

Yükseköğretimde Kalitenin Tamamlayıcı Haritası: Eksik Veriler ve Değerlendirme Stratejileri

Complementary Map of Quality in Higher Education: Missing Data and Evaluation Strategies

Gözde Bozkurt¹ , Volkan Öngel² 

¹ İstanbul Beykent Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, İktisat Bölümü, İstanbul, Türkiye

² İstanbul Beykent Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, Uluslararası Ticaret ve Finans Bölümü, İstanbul, Türkiye

Özet

Bu çalışmada, Yükseköğretim Kalite Kurulu'nun (YÖKAK) resmi internet sitesinde paylaşılan Gösterge Değerlendirme Raporu kapsamındaki üniversitelerin temel göstergelerindeki eksik veriler incelenmiştir. Üniversitelerin gösterge raporları, stratejik yönetim ve kalite güvencesi açısından önemli bir yer tutmaktadır. Bu bağlamda, eksik verilerin yapısı tamamen rastgele kayıp (MCAR), rastgele eksik (MAR) ve rastgele olmayan kayıp (MNAR) kategorileri çerçevesinde oluşturulan panel veri seti aracılığıyla analiz edilmiştir. Bulgular, eksik verilerin ağırlıklı olarak rastgele bir yapıya sahip olduğunu göstermektedir. Çalışmada, eksik verilerin rastgele veya sistematik olup olmadığının belirlenmesinin, analiz sürecinde doğru yöntemlerin seçilmesi açısından kritik olduğu vurgulanmaktadır. Bu doğrultuda, yükseköğretim kurumlarının performanslarını daha sağlıklı değerlendirebilmesi ve kaliteyi artırmaya yönelik stratejiler geliştirilebilmesi için uygun veri tamamlama yöntemleri üzerine bir tartışma yürütülmüştür.

Anahtar Kelimeler: Yükseköğretimde Kalite, Eksik Veri Analizi, Gösterge Değerlendirme, Stratejik Yönetim, Panel Veri

Abstract

In this study, missing data in the basic indicators of universities within the scope of the Indicator Evaluation Report shared on the official website of the Higher Education Quality Council (YÖKAK) were examined. Indicator reports of universities have an important place in terms of strategic management and quality assurance. In this context, the structure of missing data was analyzed through a panel data set created within the framework of completely missing at random (MCAR), missing at random (MAR) and non-random missing (MNAR) categories. The findings show that the missing data predominantly have a random structure. The study emphasizes that determining whether the missing data is random or systematic is critical in terms of choosing the right methods in the analysis process. In this context, a discussion was carried out on appropriate data completion methods so that higher education institutions can evaluate their performance more accurately and develop strategies to increase quality.

Keywords: Quality in Higher Education, Missing Data Analysis, Indicator Evaluation, Strategic Management, Panel Data

Yükseköğretim kurumları, sürdürülebilir kalite standartlarına ulaşmak ve kurumsal gelişmelerini güçlendirmek adına stratejik planlamalarını ulusal ve uluslararası kalite güvencesi ilkeleri doğrultusunda şekillendirmektedir. Türkiye'de bu sürece yön veren temel kurumsal aktörlerden biri olan Yükseköğretim Kalite Kurulu (YÖKAK), yükseköğretim sisteminde kalite kültürünü kurumsallaştırmak amacıyla eğitim, araştırma, idari işler ve toplumsal hizmet alanlarında performans göstergelerine dayalı izleme ve değerlendirme mekanizmaları geliştirmiştir. Bu faaliyet alanları, Avrupa Yükseköğretim Alanı çerçevesinde benimsenen ESG

2015 iç kalite güvencesi standartlarına paralel olarak yapılandırılmakta ve kalite yönetiminin sadece öğretimle sınırlı kalmayıp tüm kurumsal süreçleri kapsamı gerektiğini vurgulamaktadır (ENQA, 2015).

Yükseköğretim sisteminin performansını değerlendirmede kullanılan istatistiksel göstergeler, kurumsal stratejilerin tasarlanması, izlenmesi ve iyileştirilmesi açısından yönetsel karar alma süreçlerinde belirleyici bir rol oynamaktadır (Harvey & Green, 1993; García-Aracil & Palomares-Montero, 2013). Ancak, söz konusu göstergelerin değerlendirilmesinde sıkça karşılaşılan eksik veri problemi,

İletişim / Correspondence:

Dr. Öğr. Üyesi Gözde Bozkurt
İstanbul Beykent Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, İktisat Bölümü, İstanbul / Türkiye
e-posta: gozdebozkurt@beykent.edu.tr

Yükseköğretim Dergisi / TÜBA Higher Education Research/Review (TÜBA-HER), 15(ÖS), 141-149. © 2025 TÜBA
Geliş tarihi / Received: Ağustos / August 26, 2024; Kabul tarihi / Accepted: Nisan / April 17, 2025

Bu makalenin atıf künyesi / How to cite this article: Bozkurt, G. & Öngel, V. (2025). Yükseköğretimde Kalitenin Tamamlayıcı Haritası: Eksik Veriler ve Değerlendirme Stratejileri. *Yükseköğretim Dergisi*, 15(ÖS), 151-159. <https://doi.org/10.53478/yuksekogretim.1538863>

ORCID: G. Bozkurt: 0000-0001-8413-1099; V. Öngel: 0000-0001-8881-2465

analiz sonuçlarının geçerliliğini ve karar süreçlerinin doğruluğunu tehdit etmektedir. Eksik veriler, yalnızca bilgi kaybına yol açmakla kalmaz, aynı zamanda tahmin hataları, model sapmaları ve politika geliştirmede yanlış yönlendirmelere neden olabilmektedir (Allison, 2002; Rubin, 1976; Enders, 2010; Little & Rubin, 2020). Bu nedenle, eksik veri yapısının istatistiksel olarak doğru şekilde belirlenmesi ve buna uygun analiz stratejilerinin geliştirilmesi büyük önem taşımaktadır.

Eksik veri yapısı literatürde, Rubin (1976) tarafından ortaya konan üçlü sınıflandırma ile tanımlanmaktadır: Tamamen rastgele eksik veri (MCAR), rastgele eksik veri (MAR) ve rastgele olmayan eksik veri (MNAR). Bu yapıların her biri farklı analiz stratejileri gerektirmekte olup, özellikle MNAR durumunda, eksikliğin sistematik nedenlerle ortaya çıkması analizlerde ciddi yanlışlıklar yaratabilmektedir (Leke & Marwala, 2019). Bu bağlamda, çalışmalarda kullanılan veri setinin eksiklik yapısı analiz edilmeden doğrudan tamamlayıcı tekniklerin uygulanması, modelin güvenilirliğini zedeleyebilmektedir.

Bu çalışma, YÖKAK Gösterge Değerlendirme Raporları temelinde oluşturulan panel veri seti üzerinden, eksik verilerin sistematik bir yapıya sahip olup olmadığını analiz etmeyi amaçlamaktadır. Örneklem olarak Türkiye'deki 208 üniversitenin (126 devlet, 82 vakıf) 2015–2022 yıllarına ait performans göstergeleri incelenmiş ve özellikle yabancı öğrenci sayısı verisinde gözlemlenen eksiklikler üzerine odaklanılmıştır. Yabancı öğrenci sayısı, yükseköğretimde uluslararasılaşma süreçlerinin göstergesi olması açısından stratejik öneme sahiptir (de Wit, 2020; Knight, 2004;). Bu bağlamda, öğrenci çeşitliliği, kültürel etkileşim, uluslararası araştırma işbirlikleri ve ekonomik katkılar gibi çok boyutlu etkiler oluşturmakta, kurumsal performans pozitif katkı sağlamaktadır (Baydemir & Bolat, 2024; Gümüş & Onurlubaş, 2019; Lee, 2010). Ancak, yabancı öğrencilere ilişkin verilerdeki eksiklikler, yalnızca analiz güvenilirliğini değil, aynı zamanda stratejik planlama, politika üretimi ve kalite değerlendirme süreçlerinin bütüncül biçimde yönetilmesini de güçleştirmektedir. Örneğin, öğrenci destek hizmetlerinin güçlendirilmesi, araştırma işbirliklerinin planlanması ve toplumsal katkı hedeflerinin belirlenmesi gibi alanlarda, yabancı öğrenci verilerinin eksikliği, kurumların uluslararasılaşma performanslarını doğru şekilde izleyememelerine neden olabilmektedir. Eksik veriler bu bağlamda yalnızca teknik bir sorun değil, aynı zamanda veriye dayalı yönetim açısından kritik bir zafiyet olarak karşımıza çıkmaktadır (Batini & Scannapieca, 2016; Gartner, 2017; Karr ve ark., 2006).

Çalışmada, eksik veri yapısı özellikle Little's MCAR testi ile analiz edilerek, rastgelelik düzeyi belirlenmiş ve verinin sistematik olarak eksik olup olmadığı değerlendirilmiştir. Eğer eksik veriler MCAR dışı bir yapıya sahipse (örneğin MAR veya MNAR), bu eksikliklerin belirli üniversite

türlerine, yıllara veya göstergelere sistematik olarak yoğunlaştığı anlamına gelir ki bu da analiz sonuçlarının genellenebilirliğini ciddi biçimde sınırlar. Bu bağlamda, eksik veri yapısının doğru tespiti ve buna uygun eksik veri yönetimi stratejilerinin uygulanması, hem kurum içi kalite güvencesi hem de ulusal yükseköğretim politikaları açısından kritik öneme sahiptir.

Sonuç olarak bu çalışma, yükseköğretim kurumlarının kalite göstergelerine ilişkin eksik veri yapısını sistematik biçimde analiz ederek, eksikliklerin veri yönetimi süreçlerine ve değerlendirme modellerine olan etkilerini ortaya koymayı amaçlamaktadır. Çalışmanın sonucunda, eksik veri sorununa yönelik öneriler ve uygulama stratejileri sunulurken, yükseköğretimde kalite güvencesine katkı sağlanması hedeflenmektedir.

Literatür

Yükseköğretim kurumlarında kurumsal gelişimi sürdürülebilir kılmak ve toplumsal sorumlulukları etkin biçimde yerine getirmek amacıyla uygulanan kalite güvencesi sistemleri, Türkiye'de son yıllarda Yükseköğretim Kalite Kurulu (YÖKAK) öncülüğünde yeniden yapılandırılmıştır. YÖKAK, yükseköğretim kurumlarının eğitim-öğretim, araştırma, toplumsal katkı ve idari süreçlerine yönelik iç ve dış değerlendirme mekanizmaları geliştirerek kalite kültürünün kurumsallaşmasını hedeflemektedir.

Bu bağlamda, stratejik yönetim, kurumların misyon, vizyon ve temel değerleri doğrultusunda hedeflerine ulaşmasını sağlamak amacıyla planlama, uygulama, izleme ve değerlendirme süreçlerini içeren bütüncül bir yönetim anlayışı sunmaktadır. Stratejik yönetim ile kalite güvencesi arasındaki ilişki, karar alma süreçlerinin sürekli iyileştirme temelinde sistematik biçimde yürütülmesini sağlamaktadır. Kalite güvencesi yalnızca kurum faaliyetlerinin belirlenen hedef ve standartlara uygunluğunu değil, aynı zamanda bu faaliyetlerin şeffaf, izlenebilir ve geliştirilebilir olmasını da güvence altına almaktadır (ENQA, 2015; Harvey & Green, 1993).

Yükseköğretimde stratejik yönetim ve iç kalite güvencesi, kurumların performanslarını sürdürülebilir şekilde izlemelerine ve geliştirmelerine olanak tanımaktadır. Bu sistemler, yükseköğretim kurumlarının eğitim, araştırma ve toplumsal katkı alanlarındaki faaliyetlerini belirli standartlar çerçevesinde düzenli olarak değerlendirmesine olanak sağlamaktadır. Performans göstergeleri bu bağlamda kilit rol oynamakta olup; öğrenci memnuniyeti, mezuniyet oranları, akademik yayın sayısı, istihdam oranları ve uluslararası öğrenci sayısı gibi alanlara yayılmaktadır (Hazelkorn, 2015; Jongbloed ve ark., 2008; Sadırlı, 2023). Türkiye'de bu göstergeler, YÖKAK tarafından yayımlanan Gösterge Değerlendirme Raporları ile düzenli olarak izlenmekte ve raporlanmaktadır.



Uluslararası düzeyde ise, European Standards and Guidelines (ESG) çerçevesinde benzer iç kalite güvencesi süreçleri yürütülmekte; Birleşik Krallık'ta Higher Education Statistics Agency (HESA) ve Amerika Birleşik Devletleri'nde National Center for Education Statistics (NCES) gibi kuruluşlar yükseköğretim kurumlarının performans verilerini yayımlamaktadır.

Bu sistemlerin güvenilirliğini tehdit eden başlıca unsurlardan biri eksik veri problemidir. Eksik veriler, yükseköğretim performans göstergelerinin analitik olarak değerlendirilmesini zorlaştırmakta ve sonuçların geçerliliğini tehlikeye atabilmektedir. Little ve Rubin (2002) ile Enders (2010) tarafından yapılan çalışmalar, eksik veri yapısının (örneğin MCAR, MAR, MNAR) doğru belirlenmesinin ve buna uygun analiz yöntemlerinin seçilmesinin istatistiksel sonuçlar üzerindeki etkisine dikkat çekmektedir. Özellikle eksik verilerin yanlış sınıflandırılması veya uygun olmayan yöntemlerle tamamlanması, kalite değerlendirme süreçlerinde sistematik hatalara neden olabilmektedir (Allison, 2002).

Bu nedenle, eksik verilerin türüne göre uygun yöntemlerle ele alınması ve bu sürecin kalite güvence sistemlerine entegre edilmesi gerektiği vurgulanmaktadır (García-Aracil & Palomares-Montero, 2013). Eksik veri yönetimi, yalnızca analiz düzeyinde değil, aynı zamanda kurumsal planlama ve karar alma süreçlerinde de dikkate alınması gereken kritik bir veri kalite unsurudur.

Yöntem

Bu bölümde, araştırmanın yöntemsel çerçevesi açıklanmaktadır. Araştırmanın temel odak noktalarından biri olan eksik veri sorunu, literatürdeki teorik çerçeveye birlikte ele alınmış, eksik veri mekanizmalarının yapısı nedensel diyagramlarla örneklendirilmiştir. Ayrıca, eksik veri türlerinin belirlenmesine yönelik olarak uygulanan Little's MCAR testi ile analiz sürecinin yönlendirilme biçimi açıklanmış ve panel veri analizlerinde eksik verilerle nasıl başa çıkılabileceğine dair yöntemsel önerilerde bulunulmuştur.

Araştırma Modeli

Bu çalışma, nicel araştırma yöntemleri kapsamında tasarlanmış olup, herhangi bir değişkenler arası neden-sonuç ilişkisi test edilmeden, mevcut durumun sistematik biçimde tanımlandığı betimsel tarama modeli çerçevesinde yürütülmüştür. Araştırmanın temel amacı, yükseköğretim kurumlarının kalite göstergelerine ilişkin mevcut durumu geniş kapsamlı bir veri seti üzerinden değerlendirmek ve özellikle eksik veri yapısını incelemektir.

Bu doğrultuda, çalışmada bütünlük kavramı temel bir ilke olarak ele alınmıştır. Bütünlük, veri üzerinde gerçekleştirilecek analizlerin geçerliliğini etkileyen ve verinin genişlik, kapsama ve doğruluk derecesini ifade

eden bir kavramdır (Batini & Scannapieca, 2016). Özellikle istatistiksel bütünlük, her bir veri kaydının tam olması, yani eksik veri içermemesi durumunu tanımlar (Karr ve ark., 2006). Bu çalışmada, veri bütünlüğünü bozan en önemli unsur olarak ele alınan eksik veriler, analizlerin güvenilirliğini etkileyebilecek potansiyel bir problem olarak değerlendirilmiştir.

Eksik Veri Tanımı ve Önemi

Veri bütünlüğü çerçevesinde değerlendirilen eksik veri sorunu, analizlerin güvenilirliğini etkileyen önemli bir unsurdur. Eksik veri; gözlemlenmemiş ancak analiz açısından anlamlı olabilecek değerlere karşılık gelir (Little & Rubin, 2020). Eksik veri, yalnızca bilgi kaybı yaratmakla kalmaz, aynı zamanda hatalı tahminler, geçersiz sonuçlar ve genelleme problemlerine neden olabilir (Allison, 2002; Rubin, 1978). Bu nedenle çalışmada, veri setindeki eksik verilerin sistematik bir yapıya sahip olup olmadığını belirlemek amacıyla eksik veri yapısı analiz edilmiştir. Özellikle eksik verilerin rastgele olup olmadığı, yani MCAR, MAR veya MNAR yapısına sahip olup olmadığı tartışılmış ve yönetime bu doğrultuda yön verilmiştir.

Eksik verilerin tanımlanması için literatürde en yaygın olarak kullanılan sistem, Rubin tarafından geliştirilmiştir (Rubin, 1976, s. 582). Bu çalışmada kullanılan panel veri setinde eksik gözlemlerin doğası, eksiklik mekanizmaları açısından Rubin'in üçlü sınıflandırması çerçevesinde ele alınmıştır. Rubin, eksik veri tiplerini tamamen rastgele eksik veri (MCAR - Missing Completely At Random), rastgele eksik veri (MAR - Missing At Random) ve rastgele olmayan eksik veri (MNAR - Missing Not At Random) olmak üzere üç kategoride ele almıştır (Little & Rubin, 2014, s. 21).

■ MCAR Varsayımı:

Bu varsayım, bir değişkende eksik veri olma olasılığının, bu değişkenin kendi değeri ya da veri setindeki diğer değişkenlerle ilişkisi olmadığı durumlarda meydana geldiği eksiklik türünü ifade etmektedir.

■ MAR Varsayımı:

Bu varsayım, bir değişkende eksik veri olma olasılığının, analizdeki diğer değişkenler kontrol altına alındığında, bu değişkenin kendi değeri ile ilişkisiz olduğu durumu ifade etmektedir.

■ MNAR Varsayımı:

Göz ardı edilemeyen eksik veri mekanizmasıdır. Bu varsayım, bir değişkende eksik veri olma olasılığının değişkenin kendi değerine bağlı olduğu durumu ifade etmektedir. Bu tür eksik verilerin bulunduğu veri setleriyle yapılan model geliştirme ve tahminler genellikle yanlış çalışabilmektedir (Leke & Marwala, 2019, s. 18).

MCAR, MAR'a kıyasla daha kısıtlayıcı bir eksik veri düzenine işaret etmektedir. MCAR durumunda eksik veri olasılığı, herhangi bir değişkenle ilişkilendirilmemektedir; ancak MAR durumunda eksik veri olasılığının diğer değişkenlere bağlı olduğu görülmektedir. MAR, eksik verilerin ihmal edilebilir

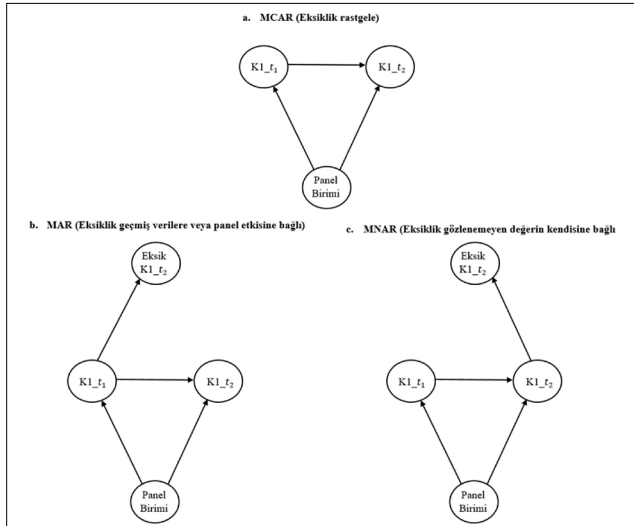
olduğu varsayımıyla birlikte literatürde çeşitli analizlerde sıkça kullanılmıştır. Rubin (1976), örneklem tabanlı çözümleme yöntemlerinin MCAR durumunda geçerli olabileceğini göstermiştir.

Uygulamalı panel veri araştırmalarında, genellikle bazı değerlerin eksik olduğu ve bu durumun dengesiz panellere yol açtığı bilinmektedir. Eksik gözlemlerin sebepleri çeşitlidir; örneğin, bazı veriler yıllık bazda elde edilirken, diğerleri çeyreklik bazda toplanabilmekte ve bu durum, yalnızca yıllık gözlemlerin mevcut olduğu dönemlerde çeyrek tarihlerinde eksik değerlere neden olabilmektedir. Bu çalışmada üniversitelere ait göstergelerin incelenmesinden hareketle; üniversitelerin eksik verileri eğer ki veri setinde belirli bir düzen içinde ortaya çıkıyorsa (örneğin, belirli bir üniversitenin belirli bir dönemde verilerin eksik olması gibi), bu durum panel veri analizlerinde sistematik bir hataya neden olabilecektir.

Not: K1, üniversitelerin stratejik planlarında yer alan eğitim ve öğretim faaliyetlerine ilişkin hedefleri gerçekleştirme yüzdesini ifade etmektedir. Panel veri setinde K1 değişkeni yıllara göre üniversiteler bazında gözlemlenmiş olup, bazı dönemlerde eksik gözlemler bulunmaktadır. Diyagramda K1_t1 geçmiş yılı, K1_t2 ise mevcut yılı temsil etmektedir.

Şekil 1

Panel Veride Eksiklik Mekanizmalarının Nedensel Diyagramlarla Gösterimi



Şekil 1'deki diyagram, sırasıyla MCAR, MAR ve MNAR mekanizmalarının nedensel ilişkiler bağlamında nasıl şekillendiğini göstermektedir. **Şekil a (MCAR):** Eksiklik, geçmiş veya mevcut K1 değerinden bağımsızdır. Eksiklikler tamamen rastgele oluşmaktadır. Bu durumda, eksik gözlemlerin analize dâhil edilmesi sonucu yanlılığa neden olmaz. Ancak uygulamada bu mekanizma oldukça nadirdir. **Şekil b (MAR):** Eksiklik, geçmiş döneme ait K1 değerine ve/veya panel birime (üniversite) bağlıdır. Yani, belirli geçmiş performans düzeylerine sahip

üniversitelerde eksik veri olma olasılığı daha yüksektir. Bu durumda eksiklik mekanizması, gözlemlenebilir değişkenlerle modelleme yoluyla yönetilebilmektedir. **Şekil c (MNAR):** Eksiklik, gözlemlenemeyen mevcut dönem K1 değeri ile doğrudan ilişkilidir. Örneğin, K1 değeri çok düşük olduğu için üniversite ilgili veriyi dikkate almamış olabilir. Bu durumda eksik veri mekanizması gözlemlenemeyen bir değere bağlı olduğu için klasik tahmin teknikleri ciddi yanlılıklar yaratabilir. Bu tür eksiklikler için model temelli yaklaşımlar kullanılabilir.

Eksik veriler, üniversitelerin performansını ve diğer önemli değişkenleri değerlendiren panel veri araştırmalarında bütünlük ve kapsam sorunlarına yol açabilecektir. Eğer eksiklikler rastgele değilse ve belirli bir tip üniversitelerde daha sık görülüyorsa, bu durum genellemelerin yapılmasını zorlaştıracak bir unsura dönüşecektir. Bu noktada, eksik verilerle başa çıkma stratejileri (imputation, veri toplama süreçlerini geliştirme vb.) eksik verilerin etkilerini azaltabilecektir. Eğer eksik veriler rastgele değilse ve bir sistematığı varsa, eksik verileri olduğu şekilde ele almak analizde daha güvenilir sonuçlar elde edilmesine yardımcı olacaktır. Ancak eksik verilerin sistematik bir düzeni yok ise eksik verileri tamamlamak daha uygun bir yaklaşım olacaktır. Bu çalışmada, eksik veri probleminin varlığı ve türü incelenerek analiz süreçlerinde nasıl ele alınması gerektiği değerlendirilmektedir. Panel veri setinde eksik verinin sistematik olup olmadığını test etmek için Little's MCAR testi uygulanmıştır. Yapının değerlendirilmesi için kullanılan Little's MCAR testi, eksik verilerin tamamen rastgele olup olmadığını belirlemeye yönelik bir yöntemdir. Testin temel hipotezleri şunlardır:

H_0 : Eksik veriler tamamen rastgele (MCAR) dağılmıştır.
 H_a : Eksik veriler tamamen rastgele değildir; yani, MCAR değildir.

Eğer elde edilen *p*-değeri belirlenen anlamlılık düzeyinden küçükse (genellikle 0.05), reddedilmekte ve eksik verilerin MCAR olmadığı anlaşılmaktadır. Little's MCAR test sonuçlarına göre; eğer eksik verilerin rastgele olmadığı (MCAR dışı) tespit edilirse, bu eksikliklerin belirli üniversite türlerinde veya belirli yıllarda daha sık görüldüğü anlamına gelmektedir. Bu durumda eksik verilerin olduğu şekilde modele dâhil edilmesi daha uygun olacaktır. Eğer eksik verilerin rastgele olduğu sonucuna ulaşırsa, eksik veriler çoklu atama yöntemi gibi tekniklerle tamamlanarak analizlere devam edilmektedir. Bu çalışma kapsamında oluşturulan panel veri setinde kullanılan değişkenler yıllara ve üniversitelere göre dengesiz şekilde eksiklik içermektedir. Bu nedenle, veri setindeki eksiklik yapısı değerlendirilmiş ve analiz süreciyle ilgili eksik verilerin doğasına uygun olarak önerilerde bulunulmuştur.



Bulgular

Bu bölümde, çalışmada oluşturulan panel veri setine ilişkin betimsel bilgiler ve eksik veri yapısına yönelik yapılan analizlerin bulguları sunulmaktadır. İlk olarak, panel veri setinin kapsamı ve veri kaynakları açıklanmakta; ardından, eksik veri sorununun yapısını anlamaya yönelik olarak uygulanan Little's MCAR testi sonuçları değerlendirilmektedir. Elde edilen bulgular eksik verilerin tamamen rastgele dağıldığını (MCAR) göstermekte ve bu doğrultuda eksik veriyle başa çıkma stratejileri kuramsal düzeyde tartışılmaktadır.

Veri Seti ve Kapsam

Çalışma kapsamında; Yükseköğretim Kalite Kurulu resmi internet sitesinde paylaşılan YÖKAK Gösterge Değerlendirme Raporu veri tabanından 208 üniversitenin (82'si vakıf ve 126'sı devlet üniversitesi olmak üzere) 2015-2022 zaman periyodunda ■ Tablo 1'de verilen göstergeler aracılığıyla panel veri seti oluşturulmuştur.

■ Tablo 1
Panel Veri Seti Değişkenleri

Gösterge Değerleri Raporu	Kapsamı	Kısaltma
2- Kalite Güvence Sistemi 1- Kurumun misyon, vizyon, stratejik hedefleri ve Kalite Komisyonunun faaliyetleri 1-2-3-4	Kurumun stratejik planında yer alan eğitim ve öğretim faaliyetlerine ilişkin hedefleri gerçekleştirme yüzdesi (% olarak)	K1
	Kurumun stratejik planında yer alan araştırma faaliyetlerine ilişkin hedefleri gerçekleştirme yüzdesi (% olarak)	K2
	Kurumun stratejik planında yer alan idari faaliyetlerine ilişkin hedefleri gerçekleştirme yüzdesi (% olarak)	K3
	Kurumun stratejik planında yer alan toplumsal hizmet faaliyetlerine ilişkin hedefleri gerçekleştirme yüzdesi (% olarak)	K4
1- Kuruma Ait Bilgiler 1- Birim/Öğrenci/Personel Sayıları, Fiziki Alanlar 16- Yabancı Uyruklu Öğrenci Sayısı	Yabancı Öğrenci Sayısı	Y1

Veri Kaynakları*

- YÖKSİS => istatistik.yok.gov.tr'de yayımlanan verileri
- Kurum => Kurumun kendisi tarafından doldurulan verileri
- ÖSYM => ÖSYM'den alınan verileri
- YÖKAK => Değişik (URAP, THE, WOS ya da kurul tarafından hesaplanan) kaynaklardan alınan verileri ifade etmektedir.
- Araştırma ve Geliştirme başlığı altında yer alan göstergeler Nisan ayı içerisinde, Q1 yayın sayısı ve Q1 yayın oranına ilişkin veriler ise Ağustos ayı içerisinde Scival ve Incites veri kümelerinden alınmıştır.
- ARWU, THE, QS, NTU göstergeleri sıralama aralığının alt sınırı olarak hazırlanmıştır.

Not:* veri kaynakları <https://yokak.gov.tr/raporlar/IndicatorValuesReport> adresindeki Veri Kaynakları bölümünden elde edilmiştir.

Eksik Veri Yapısının Tespiti

Panel veri setinde yer alan bazı değişkenlerde eksik gözlemler bulunduğu tespit edilmiştir. Bu eksikliklerin belirli bir örüntüye dayanıp dayanmadığını, yani sistematik olup olmadığını değerlendirmek amacıyla Little's MCAR (Missing Completely At Random) testi uygulanmıştır. Bu test, eksik verilerin tamamen rastgele (MCAR), rastgele (MAR) veya rastgele olmayan (MNAR) şekilde dağılıp dağılmadığını istatistiksel olarak değerlendirmeyi amaçlamaktadır.

Çalışmada oluşturulan panel veri setinde eksik verilerin dağılım yapısı, eksiklik türüne göre analiz edilmiş; eksik verilerle nasıl başa çıkılması gerektiğine yönelik yöntemsel tartışma yürütülmeden önce bu yapının doğru şekilde tespit edilmesi hedeflenmiştir. Bu doğrultuda uygulanan Little's MCAR testinin bulguları ■ Tablo 2'de sunulmuştur.

■ Tablo 2
Little's MCAR Testi Bulguları

Örneklem Hacmi	n=1664	n=1008	n=656
	Üniversiteler	Devlet Üniversiteleri	Vakıf Üniversiteleri
Değerler	14.6848	29.1604	57.8286
Serbestlik Derecesi	22	47	44
p-değeri	0.8754	0.9809	0.0789
Karar	reddedilemez	reddedilemez	reddedilemez

Little'in MCAR testi için hesaplanan *p-değerlerinin* istatistiksel olarak anlamlı olmaması yani 0,05'ten büyük olması nedeniyle eksik verilerde herhangi bir örüntü bulunmadığını, diğer bir ifadeyle tamamen rastgele dağıldığını ifade eden yokluk hipotezinin reddedilmediği bulgusuna ulaşılmaktadır. Buradan hareketle; veri setinde yer alan eksik verilerin tamamen rastgele bir dağılım gösterdiği sonucuna ulaşılmıştır. Başka bir deyişle, üniversitelerin kalite hedeflerini etkileyen faktörlerin eksik veri oluşumuyla sistematik bir şekilde bağlantılı olmadığı değerlendirilmektedir.

MCAR Yapısının Oluşmasına Sebep Olabilecek Nedenler

Üniversitelerde incelenen zaman periyodunda MCAR (Missing Completely At Random) yapısının oluşmasına neden olabilecek çeşitli faktörler bulunmaktadır. Analiz sonucunda elde edilen MCAR yapısı, bazı rastgele faktörlerle açıklanabilir:

Veri Girişi ve Anket Hataları: Anketlerin eksik doldurulması veya veri girişinde oluşan insan hataları (örn. veri giriş personelinin raporlamayı unutması).
Teknolojik Aksaklıklar: Yazılım arızaları, sunucu çökmeleri gibi sistemsel problemler.
Doğal Afetler veya Pandemi: Pandemi gibi olağanüstü durumlarda raporlama yapılamaması.
Personel veya Öğrenci Hareketliliği: Geçici görevlendirmeler veya dönemsel kayıt aksaklıkları.
Tesadüfi Kayıplar: Raporlama sürecinde rastgele unutulmalar.

Eksik Verilerle Başa Çıkma Stratejileri Üzerine Değerlendirme

Veri eksikliklerinin tamamen rastgele olduğunun (MCAR) belirlenmesi, silme ya da basit atama yöntemlerinin istatistiksel olarak kullanılabilir olduğunu göstermektedir. Bu doğrultuda, ortalama değer atama yöntemi, MCAR veriler için önerilen geleneksel bir yöntem olup, analizlerin güvenilirliğini olumsuz yönde etkilemeyeceği belirtilmektedir (Kelly, 2011). Ancak eksik veri oranı %5'in üzerine çıktığında, daha gelişmiş yöntemlerin (örneğin çoklu atama, EM algoritması) tercih edilmesi önerilmektedir (Osman ve ark., 2018).

Diğer taraftan, silme yöntemlerinin yalnızca MCAR koşulunda geçerli olduğu, aksi takdirde bilgi kaybına ve tahminlerde sapmalara yol açabileceği unutulmamalıdır. Bu nedenle, eksik verilerin sistematik bir düzene sahip olduğu (MAR/MNAR) durumlarda silme yöntemleri yerine atama veya modelleme teknikleri kullanılmalıdır.

Eksik verilerin tamamlanmaksızın analiz edilmesi durumunda, veri seti dengesiz panel veri yapısında olacaktır. Dengesiz panel veri setlerinin kullanımı, bazı avantajlar ve dezavantajlar içermektedir. Avantaj olarak, dengesiz panel yapısı gerçek hayattaki veri koşullarını daha iyi yansıtabilir; tüm birimlerin her dönemde tam veri sunamaması, özellikle sosyal bilimlerde sıkça karşılaşılan bir durumdur. Ayrıca, daha fazla birim ve yıl gözlemi içerebildiği için örneklem çeşitliliğini artırabilir. Bununla birlikte, bu yapı bazı dezavantajları da beraberinde getirir. Özellikle ekonometrik modelleme sürecinde dengesizlik, bazı yöntemlerin uygulanabilirliğini sınırlandırabilir; tahminlerin tutarlılığı ve anlamlılığı dengesizliğe bağlı olarak zayıflayabilir. Ayrıca eksik gözlemlere yapılan atamalar, parametre tahminlerinde sapmalara neden olabilir. Bu nedenle, dengesiz panel veri setleri kullanılabilir olsa da, eksik verilerle etkin şekilde başa çıkılmadığı takdirde analizlerin güvenilirliği zedelenabilir. Sonuç olarak, eksik veri sorunu içeren panel veri analizlerinde doğru modelleme tekniklerinin ve eksik veri yönetimi stratejilerinin özenle seçilmesi, analiz sonuçlarının geçerliliği ve güvenilirliği açısından kritik öneme sahiptir.

Sonuç ve Tartışma

Elde edilen bulgulara göre, eksik veriler tamamen rastgele (MCAR) yapıda dağılmaktadır. Bu bulgu, analizlerin genellenebilirliği açısından avantaj sağlamakla birlikte, veri kalitesine yönelik iyileştirme gerekliliğini ortadan kaldırmamaktadır. Zira Demir ve Parlak (2012) tarafından yapılan kapsamlı incelemede, Türkiye'deki eğitim araştırmalarının büyük çoğunluğunda eksik veri sorununun dikkate alınmadığı ve yalnızca %1,7'sinde eksik veriye özel bir analiz tekniği uygulandığı tespit edilmiştir. Bu durum, akademik çalışmaların yöntemsel sağlamlığını ve raporlama kalitesini olumsuz etkileyebilecek önemli bir eksiklik olarak değerlendirilmektedir. Benzer biçimde, Cox ve arkadaşları (2014) tarafından yükseköğretim alanında yapılan ampirik bir çalışmada da, eksik verilerin türüne göre uygulanacak stratejilerin analiz sonuçları üzerinde belirleyici olduğu vurgulanmıştır. Bu bağlamda, eksik veri mekanizmasının analiz öncesi test edilmesi, eksikliğin rastgele mi (MCAR), gözlemlenebilir değişkenlere bağlı mı (MAR), yoksa gözlemlenemeyen değerlere mi bağlı (MNAR) olduğunun anlaşılması, uygulanacak yöntemin doğruluğu açısından kritik önemdedir. Soğuksu (2024), eğitim araştırmalarında kullanılan IRTree (Item Response Tree) modelleriyle eksik veri mekanizmasının türünü belirlemenin mümkün olduğunu ve analizlerin güvenilirliğini artırdığını ortaya koymuştur. Yine benzer şekilde, Anahideh ve arkadaşları (2021), yükseköğretim verilerinde yapılan öngörücü analizlerde uygulanan atama (imputation) yöntemlerinin sadece doğruluk açısından değil, aynı zamanda analizlerin adalet ve eşitlik boyutları üzerinde de önemli etkileri olduğunu vurgulamıştır. Bu da eksik veriyle başa çıkma stratejilerinin hem istatistiksel hem de etik sorumluluklar taşıdığını göstermektedir. Bu bağlamda, eksik veri yapısının MCAR olması, klasik analiz yöntemlerinin uygulanabilirliğini desteklemekte ve örnekleme yanlılığı riskini azaltmaktadır. Ancak veri kalitesinin artırılması ve eksiklik oranlarının azaltılması yönünde önleyici adımlar atılması hâlâ önemini korumaktadır.

Günümüzde yükseköğretim kurumları, sürekli kaliteyi artırmaya ve sürdürülebilir bir eğitim ortamı oluşturmaya çalışmaktadır. Bu çabaların merkezinde kanıta dayalı arşivleme sistemleri ve bütünsel bilgi yönetim sistemleri yer almakta; bu sistemler, kurumların büyük ölçekli verileri merkezi olarak saklamasını, analiz etmesini ve karar destek sistemleriyle entegre biçimde kullanmasını sağlamaktadır. Böylece, akademik ve idari performans değerlendirmeleri daha etkin yapılabilmekte, stratejik planlama süreçlerine veriye dayalı katkı sunulabilmektedir. Ancak bu sistemlerin başarısı, büyük ölçüde veri bütünlüğüne ve eksiksizliğine bağlıdır. Eksik veriler; veri toplama sürecindeki aksaklıklar, farklı dönemsel raporlama standartları veya teknik sorunlar nedeniyle oluşabilmekte ve analizlerin doğruluğunu ciddi biçimde etkileyebilmektedir. Bu durum, üniversitelerin hem performans değerlendirme süreçlerini hem de uluslararası sıralamalardaki konumlarını zayıflatabilecek potansiyele sahiptir.



Bu çalışma kapsamında, Yükseköğretim Kalite Kurulu (YÖKAK) tarafından yayımlanan Gösterge Değerlendirme Raporları incelenmiş; üniversitelere ait göstergelerdeki eksiklikler MCAR, MAR ve MNAR mekanizmaları temelinde sınıflandırılmıştır. Panel veri analizinden önce yapılan Little's MCAR testi, eksik verilerin rastgele bir yapıda dağıldığını (MCAR) ortaya koymuştur. Bu bulgu, eksik verilerin belirli bir sistematik örüntüye bağlı olmadığını ve klasik panel veri analizlerinin güvenle uygulanabileceğini göstermektedir.

Bununla birlikte, veri eksikliklerinin sistematik hâle gelmemesi adına veri giriş standartlarının üniversiteler arasında eşgüdümlü hâle getirilmesi, veri toplama süreçlerinin dijital altyapılarla desteklenmesi ve bütünlük bilgisi yönetim sistemlerinin daha etkin kullanılması önerilmektedir. Ayrıca, eksik veriyle başa çıkmak için sadece teknik çözümler değil, kurumsal düzeyde farkındalık ve sürdürülebilir veri politikaları da geliştirilmelidir.

Sonuç olarak, bu çalışma; eksik verilerin rastgeleliğini doğrulamakla kalmayıp, veri kalitesinin yükseköğretimde stratejik karar alma süreçlerine olan etkisini de ortaya koymaktadır. Bu bağlamda, hem teknik hem de yönetsel düzeyde eksik veriye karşı önlem alınması, yükseköğretim sisteminde kalite güvence mekanizmalarının sürdürülebilirliği açısından kritik önem taşımaktadır.

Extended Abstract

Complementary Map of Quality in Higher Education: Missing Data and Evaluation Strategies

Quality assurance in higher education systems is a fundamental element for sustainable academic and administrative success. The Higher Education Quality Board (YÖKAK), the main institutional structure that directs this process in Türkiye, monitors and evaluates quality through performance indicators. Indicator Assessment Reports, which include these indicators, are of critical importance for the strategic planning and institutional improvement processes of universities. However, the lack of data in these reports is a problem that can undermine the validity and reliability of quality assessment. This study aims to analyze the nature of missing data in the indicators in the YÖKAK database and to present methodological and strategic suggestions to increase data quality.

Literature Review

Missing data analysis is an important research topic in social sciences. The missing data classification proposed by Rubin (1976) – Completely Missing Data at Random (MCAR), Missing Data at Random (MAR) and Missing Data Not at Random (MNAR) – constitute the basic framework in this field. It is stated that in educational research, the problem of missing data is not given enough attention and this situation negatively affects the validity of the analyses (Demir & Parlak, 2012; Cox et al., 2014). At the same time, it is emphasized that appropriate methods should be selected according to the missing data structure and that incorrect classifications may lead to systematic errors (Enders, 2010; Allison, 2002). In this context, it is vital to correctly define the missing data structure in order to reliably evaluate higher education performance data.

Method

The research is based on a panel dataset created from YÖKAK Indicator Evaluation Reports of 208 universities in Türkiye between 2015 and 2022. Missing data were analyzed specifically for the number of foreign students, one of the indicators in the dataset. Little's MCAR test was applied to determine the missing data structure and statistically evaluated whether the missing data were random or not. The study was carried out within the framework of a descriptive screening model with a quantitative approach; causal diagrams regarding the missing data structure were also used in the analyses.

Findings

As a result of the analyses, it was determined that the distribution of missing data was completely random (MCAR). The results of the MCAR test, which was applied separately to the state and foundation university groups,

similarly show that the missing data are not based on a systematic structure. This situation supports the validity of classical methods such as deletion or averaging, as the risk of sampling bias in the analyses is low due to missing data. However, factors such as data entry errors, technological disruptions or extraordinary circumstances can cause missing data to be distributed randomly.

Discussion

The fact that missing data are in the MCAR structure is a situation that supports statistical integrity in the analyses. However, this finding does not mean that data quality problems can be ignored. Strategies for dealing with missing data should be supported not only by technical tools but also by sustainable data policies at the institutional level. In addition, it is frequently emphasized in the literature that classical methods are inadequate and model-based approaches may be required in cases where missing data are not random (MAR or MNAR) (Little & Rubin, 2020; Soğuksu, 2024). Therefore, testing the missing data structure before each analysis is a critical step.

Conclusion and Recommendations

This study analyzed whether the data belonging to the indicators used in quality assessments in higher education were missing and revealed that the missing data were completely random (MCAR). This finding shows that classical panel data analyses can be applied safely. However, in order to minimize data missingness, it is recommended that data entry standards be harmonized among universities, digital data collection systems be disseminated, and integrated information management systems be used more effectively. In addition, dealing with missing data is not only a technical necessity but also an ethical and governance responsibility. In future studies, it is recommended that the structure of missing data be determined as well as the impact of this structure on data analyses be evaluated.



Kaynakça

- Allison, P. D. (2002). *Missing data*. Sage Publications.
- Anahideh, H., Ye, X., Vate-U-Lan, P., & Ouyang, Y. (2021). Auditing the imputation effect on fairness of predictive analytics in higher education. *arXiv preprint*, arXiv:2109.07908. <https://doi.org/10.48550/arXiv.2109.07908>
- Batini, C., & Scannapieca, M. (2016). *Data and information quality: Dimensions, principles and techniques*. Springer International Publishing.
- Baydemir, M., & Bolat, M. A. (2024). Uluslararası öğrencilerin Türk ekonomisine katkıları: İstanbul'da kendi hesabına okuyan öğrenciler örneği. *Turkish Journal of Diaspora Studies*, 4(2), 203–220. <https://doi.org/10.52241/TJDS.2024.0078>
- Cox, B. E., McIntosh, K., Reason, R. D., & Terenzini, P. T. (2014). Working with missing data in higher education research: A primer and real-world example. *The Review of Higher Education*, 37(3), 377–402. <https://doi.org/10.1353/rhe.2014.0026>
- de Wit, H. (2020). Internationalization of higher education: The need for a more ethical and qualitative approach. *Journal of International Students*, 10(1), i–iv. <https://doi.org/10.32674/jis.v10i1.1893>
- Demir, E., & Parlak, B. (2012). Türkiye'de eğitim araştırmalarında kayıp veri sorunu. *Journal of Measurement and Evaluation in Education and Psychology*, 3(1), 230–241.
- Enders, C. K. (2010). *Applied missing data analysis*. Guilford Press.
- European Association for Quality Assurance in Higher Education. (2015). *Standards and guidelines for quality assurance in the European Higher Education Area (ESG)*. ENQA. https://www.enqa.eu/wp-content/uploads/2015/11/ESG_2015.pdf
- García-Aracil, A., & Palomares-Montero, D. (2013). Examining performance indicators in higher education: Efficiency and quality in teaching and research. *Journal of Higher Education Policy and Management*, 35(6), 610–623. <https://doi.org/10.1080/1360080X.2013.844666>
- Gartner. (2017). *How to create a business case for data quality improvement*. <https://www.gartner.com/smarterwithgartner/how-to-create-a-business-case-for-data-quality-improvement>
- Gümüş, N., & Onurlubaş, E. (2019). Uluslararası öğrencilerin tüketim harcamalarının şehir ekonomisine katkısının incelenmesi: Kastamonu Üniversitesi örneği. *Düzce Üniversitesi Sosyal Bilimler Dergisi*, 9(2), 174–192.
- Harvey, L., & Green, D. (1993). Defining quality. *Assessment & Evaluation in Higher Education*, 18(1), 9–34. <https://doi.org/10.1080/0260293930180102>
- Hazelkorn, E. (2015). *Rankings and the reshaping of higher education: The battle for world-class excellence*. Palgrave Macmillan.
- Jongbloed, B., Enders, J., & Salerno, C. (2008). Higher education and its communities: Interconnections, interdependencies and a research agenda. *Higher Education*, 56(3), 303–324. <https://doi.org/10.1007/s10734-008-9128-2>
- Karr, A. F., Sanil, A. P., & Banks, D. L. (2006). Data quality: A statistical perspective. *Statistical Methodology*, 3(2), 137–173. <https://doi.org/10.1016/j.stamet.2005.08.006>
- Kelly, S. (2011). Do homes that are more energy efficient consume less energy?: A structural equation model of the English residential sector. *Energy*, 36(9), 5610–5620. <https://doi.org/10.1016/j.energy.2011.07.004>
- Knight, J. (2004). Internationalization remodeled: Definition, approaches, and rationales. *Journal of Studies in International Education*, 8(1), 5–31. <https://doi.org/10.1177/1028315303260832>
- Leke, C. A., & Marwala, T. (2019). *Deep learning and missing data in engineering systems*. Studies in Big Data (Vol. 48) içinde. Springer Nature. https://doi.org/10.1007/978-3-030-04152-7_2
- Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data* (2nd ed.). John Wiley & Sons.
- Little, R. J. A., & Rubin, D. B. (2014). *Statistical analysis with missing data* (3rd ed.). John Wiley & Sons.
- Little, R. J. A., & Rubin, D. B. (2020). *Statistical analysis with missing data* (3rd ed.). John Wiley & Sons.
- Osman, M. S., Abu-Mahfouz, A. M., & Page, P. R. (2018). A survey on data imputation techniques: Water distribution system as a use case. *IEEE Access*, 6, 63279–63291. <https://doi.org/10.1109/ACCESS.2018.2876558>
- Rubin, D. B. (1976). Inference and missing data. *Biometrika*, 63(3), 581–592. <https://doi.org/10.1093/biomet/63.3.581>
- Sadırlı, S. (2023). *Yükseköğretim kalite güvencesi modeli: Yeni kurulan bir devlet üniversitesinde uygulamalar* [Yüksek lisans tezi, Tarsus Üniversitesi].
- Soğuksu, Y. B. (2024). The efficacy of the IRTree framework for detecting missing data mechanisms in educational assessments. *Journal of Measurement and Evaluation in Education and Psychology*, 15(3), 209–220. <https://doi.org/10.21031/epod.1514741>
- Yükseköğretim Kalite Kurulu (YÖKAK). (2024). *Gösterge Değerleri Raporu*. <https://yokak.gov.tr/raporlar/IndicatorValuesReport>

Bu makale Creative Commons Attribution-NonCommercial-NoDerivs 4.0 Unported (CC BY-NC-ND 4.0) Lisansı standartlarında; kaynak olarak gösterilmesi koşuluyla, ticari kullanım amacı ve içerik değişikliği dışında kalan tüm kullanım (çevrimiçi bağlantı verme, kopyalama, baskı alma, herhangi bir fiziksel ortamda çoğaltma ve dağıtma vb.) haklarıyla açık erişim olarak yayımlanmaktadır. / This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 Unported (CC BY-NC-ND 4.0) License, which permits non-commercial reuse, distribution and reproduction in any medium, without any changing, provided the original work is properly cited.

Yayıncı Notu: Yayıncı kuruluş olarak Türkiye Bilimler Akademisi (TÜBA) bu makalede ortaya konan görüşlere katılmak zorunda değildir; olası ticari ürün, marka ya da kuruluşlarla ilgili ifadelerin içerikte bulunması yayıncının onayladığı ve güvence verdiği anlamına gelmez. Yayıncının bilimsel ve yasal sorumlulukları yazar(lar)ına aittir. TÜBA, yayınlanan haritalar ve yazarların kurumsal bağlantıları ile ilgili yargı yetkisine ilişkin iddialar konusunda tarafsızdır. / *Publisher's Note: The content of this publication does not necessarily reflect the views or policies of the publisher; nor does any mention of trade names, commercial products, or organizations imply endorsement by Turkish Academy of Sciences (TÜBA). Scientific and legal responsibilities of published manuscript belong to their author(s). TÜBA remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.*

