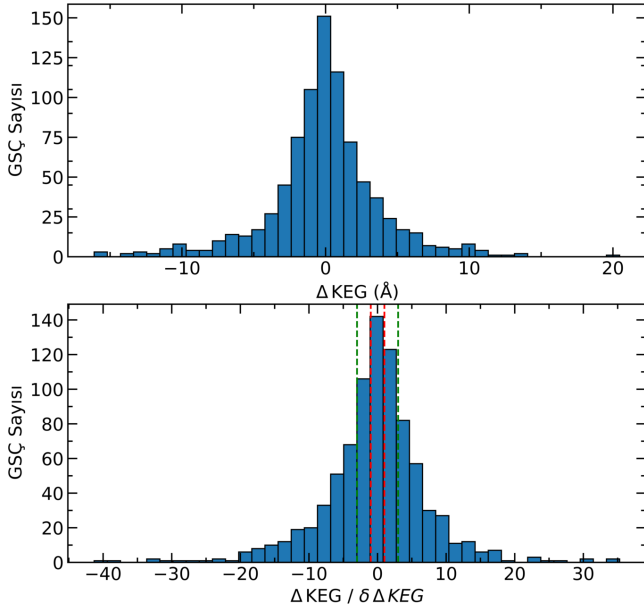
2020). SDSS tarafından en az 2 farklı zamanda elde edilmiş tayfların kıyaslanmasıyla GSÇ yapılarının değişimleri belirlenebilmektedir. Filiz Ak ve diğ. (2013)'te elde edilen sonuçlar bazı GSÇ yapılarının çok güçlü değişim gösterirken bazılarının ise kayda değer bir değişim göstermediğini işaret etmektedir.

GSÇ yapılarının değişimine neden olan mekanizmaların araştırılması için yapılan çalışmalar birkaç model üzerinde yoğunlaşmıştır. Bu modeller arasında GSÇ değişimlerinin rüzgarlar tarafından taşınan soğurucu maddenin (a) kendi yapısal özelliklerine bağlı olarak, (b) iyonizasyon durumundaki değişikliklere bağlı olarak ve (c) merkez etrafında yığılma diskiyle beraber dolanma hareketine bağlı olarak gerçekleştiği fikirleri öne çıkmaktadır (Arav ve diğ. 1999; Hamann ve diğ. 2000; Filiz Ak ve diğ. 2012, 2013, 2014; Rogerson ve diğ. 2016, 2018).

GSÇ değişimlerinin GSÇ'lerin kendi özelliklerine bağlı olup olmadığını incelemek için bu çalışmada değişim gösteren ve göstermeyen GSÇ yapılarının özellikleri kıyaslanmıştır. Bir yaklaşım olarak GSÇ yapılarının eş değer genişlik, derinlik, genişlik ve hız gibi parametrelerini kullanarak değişim gösterip göstermeyeceklerini tahmin etmeye yönelik makine öğrenmesi modelleri test edilmiştir. Böyle bir modelin yüksek duyarlılıkla elde edilebilmesi bu parametrelerin değişim üzerinde belirleyici bir etkisi olduğunu destekler nitelikte olacaktır.

Çalışmada kullanılan örneklemin seçimi ve sınıflandırılması,

\* sthokelek@gmail.com



**Şekil 1.** Bu çalışmada ele alınan 454 GSC yapısı için elde edilen kesirsel eşdeğer genişlik değişimi  $\Delta KEG$  parametresinin dağılımı üst panelde verilmiştir. Alt panelde ise değişen ve değişmeyen GSC yapılarının seçim kriteri olarak kullanılan  $\Delta KEG/\delta\Delta KEG$  parametresinin dağılımı verilmiştir. Şekilde verilen kırmızı kesikli çizgilerin arasında kalanlar değişmeyen GSC ve yeşil dik çizgilerin dışında kalanlar ise değişen GSC olarak tanımlanmıştır.

kullanılan makine öğrenmesi yöntemleri ikinci bölümde ve elde edilen sonuçlar ise üçüncü bölümde verilmiştir.

## 2 Verilerin Elde Edilmesi, Sınıflanması ve Kullanılan Yöntemler

Bu çalışmada kullanılan parametreler, parametrelerin ölçüm yöntemleri ve hata hesapları bu kısımda ele alınmıştır. Ayrıca yapılan ölçümlere göre GSC yapılarının "Değişen" ve "Değişmeyen" şeklinde gruplanması için kullanılan kriterler belirlenmiştir. Son kısımda, kullanılan makine öğrenme yöntemleri hakkında bilgi verilmiştir.

### 2.1 Verilerin Elde Edilmesi

Kuazarların GSC yapılarındaki değişimlerin incelenmesi için Filiz Ak ve diğ. (2014) tarafından elde edilen veriler kullanılmıştır. Filiz Ak ve diğ. (2014) çalışmasında, SDSS tarafından en az iki tayfsal gözlemi alınmış olan kuazarların GSC yapılarının hız ( $v_m$ ), eşdeğer genişlik ( $EG$ ), derinlik ( $d$ ) ve genişlik ( $\Delta v$ ) ölçümleri yapılmıştır. Örnekleme yer alan toplam 614 kuazarın iki farklı zamanda alınmış tayflarında belirlenen toplam 851 GSC için bu parametrelerin ölçümleri Filiz Ak ve diğ. (2014) tarafından yapılmıştır. Yapılan ölçümler arasındaki farklar ise her bir parametrenin değişimini tanımlamaktadır. Örneğin EG değişimi

$$\Delta EG = EG_2 - EG_1$$

şeklinde tanımlanmıştır. Parametrelerin değişimlerinin hataları, ölçüm hatalarının yayılımı ile hesaplanmıştır. Örneğin, bir GSC yapısının şiddetinin belirteci olan eşdeğer genişlik değişiminin hatası

$$\delta\Delta EG = \sqrt{\delta EG_2^2 + \delta EG_1^2}$$

şeklinde hesaplanmıştır.

Bu çalışmada özellikle güçlü rüzgar yapılarının değişimlerinin dikkate alınması hedeflendiğinden, örnekleme zayıf GSC yapıları barındıran kuazarlar kullanılmamıştır. Kuazarların tayfında geniş soğurma çizgileri, C IV, Si IV ve Al III elementlerinin salma çizgilerinin mavi tarafında görülmektedir. Bu çalışma için seçilen kuazarlarda, C IV, Si IV ve Al III soğurma çizgileri  $2000 \text{ km s}^{-1}$  üzerinde olanlar tercih edilmiştir. Bu kriter kullanılan kuazar sayısının toplam 424 ve ölçümleri değerlendirilen GSC yapılarının toplam 454 olmasını sağlamıştır.

### 2.2 Değişen ve Değişmeyen GSC Kriterleri

Ana örneklem içinde yer alan 454 GSC'nin özelliklerinin değişim gösterip göstermeyeceğini inceleyebilmek için öncelikle değişen GSC ve değişmeyen GSC tanımlamalarının yapılması gerekmektedir. Bir GSC'nin değişimi, şiddetinin göstergesi olan eşdeğer genişlik ( $EG$ ) parametresinin zamana bağlı değişimi ile ölçülür. Ancak, eşdeğer genişlik değişimini kesirsel olarak ifade etmek daha anlamlı sonuçlar vermektedir (Filiz Ak ve diğ. 2013). Bunun sebebi, eşdeğer genişliği çok küçük ve çok büyük iki örnek ele alındığında,  $\Delta EG=1 \text{ \AA}$  olarak ölçülen değişim, GSC'nin kendi eşdeğer genişliğine oranlandığında değişimin şiddeti kıyaslanabilir. Bu nedenle değişimi tanımlamak için kesirsel değişim parametresi olan  $\Delta KEG=\Delta EG/EG_1$  tanımlanmıştır. Bu parametre için hata hesabı ise

$$\frac{\delta\Delta KEG}{\Delta KEG} = \left[ \left( \frac{\delta\Delta EG}{\Delta EG} \right)^2 + \left( \frac{\delta EG_1}{EG_1} \right)^2 \right]^{1/2} \quad (1)$$

ile hesaplanmıştır. Burada  $\delta$  notasyonu ölçüm hataları için kullanılmıştır.

Bu çalışmada yer alan GSC'lerin  $\Delta KEG$  ve  $\frac{\delta\Delta KEG}{\Delta KEG}$  dağılımları, Şekil 1'de verilmiştir. Dağılımların sıfır değeri etrafında simetrik bir yapıya sahip olması, örneklem içerisinde zayıflanan ve şiddetlenen GSC yapılarının eşit olduğunu ortaya koymaktadır.  $\Delta KEG$  dağılımının ortalaması,  $-0.0623 \text{ \AA}$ , ve standart sapması  $3.9325 \text{ \AA}$ 'dır.  $\delta\Delta KEG/\Delta KEG$  dağılımının ortalama ve standart sapma değerleri sırasıyla  $-0.3716$  ve  $7.7112$  olarak elde edilmiştir.

Bir GSC yapısının ölçüm hatalarına göre anlamlı bir kesirsel değişim gösterip göstermediğini dikkate almak, seçilen örneklemelerin daha güvenilir olmasını sağlamaktadır. Bu nedenle, aşağıdaki kriterler belirlenmiş ve buna göre sınıflama yapılmıştır:

$$\text{Değişen GSC: } \left| \frac{\Delta KEG}{\delta\Delta KEG} \right| > 3 \text{ (değişimi } 3\sigma \text{ dan büyük)}$$

$$\text{Değişmeyen GSC: } \left| \frac{\Delta KEG}{\delta\Delta KEG} \right| < 1 \text{ (değişimi } 1\sigma \text{ dan küçük)}$$

Yukarıda verdiğimiz tanımlamalara göre verileri gruplandırdığımızda ana örneklemimizdeki 454 GSC'lerin kriterler uygulandıktan sonra 332'ye azalmıştır. Bu GSC'lerin 238 tanesi Değişen ve 94 tanesi Değişmeyen GSC olarak sınıflanmıştır. Arada kalanlar ise iki sınıfın daha net ayrışması için hesaplamaların dışında bırakılmıştır.

Çalışmada kullanılan Değişen ve Değişmeyen GSC yapılarının genişlik, merkezi hız, derinlik ve eşdeğer genişlik gibi parametrelerinin ortalama değerleri Tablo 1'de verilmiştir.

**Çizelge 1.** Değişen ve Değişmeyen GŞÇ'lerin ortalama özellikleri

Parametre	Birim	Değişen GŞÇ	Değişmeyen GŞÇ
$\Delta v$	km s <sup>-1</sup>	9034.70	7345.63
$v_m$	km s <sup>-1</sup>	-11757.60	-12774.12
$d$		0.43	0.36
EG	Å	23.69	14.25

### 2.3 Makine Öğrenme Yöntemleri

GŞÇ yapılarının şiddetinin zamanla kayda değer bir değişim gösterip göstermeyeceğini makine öğrenmesi yöntemleri ile tahmin edebilmek için Değişen ve Değişmeyen olarak sınıflanan GŞÇ yapılarının hız genişliği ( $\Delta v$ ), ağırlıklı merkezi hız ( $v_m$ ), derinlik ( $d$ ) ve eşdeğer genişlik ( $EG$ ) parametreleri kullanılmıştır. Makine öğrenmesi yöntemlerinde bu parametrelerin her biri bir boyut olarak ele alınmaktadır. Ayrıca değişen ve değişmeyen tanımlamaları da bilinen sınıf parameresi olarak kullanılmaktadır.

Doğru sınıflama tahmini yapabilen bir model üretmek için Doğrusal Ayırma (linear discriminant – LD), Destek Vektör Modeli (support vector model – SVM) ve Çekirdek Bayes (Kernel Naive Bayes – KNB) yöntemleri dikkate alınmıştır.

Doğrusal Ayırma Yöntemi, veri uzayı içerisindeki boyutları doğrusal olarak ayırt ederek boyut azaltması sağlayan bir yöntemdir. Bunun için bir sınıf içerisindeki varyansı olabildiğince küçültmeye çalışırken sınıflar arasındaki varyansı olabildiğince artırmaya çalışır (Feigelson & Babu 2012).

Çekirdek Bayes, veri uzayı içerisindeki boyutlar birbirinden bağımsız kabul etmektedir ve her bir boyut için bayes istatistiğini kullanarak olasılık hesaplayan bir yöntemdir. Ele alınan veride gerçekte boyutlar arasında korelasyonun olduğu durumlarda bile bu yöntem başarılı sınıflandırmalar yapabilmektedir (du Buisson ve diğ. 2015).

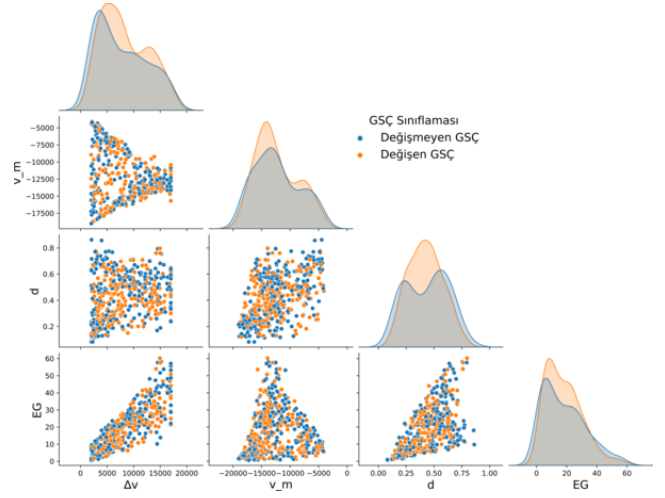
Destek vektör modeli, sınıflar arasındaki sınırları belirlerken hem sınıfların doğru ayrıştırılmasına hem de aralarındaki uzaklığın maksimum olmasına dikkat eden iteratif bir yöntemdir. Tanımlanan sınıfların birbirlerine en yakın marjinlerini kullanarak destek vektörleri üretir ve iterasyonlarla destek vektörlerinin orta değerlerini sınıflama kriteri olarak belirler. Bu yöntem ile ortaya keskin ve ihtimale dayanmayan sınıflar ortaya çıkar. (Ivezić ve diğ. 2014).

Bu modeller Şekil 2'da gösterilen sınıfların dağılımları dikkate alınarak seçilmiştir. Sınıfların tüm boyutlarda yakın dağılımlara neden olması nedeniyle karar ağacı tabanlı yöntemlerden kaçınılmıştır.

Bu modellerin tamamı için değişen ve değişmeyen GŞÇ'ler içinden ratsgele seçilen %15'lik kısmı modelin eğitilmesi için ayrılmaktadır.

Modellerin eğitilmesi için kullanılan veri setinin hem tüm örnekleri temsil edecek kadar yeterli büyüklükte hem de marjin değerlerinin etkisini azaltacak kadar küçük olması gerektiği düşünüldüğünden toplam veri setinin %85'i test için ayrılmıştır. Sonuçların güvenilirliğini artırabilmemiz için rastgele test örnekleme seçilerek modellerin oluşturulması 5 kez tekrar edilmiştir. Çapraz doğrulama (k-fold) sayısı 5 alınarak, örneklem içerisindeki marjin değerlerinin doğru ağırlıkta dikkate alınması sağlanmıştır (Ivezić ve diğ. 2014).

Bu analizler için MATLAB programının sınıflama öğrenme uygulaması kullanılmıştır.



**Şekil 2.** Çalışmada ele alınan 454 GŞÇ yapısının parametrelerinin değişen ve değişmeyenler için dağılımı en üstte verilmiştir. Alt panellerde parametrelerin birbirleriyle ilişkileri gösterilmiştir. Şekilde görüldüğü üzere parametreler birbirlerinden kesin olarak ayrılmamaktadırlar.

### 3 Sınıflama Sonuçları

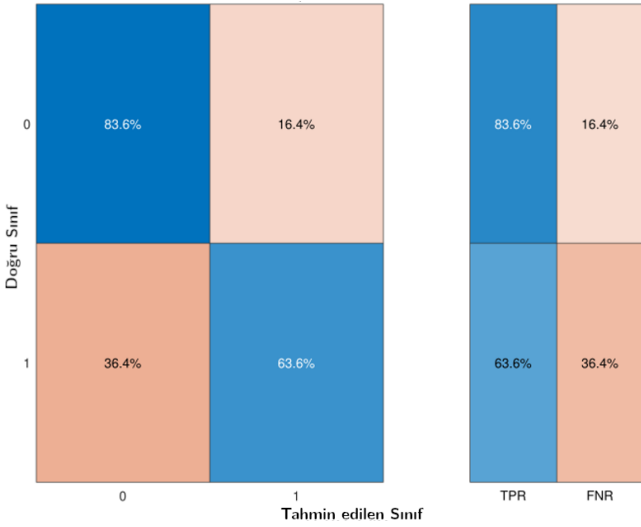
Örnekleme içinde yer alan Değişen ve Değişmeyen GŞÇ yapılarının sınıflama tahmini için ilk tayf gözleminde elde edilen GŞÇ'nin derinlik, genişlik, hız ve eşdeğer genişlik ölçümleri girdi olarak verilmiş ve sınıflama tahminleri yapılmıştır. Gerçekleştirilen tahmin sınıflamalarından en iyi sonucu veren yöntem LD olarak belirlenmiştir.

Şekil 3'de, LD yöntemi kullanılarak oluşturulan modelin sonuçları verilmiştir. Buna göre, Gerçek sınıfı değişen (1) olan GŞÇ'lerin %83.6'sı için ve gerçek sınıfı değişmeyen (0) olan GŞÇ'lerin %63.6'sı için doğru sınıflama tahmini yapılabilmektedir. Yanlış sınıflama oranları ise değişen ve değişmeyen GŞÇ'ler için sırasıyla %16.4 ve %36.4 bulunmuştur. Modelin Gerçek Doğru Oranı (True Positive Rate – TPR) ve Yanlış Hatalı Oranı (False Negative Rate – FNR) için dağılımlar Şekil 3'nin sağ panelinde verilmiştir.

### 4 Sonuç

Filiz Ak ve diğ. (2014) makalesinde, 614 kuazarın farklı zamanlarda alınmış 851 tayfında yer alan GŞÇ yapılarının ölçümleri verilmiştir. Bu çalışmada, Filiz Ak ve diğ. (2014) çalışmasında yer alan, CIV, SiIV ve AIII bölgesinde bulunan GŞÇ yapılarından, genişliği 2000 km s<sup>-1</sup> üzerinde olan toplam 454 tanesi dikkate alınmıştır. Birinci ve ikinci tayftan ölçülen eşdeğer genişliklerin kesirsel değişimleri dikkate alınarak 238 Değişen GŞÇ ve 94 Değişmeyen GŞÇ belirlenmiştir.

GŞÇ'lerin ilk tayflarından elde edilen hız genişliği ( $\Delta v$ ), ağırlıklı merkezi hız ( $v_m$ ), derinlik ( $d$ ) ve eşdeğer genişlik ( $EG$ ) parametreleri dikkate alınarak değişim gösterip göstermeyeceklerini tahmin etmek için makine öğrenmesi yöntemleri kullanılmıştır. Listelenen parametrelerin veri seti olarak kullanıldığı modellerde toplam veri sayısının rastgele seçilen %15 kadarı modelin eğitimi için kullanılmıştır. Çapraz doğrulama için kullanılan 5 kat ile farklı yöntemlerin sonuçları kıyaslanmıştır. LD, SVM ve KNB yöntemlerinin birbirine yakın sonuçlar verdiği ve bunların içinden en yüksek doğruluk oranının (%73.6) LD yönteminde elde edildiği görülmüştür.



**Şekil 3.** LD yöntemi kullanılarak, GŞÇ yapılarının özellikleri ile değışim sınıflaması tahmin eden modelin sonuçları yüzdelik olarak verilmiştir.

LD yöntemi kullanılarak oluşturulan modelin genel doğruluk oranı görece yüksek bir deęer olsa da FNR deęerinin %20 üzerinde olması, kullanılan veri setinin sınıflamanın doğru tahmini için yeterli bilgi içermediğinin bir göstergesidir.

Bu sonuçlar, kuazarların GŞÇ parametrelerinin rüzgarın belirgin bir değışim gösterip gösteremeyeceğini tahmin etmek için yeterli bilgi sağlamadığına işaret etmektedir. Diğer bir ifade ile rüzgar değışimine sebep olan mekanizmanın rüzgarın kendi özellikleri ile sınırlı olmadığı düşünülmektedir. Elde ettiğimiz sonuçlar, rüzgar değışimlerine neden olan ana mekanizmanın soęurucu gazın özelliklerinden bağımsız bir dış etken olduğu görüşünü destekler niteliktedir. Bu dış etkenlerin neler olduğu ve değışime ne ölçüde etki ettiği konusunda araştırmalar yapılması kuazar rüzgarlarının oluşum ve evrim süreçlerinin anlaşılması için önemli bilgiler sağlayacaktır.

### Kaynaklar

- Allen J. T., Hewett P. C., Maddox N., Richards G. T., Belokurov V., 2011, *MNRAS*, 410, 860
- Arav N., Becker R. H., Laurent-Muehleisen S. A., Gregg M. D., White R. L., Brotherton M. S., de Kool M., 1999, *ApJ*, 524, 566
- Borguet B. C. J., Arav N., Edmonds D., Chamberlain C., Benn C., 2013, *ApJ*, 762, 49
- Capellupo D. M., Hamann F., Shields J. C., Rodríguez Hidalgo P., Barlow T. A., 2011, *MNRAS*, 413, 908
- Feigelson E. D., Babu G. J., 2012, *Modern Statistical Methods for Astronomy: With R Applications*. Cambridge University Press, doi:10.1017/CBO9781139015653
- Filiz Ak N., ve diğ., 2012, *ApJ*, 757, 114
- Filiz Ak N., ve diğ., 2013, *ApJ*, 777, 168
- Filiz Ak N., ve diğ., 2014, *ApJ*, 791, 88
- Gibson R. R., ve diğ., 2009, *ApJ*, 692, 758
- Hamann F. W., Netzer H., Shields J. C., 2000, *ApJ*, 536, 101
- Ivezić Ž., Connolly A. J., VanderPlas J. T., Gray A., 2014, *Statistics, Data Mining, and Machine Learning in Astronomy: A Practical Python Guide for the Analysis of Survey Data*. Princeton University Press, doi:10.1515/9781400848911
- Lyke B. W., ve diğ., 2020, *ApJS*, 250, 8
- Rogerson J. A., Hall P. B., Rodríguez Hidalgo P., Pirkola P., Brandt W. N., Filiz Ak N., 2016, *MNRAS*, 457, 405

Rogerson J. A., Hall P. B., Ahmed N. S., Rodríguez Hidalgo P., Brandt W. N., Filiz Ak N., 2018, *ApJ*, 862, 22

du Buisson L., Sivanandam N., Bassett B. A., Smith M., 2015, *MNRAS*, 454, 2026

Access:

M23-0381: *Turkish J.A&A* — Vol.4, Issue 3.