Int. J. Adv. Eng. Pure Sci. 2025, 37(UYIK 2024 Special Issue): <21-29>
DOI: 10.7240/jeps.1506790

RESEARCH ARTICLE / ARAŞTIRMA MAKALESİ

Türkiye, Hindistan ve Brezilya'daki Petrol Kiralarının Otoregresif Kesirli Bütünleşik Hareketli Ortalama (ARFIMA) Modeli ile Belirlenmesi

Determination of Oil Rents in Turkey, India and Brazil with Autoregressive Fractionally Integrated Moving Average (ARFIMA) Model

Semanur SARIÇAM [1], **Barış AŞIKGİL** [2]

[1]*Marmara Üniversitesi, Fen Fakültesi, İstatistik Bölümü, İstanbul, Türkiye*
[2]*Mimar Sinan Güzel Sanatlar Üniversitesi, Fen-Edebiyat Fakültesi, İstatistik Bölümü, İstanbul, Türkiye*

Öz
Petrol kiraları, üretim, ithalat ve tüketim vergileri üzerinden elde edilen gelir olduğu için enerji politikalarının bir parçası olarak değerlendirilir. Türkiye, Hindistan ve Brezilya gibi ülkelerin enerji politikaları, petrol ürünlerine yüksek vergiler uygulanmasından dolayı bu ülkelerin gelir elde etme yöntemlerinin bir parçasıdır. Petrol kiralarının ekonomide önemli rolü göz önüne alındığında, bu çalışma, bu ülkelerin 1970-2016 yıllarını kapsayan petrol kira verilerini kullanılarak uzun hafıza modellerinden biri olan otoregresif kesirli bütünleşik hareketli ortalama (ARFIMA) modeliyle incelemeyi amaçlamaktır. İlk aşamada, Türkiye, Hindistan ve Brezilya'nın petrol kira serilerinin durağan olup olmadığını belirlemek için klasik birim kök testleri uygulanmıştır. Sonrasında, serilerin uzun hafızaya sahip olup olmadığını değerlendirmek amacıyla Hurst'ün R/S istatistiği kullanılmıştır. İkinci aşamada ise p ve q değerleri 2'den küçük veya eşit olacak şekilde olası tüm otoregresif hareketli bütünleşik ortalama (ARIMA) ve ARFIMA modelleri test edilmiştir. En iyi model belirlenebilmesi için Akaike Bilgi Kriteri (AIC) ve Bayesci Bilgi Kriteri (BIC) iki temel bilgi kriterlerinden yararlanılmıştır. En uygun model belirlendikten sonra, bu model aracılığıyla geleceğe yönelik tahminler yapılmıştır.
Anahtar kelimeler: Petrol kiraları, Uzun hafıza modelleri, ARFIMA modeli, ARIMA modeli.

Abstract
Oil rents are considered as part of energy policies as they are revenues from taxes on production, imports and consumption. The energy policies of countries such as Turkey, India and Brazil are part of the revenue generation methods of these countries due to high taxes on petroleum products. Considering the important role of oil rents in the economy, this study aims to analyse the oil rent data of these countries covering the years 1970-2016 with the autoregressive fractionally integrated moving average (ARFIMA) model, which is one of the long memory models. In the first stage, classical unit root tests are applied to determine whether the oil rent series of Turkey, India and Brazil are stationary. Then, Hurst's R/S statistic is used to assess whether the series have long memory. In the second stage, all possible autoregressive moving average (ARIMA) and ARFIMA models are tested for p and q values less than or equal to 2 In order to determine the best model, two basic information criteria Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) are used. After the most appropriate model is determined, forecasts are made by using this model.
Key words: Oil rents, Long memory models, ARFIMA model, ARIMA model.

I.GİRİŞ

Zamana bağlı olarak kaydedilen finansal ve makroekonomik serilerin modellenmesi ve geleceğe yönelik tahminler yapılabilmesi için ekonometriciler ve istatistikçiler genellikle ARIMA modellerini kullanırlar. ARIMA modelleri, bugünkü gözlemler kullanılarak yarının tahmin edilmesine olanak sağlayan kısa hafızalı modellerdir. Ancak, serinin daha eski gözlemlerinin bugünü etkilemesi sonucunda uzun hafızalı modeller geliştirilmiştir. Bu, serinin otokorelasyon fonksiyonunun hiperbolik olarak azalmasıyla karakterize edilir.

Mevsimsel olmayan Box-Jenkins modelleri genellikle ARIMA(p, d, q) olarak ifade edilir. Burada d parametresi, serinin durağan olup olmadığını belirler; eğer seri durağansa, $d = 0$ olarak tanımlanır. Çoğu zamana bağlı seriler durağan olmadığından, maksimum iki defaya kadar fark alma işlemi yapılır. Serinin birinci farkı durağansa $d = 1$, ikinci farkı durağansa $d = 2$ olur. Ancak, 0 ile 1 arasındaki kesirli değerlerin modele dahil edilmemesi bilgi kaybına yol açabilir. Bu kesirli değerlerin $[-0.5, 0.5]$ aralığında olması ARFIMA(p, d, q) modelini oluşturur. Fark parametresi d, serinin uzun hafızalı mı yoksa kısa hafızalı mı olduğunu belirlemeye yardımcı olur.

Sorumlu Yazar: SEMANUR SARIÇAM, **Tel:** 05465660376, **E-posta:** semanur.saricam@marmara.edu.tr

Gönderilme: 28.06.2024, **Düzenleme:** 30.07.2024, **Kabul:** 03.08.2024

Finansal seriler gibi zaman serisi verileri üzerine uzun hafıza kavramının varlığını araştıran çok sayıda çalışma mevcuttur. Kutlar ve Turgut, ARFIMA modelleri ile tahmin ve öngörülebilirlik üzerine Türkiye'deki başlıca ekonomik serilere odaklanmışlardır [1]. Barışık ve Çevik, işsizlik verisi üzerine Geweke ve Porter-Hudak (GPH) ve modifiye edilmiş log periodogram (MLP) metotlarını kullanarak ARFIMA modellerini incelemişlerdir [2]. Ayrıca, 2001 kriz sonrası 2003-2007 yıllarını kapsayan Türk bankacılık sektörüne ait hisse senetlerini Çevik ve Erdoğan, güçlü hafıza modelleri ile analiz etmişlerdir [3]. Portföy optimizasyonu üzerine Pekkaya, ARFIMA ve kesirli bütünleşik genelleştirilmiş otoregresif koşullu değişen varyans (FIGARCH) modellerini kullanarak inceleme yapmıştır [4]. Çevik, İstanbul Menkul Kıymetler Borsası'nda (İMKB) işlem gören 10 sektör endeksini kullanarak uzun hafıza modellerinin etkin piyasa hipotezinde kullanılmasının önemini vurgulamıştır [5]. Karia, Bujang ve Ahmad, Malezya'daki ham hurma yağı fiyatlarını kullanarak tahmin hesaplamasında ARFIMA modelin daha iyi sonuçlar gösterdiğini kanıtlamışlardır [6]. Pekkaya ve Albayrak, IMKB-30 Endeks verileri üzerinde getiri öngörüleri için ARFIMA modelini ve varyans öngürüleri için de FIGARCH modelini kullanarak tahmin yapmışlardır [7]. Yılmaz, Brezilya, Rusya, Hindistan ve Çin'deki hisse senedi verilerini çalışmasına dahil ederek uzun hafıza davranışının sergilenip sergilenmediğini ve volatilitenin olup olmadığını araştırmak için bir inceleme yapmıştır [8]. Boateng, ve ark., Gana'daki tüketici fiyat endeksi (CPI) için uzun hafızanın sergilendiğini, yarı parametrik ve parametrik yöntemler kullanılarak ARFIMA modelinin en iyisi olduğunu keşfetmişlerdir [9]. Omeraka ve diğ., ARIMA ve ARFIMA modellerini kullanıp Nijerya'daki ticari bankaların likidite oran verisi üzerinde en iyi modelin ARFIMA model olduğunu ileri sürmüşlerdir [10]. Fahreddinoğlu, Azerbaycan'daki Tüketici Fiyat Endeksi (TÜFE) serisini kullanarak ARIMA ve ARFIMA modellerinin performanslarını kıyaslamıştır ve en iyi modelin ARFIMA model olduğuna karar vermiştir [11].

Bu çalışmanın sonraki bölümleri aşağıdaki şekilde düzenlenmiştir. Bölüm 2'de, uzun hafızalı model hakkında temel bilgiler sunulmuştur. Bölüm 3'te, petrol kiraları üç ülke üzerinde mevcut yöntemlerle kıyaslanarak değerlendirilmiştir. Son bölüm makaleyi sonlandırmakta ve sunulan modelinin nasıl daha da genişletilebileceğine dair bazı fikirler sunmaktadır.

II.METODOLOJİ

2.1.Uzun Hafıza Modelleri

Gerçek hayatta, gözetim videoları, gerçek zamanlı sensörler, hisse senedi fiyatları ve astrofizik gözlemleri gibi zaman serisi verileriyle sıkça karşılaşırız. Zaman serisi verilerini analiz ederken, bu verilerdeki çeşitli noktalardaki bağımlılıklar dikkate alınması gereken en önemli unsurlardan biridir. Zaman serisi verileri arasında herhangi bir bağımlılık var mı? Eğer öyleyse, bu değerler birbirlerini etkilemek için zaman açısından ne kadar uzak olmalıdır? Bu sorulara cevap ararken, uzun hafıza (long memory) kavramı devreye girer.

Uzun hafıza ve kısa hafıza varlığını anlamanın çeşitli yolları vardır. Kısa hafıza varlığında, noktalar arasındaki zaman farkı arttıkça farklı zamanlardaki değerler arasında bağımlılık hızla azalır. Bu durumda, otokorelasyon fonksiyonu üstel bir şekilde azalır veya belirli bir zaman gecikmesinden sonra noktalar bağımsız hale gelerek 0'a düşer. Diğer bir ifadeyle, eğer bugünün gözlemi ile yarını tahmin etmek istenirse, serinin ilk gecikmesinde otokorelasyon fonksiyonunda kısa hafızalı bir süreç bulunmalıdır. Uzun hafıza süreçlerinde ise otokorelasyon fonksiyonunun azalması daha yavaş olduğundan, yani bağımlılık daha güçlü olduğundan, serinin birkaç zaman önceki gözlemleri bugünü etkiler. Bu durum, sürecin uzun hafızalı süreç olduğunu gösterir [12].

Zaman serilerinin istatistiki çıkarımlarda öngörüler yapılabilmesi için serinin durağanlık varsayımı üzerine kurulması gerekmektedir. Diğer bir deyişle, ilk önce zaman serisinin birim kök içerip içermediğine bakılarak durağanlık varsayımı kontrol edilmelidir. Durağan olmayan serilerde varyans, zamanın bir fonksiyonu şeklinde hareket eder. Bu da, geçmişteki beklenmedik etkilerin seri tarafından kalıcı olarak taşınmasına neden olur. Bu yüzden, durağan olmayan serilerde öngörülerin güvenilirliği sorgulanacaktır [2]. Ancak, ne tam olarak durağan yani $I(0)$ ne de tam olarak durağan olmayan yani $I(1)$ serileri gibi olan ekonomik ve finansal zaman serileri, uzun hafıza sürecinin özelliğini sergilerler [1]. Uzun hafızaya sahip olan serilerin bütünleşik derecesi "d" parametresi $[-0.5,0.5]$ arasındaki bir kesirli sayı olarak modellenmektedir [13].

2.1.1Otoregresif kesirli bütünleşik hareketli ortalama modeli

Granger ve Joyeux 1980 yılında [14] ve Hosking 1981 yılında [15], uzun hafıza sürecinin belirlenmesinde kullanılan ARFIMA modelini ortaya atmışlardır.

y_t, bir zaman serisi olarak tanımlansın. ARFIMA(p, d, q) modeli,

$$\phi(L)(1-L)^d y_t = \mu + \theta(L)\varepsilon_t, t = 1,2,\ldots,T \quad \textbf{(1)}$$

formunda verilebilir. Burada, $\varepsilon_t \sim iid(0,\sigma^2)$, L gecikme operatörü, $\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p$, $\theta(L) = 1 - \theta_1 L - \theta_2 L^2 - \cdots - \theta_q L^q$ ve μ, zamanın herhangi bir deterministik fonksiyonu olabilir. Bu modelde, d değeri kesirli olduğu için kesirli bütünleşik (FI) model olarak adlandırılır. Eğer $d = 0$ ise y_t trend-durağan; $d = 1$ (veya daha büyük tamsayı) ise y_t fark-durağan; $d \in (-\frac{1}{2}, 0) \cup (0, \frac{1}{2})$ ise FI modeller söz konusudur.

ARFIMA model, ARIMA modelin bütünleşik (farklılaştırma) parametresi olan d'yi, tamsayı ile sınırlamaktan ziyade onun herhangi bir gerçek değeri almasına olanak tanımaktadır. d parametresi, Binom açılımından yararlanarak hareketli ortalama (MA) süreçlerinin sonsuz basamağı olarak aşağıdaki gibi ifade edilebilir:

$$(1-L)^d = 1 - dL + \frac{d(d-1)}{2!}L^2 - \frac{d(d-1)(d-2)}{3!}L^3 + \cdots$$

$$= \sum_{k=0}^{\infty} \frac{\Gamma(k-d)L^k}{\Gamma(-d)\Gamma(k+1)} \quad (2)$$

Burada, Gamma fonksiyonu,

$$\Gamma(g) = \int_0^{\infty} x^{g-1} e^{-x} dx \quad (3)$$

olarak tanımlanır. ARFIMA modelinde d parametresinin rolünü görmek için y_t'nin kovaryans fonksiyonunun bilinmesi gerekmektedir. Varyans γ_0, otokovaryans fonksiyonu γ_τ ve otokorelasyon fonksiyonu ρ_τ , sırasıyla

$$\gamma_0 = \frac{\sigma^2\Gamma(1-2d)}{\Gamma(1-d)\Gamma(1-d)} \quad (4)$$

$$\gamma_\tau = \frac{\sigma^2\Gamma(1-2d)\Gamma(\tau+d)}{\Gamma(d)\Gamma(1-d)\Gamma(\tau+1-d)}, \ \tau = \pm1, \pm2, \cdots \quad (5)$$

ve

$$\rho_\tau = \frac{\Gamma(1-d)\Gamma(\tau+d)}{\Gamma(d)\Gamma(\tau+1-d)} \quad (6)$$

biçimindedir.

Eğer $\phi(L)$ ve $\theta(L)$'nin kökleri birim çemberin dışında ve $d < |0.5|$ ise y_t süreci hem durağan hem de tersinir seri olarak tanımlanır. Fakat, $\tau \to \infty$, $\rho_\tau \propto \tau^{2d-1}$ olduğundan otokorelasyon fonksiyonları sonlu bir toplama sahip olmayacaktır. Başka bir ifadeyle, $0 < d < 0.5$ iken ARFIMA modellerinde uzun hafıza sürecinin varlığından bahsedilir. Yani, benzer otoregresif hareketli ortalama (ARMA) süreçlerine kıyasla bu süreçlerin otokorelasyon fonksiyonu daha yavaş bir bozulma sergilemektedir. $0.5 < d < 1.0$ iken γ_0, varyansı sonlu bir toplama sahip olmadığından y_t süreci durağan değildir. Böylece, Hosking'in formülü olan (6) denklemindeki otokorelasyon fonksiyonu hala 0'a doğru bir azalış sergileyecektir.

Süreçlerin hafızasının sonlu olduğu durumda bir şok verildiği zaman bu süreçlerin ortalama geri döndürme eğiliminde (mean-reverting) olduğu anlamına gelmektedir. $d > 1.0$ için, süreç ortalama geri dönme eğiliminde değil ve sürece bir şok uygulandığında sürecin başlangıç noktasından sapmasına sebep olmaktadır. Böylece, bir ARFIMA modelinin hafıza özelliği önemli bir şekilde d parametresine bağlıdır ve otokorelasyon fonksiyonları, $I(0)$ serilerine nispeten yavaş bir şekilde daha küçük oranda azalış sergiler [16].

2.2.Uzun Hafıza Modellerinin Tahmin Yöntemleri

Literatürde, iki aşamalı ve tek aşamalı yöntem olmak üzere uzun hafıza tahmin yöntemleri iki grup altında toplanmaktadır. Bu çalışmada iki aşamalı tahmin yöntemi kullanılmıştır [17].

Geweke ve Porter-Hudak (GPH) tarafından önerilen yarı-parametrik iki aşamalı bir yöntem olan GPH metodunun ilk aşamasında d bütünleşik parametresi tahmin edilmektedir [18]. İkinci aşamasında ise diğer otoregresif (AR) ve MA parametreleri tahmin edilmektedir. Uygulamada da oldukça sık tercih edilen bir yöntemdir [16]. GPH yönteminin temeli log-periodogram regresyonuna dayandığından dolayı bir regresyon yöntemi olarak ifade edilir.

Durağan bir serinin spektral yoğunluğu,

$$f(\lambda) = f_0(\lambda)[2\sin(\lambda/2)]^{-2d} \quad (7)$$

biçiminde tanımlanır. (7) eşitliğinin logaritması alınarak aşağıdaki denklem elde edilir:

$$logf(\lambda_j) = logf_0(0) - dlog\left[2\sin\frac{\lambda_j}{2}\right]^2 + log\left[\frac{f_0(\lambda_j)}{f_0(0)}\right] \quad (8)$$

Burada $\lambda_j = 2\pi j / n$, Fourier frekanslardaki spektral yoğunluktur. Başka bir ifadeyle, $I(\lambda_j)$ periodogramının logaritması aşağıdaki gibi ifade edilmektedir:

$$logI(\lambda_j) = log\left[\frac{I(\lambda_j)}{f(\lambda_j)}\right] + logf(\lambda_j) \quad (9)$$

(8) ve (9) eşitliklerinden yararlanarak,

$$logI(\lambda_j) = logf_0(0) - dlog\left[2\sin\frac{\lambda_j}{2}\right]^2 + log\left\{\frac{I(\lambda_j)[2\sin(\lambda/2)]^{2d}}{f_0(0)}\right\} \quad (10)$$

denklemi elde edilir. $y_j = \log I(\lambda_j)$, $\alpha = logf_0(0)$, $\beta = -d$, $x_j = log[2\sin(\lambda_j/2)]^2$ ve $\varepsilon_j = log\left\{\frac{I(\lambda_j)[2\sin(\lambda/2)]^{2d}}{f_0(0)}\right\}$ biçiminde olmak üzere, aşağıdaki regresyon denklemi elde edilebilir:

$$y_j = \alpha + \beta x_j + \varepsilon_j, \ j = 1, \cdots, m \quad (11)$$

Bu durumda, $f(\lambda_j) \sim f_0(0)[2\sin(\lambda_j/2)]^{-2d}$ ve $\varepsilon_j \sim log\left[\frac{I(\lambda_j)}{f(\lambda_j)}\right]$ olmaktadır. Böylece, uzun hafıza parametresi en küçük kareler tahmini,

$$\hat{d}_m = -\frac{\sum_{j=1}^{m}(x_j - \bar{x})(y_j - \bar{y})}{\sum_{j=1}^{m}(x_j - \bar{x})^2} \quad (12)$$

biçiminde verilir. Burada, $\bar{x} = \sum_{j=1}^{m} x_j/m$ ve $\bar{y} = \sum_{j=1}^{m} y_j/m$'dir [13].

2.3.Model Seçiminde Bilgi Kriterleri

Uygulamalarda parametrik olarak belirlenen modeller için seçim kriterleri yardımıyla en iyi model belirlemesi yapılmaktadır [19]. ARFIMA modellerinin seçimi için çeşitli uygulamalarda AIC, BIC, Schwarz bilgi kriteri (SIC) ve Hannan–Quinn bilgi kriteri (HQC-HIC) gibi seçim yöntemleri mevcuttur [17, 20, 21, 22]. Ancak, uzun hafızalı modellere bu seçim yöntemlerin uygulanması sonucunda etkinliği ve/veya yanlılığına ilişkin herhangi bir teorik sonuç bulunmamaktadır. Ancak, son zamanlarda pek çok gelişmeler mevcuttur.

Crato ve Ray, geniş kapsamlı bir simülasyon çalışması yaparak, üç farklı tahmin yöntemi ve üç otomatik model seçimi kriterinin tahmin doğruluğuna etkilerini karşılaştırmışlardır [23]. Schmidt ve Tschemig, ARFIMA(0, d, 0) süreçlerinde yanlılığı düzeltilmiş AIC (AICc), BIC ve HQC kriterlerinin performansını değerlendirmek için küçük ölçekli bir simülasyon çalışması gerçekleştirmişlerdir [24]. Bu kriterler temelde bir Gauss ARMA(p,q) sürecinin log olabilirlik fonksiyonunun yaklaşık maksimum değeri parametre sayısı için bir ceza cinsinden ifade edilmektedir. Kriterlerin birbirinden farklı olması, cezanın aşırı uyum eğilimine kaşı koymasından kaynaklanmaktadır. Bilgi kriterleri arasından en küçük değere sahip olan en iyi model olarak belirlenir. Bu bilgi kriterleri eşitlik 13 ve 14'teki gibi tanımlanır:

$$AIC = \ln \hat{\sigma}_N^2 + 2(p+q+1) \quad \textbf{(13)}$$

$$BIC = N \ln \hat{\sigma}_N^2 + (p+q+1)\log N \quad \textbf{(14)}$$

Burada, $\hat{\sigma}_N^2$, akgürültü varyansıdır [20].

III.UYGULAMA

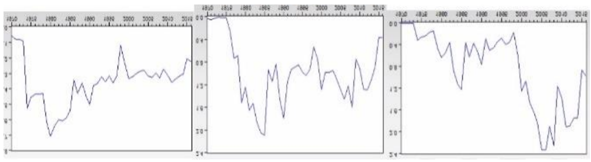
Bu çalışma için petrol kiraları (OR)'ye ait seriler dikkate alınmıştır [25]. Seriler, Türkiye, Hindistan ve Brezilya ülkeleri için 1970 ile 2016 dönemini kapsayan yıllık değerler içermektedir. Burada OR, dünyadaki ham petrol üretiminin değeri ile toplam üretim değerleri arasındaki fark olarak elde edilmiştir.

Türkiye, Hindistan ve Brezilya için serilerinin normallik varsayımının sağlanıp sağlanmadığı Tablo 1'de sunulmuştur. Türkiye serisinde normallik varsayımı sağlanmadığı için en uygun dönüşüm olan karekök dönüşümü uygulanmıştır. Türkiye, Hindistan ve Brezilya serilerinin sırasıyla tablo değerlerini incelersek karekök dönüşümü uygulanan Türkiye serisinde Jarque-Bera istatistiğine karşılık gelen $p = 0.879894$ olup 0.05 anlam düzeyinden büyüktür. Bu nedenle, bu serinin normal dağıldığı görülmektedir. Aynı şekilde, Hindistan ve Brezilya seriler için de Jarque-Bera istatistiğine karşılık gelen p değerleri, 0.05 anlam düzeyinden büyük olduğundan, bu serilerin de normal dağıldığı açıkça görülmektedir.

Tablo 1. Seriler için tanımlayıcı istatistikler

	TÜRKİYE	HİNDİSTAN	BREZİLYA
Mean	0.359	0.973	0.851
Median	0.335	0.971	0.599
Maximum	0.710	2.090	2.328
Minimum	0.059	0.002	0.000
Std. Dev.	0.151	0.542	0.693
Skewness	0.181	-0.125	0.695
Kurtosis	2.989	2.590	2.251
Jarque Bera	0.256	0.452	4.879
Probability	0.880	0.798	0.087

Türkiye, Hindistan ve Brezilya serilerinin çizgi grafikleri Şekil 1'de sunulmuştur. Her üç serinin durağan olmadığı açıkça görülmektedir. Ayrıca, Türkiye serisi için 1970 ile 1980 yılları arasında, Hindistan serisi için ise 1975 ile 1985 yılları arasında bir artış görülmektedir. Brezilya serisinde ise 2000 ile 2005 yılları arasında bir artış, 2005'ten sonrası için ise bir azalış söz konusudur.



Şekil 1. TÜRKİYE, HİNDİSTAN ve BREZİLYA serilerinin çizgi grafikleri

Tablo 2'de serilerin durağanlıklarını kontrol etmek için Augmented Dickey-Fuller (ADF) ve Phillipps-Perron (PP) birim kök testlerinin sonuçları yer almaktadır. Tablodaki serilerin ADF ve PP birim kök testlerinin test istatistiği ve olasılık değerlerine bakıldığında, Türkiye, Hindistan ve Brezilya serilerinin her iki birim kök testinde p-değerlerinin 0.05 anlam düzeyinden büyük olduğu görülmektedir. Bu nedenle, serilerin birim kök içerdiği ve dolayısıyla durağan olmadığı yorumu yapılabilir.

Tablo 2. Serilerin ADF ve PP birim kök test sonuçları

		ADF Testi		PP Testi	
		Sabit	Trend Sabit	Sabit	Trend Sabit
TÜRKİYE	t-istatistiği	-2.691	-3.152	-2.644	-3.056
	Prob	0.083	0.11	0.092	0.129
HİNDİSTAN	t-istatistiği	-2.613	-2.350	-2.475	-2.088
	Prob	0.098	0.08	0.1281	0.539
BREZİLYA	t-istatistiği	-1.968	-2.477	-1.888	-2.456
	Prob	0.300	0.338	0.335	0.347

Tablo 3'te ise birinci farkı alınan her üç serinin birim kök testlerinin sonuçları sunulmuştur. Sonuçlar incelendiğinde, serilerin p-değerleri 0.05 anlam düzeyine göre kıyaslandığında serilerinin durağan oldukları açıkça görülmektedir.

Tablo 3. Birinci fark seriler için ADF ve PP birim kök test sonuçları

		ADF Testi		PP Testi	
		Sabit	Trend Sabit	Sabit	Trend Sabit
TÜRKİYE	t-istatistiği	-7.328	-7.509	-7.450	-7.662
	Prob	0.000	0.000	0.000	0.000
HİNDİSTAN	t-istatistiği	-7.932	-8.132	-8.022	-9.323
	Prob	0.000	0.000	0.000	0.000
BREZİLYA	t-istatistiği	-7.927	-7.907	-7.937	-7.919
	Prob	0.000	0.000	0.000	0.000

Serilerin durağanlık varsayımı incelendikten sonra, uzun hafızanın varlığından bahsedebilmek için Hurst'ün R/S istatistiğinden yararlanılmıştır. Bu test istatistiğinde, güven aralıkları kıyaslanarak serilerde uzun hafıza varlığının olup olmadığı belirlenir. Tablo 4'te bu serilere ait uzun hafıza varlığının sonuçları yer almaktadır. Sonuçlar incelendiğinde, test istatistikleri değerlerinin %90, %95 ve %99 güven düzeylerinin tamamında aralığın dışında kaldığı görülmektedir. Bu durum, her üç serinin uzun hafızaya sahip olduğunu göstermektedir.

Tablo 4. Serilere ait R/S test sonuçları

Güven Düzeyleri	TÜRKİYE	HİNDİSTAN	BREZİLYA
90%	[0.861, 1.747]	[0.861, 1.747]	[0.861, 1.747]
95%	[0.809, 1.862]	[0.809, 1.862]	[0.809, 1.862]
99%	[0.721, 2.098]	[0.721, 2.098]	[0.721, 2.098]
Test İstatistiği	2.51	2.21	2.78

Bu çalışmada, $p, q \leq 2$ için tüm ARFIMA(p, d, q) ve ARIMA(p, d, q) modelleri üzerinde durulmuştur. En iyi modelin belirlenebilmesi için parametrelerin istatistiksel anlamlılığına bakılmıştır. Anlamlı parametreler * olarak işaretlenmiştir. Tablo 5'te Türkiye serisine ait ARFIMA ve ARIMA modellerinin parametre tahmin sonuçları verilmiştir. Her bir model için sonuçlar incelendiğinde, %5 anlam düzeyine göre kıyaslandığında ARFIMA$(1, d, 0)$ modelinde sabit ve AR(1) parametrelerinin anlamlı olduğu, ancak d parametresinin anlamlı olmadığı ve ARFIMA$(1, d, 1)$ modelinde ise sabit ve AR(1)'in anlamlı olduğu, fakat MA(1) ve d'nin anlamlı olmadığı görülmektedir. Bunun yanı sıra, ARFIMA(2,d,0) ve ARFIMA(0,d,2) modellerinde de bazı parametrelerin anlamlı olduğu görülmektedir. ARIMA modellerinden, ARIMA$(1,1,1)$ modelinde sadece AR(1) parametresi anlamlı iken, diğer ARIMA modellerinde ise hiçbir parametre tahminleri anlamlı değildir. Yani, serideki ilişkiyi açıklamakta ARIMA modeller yetersiz kalmaktadır. Seri için en iyi modelin ARFIMA$(0,0.423,1)$ olduğu açıkça görülmektedir.

Tablo 5. TÜRKİYE serisi için ARFIMA-ARIMA model sonuçları

	Sabit	AR(1)	AR(2)	MA(1)	MA(2)	d
ARFIMA(1,d,0)	0.323*	0.856*				-0.079
ARFIMA(0,d,1)	0.320*			0.403*		0.423*
ARFIMA(1,d,1)	0.327*	0.870*		0.105		-0.172
ARFIMA(2,d,0)	0.329*	1.008*	-0.118			-0.204
ARFIMA(0,d,2)	0.321*			0.413*	0.184	0.388
ARIMA(1,1,0)	0.004	-0.108				
ARIMA(0,1,1)	0.004			-0.144		
ARIMA(1,1,1)	0.001	0.864*		-1		
ARIMA(2,1,0)	0.004	-0.122	-0.126			
ARIMA(0,1,2)	0.004			-0.121	-0.121	

Hindistan serisine ait ARFIMA ve ARIMA modellerinin parametre tahmin sonuçları Tablo 6'da yer almaktadır. Sonuçlara göre, ARIMA modellerde %5 anlam düzeyine göre kıyaslandığında istatistiksel olarak anlamlı parametre tahmini bulunmadığından bu modeller seriyi açıklamada yetersiz kalmaktadır. ARFIMA modellerden ise ARFIMA(0,0.442,1) ile ARFIMA(0,0.342,2) modelleri dışında diğer ARFIMA modellerinin bazı parametre tahminleri anlamlı olmadığı açıkça görülmektedir.

Tablo 6. HİNDİSTAN serisine için ARFIMA-ARIMA model sonuçları

	Sabit	AR(1)	AR(2)	MA(1)	MA(2)	d
ARFIMA(1,d,0)	0.857*	0.862*				-0.127
ARFIMA(0,d,1)	0.182*			0.262*		0.442*
ARFIMA(1,d,1)	0.854*	0.858*		-0.025		-0.103
ARFIMA(2,d,0)	0.850*	0.804*	0.043			-0.079
ARFIMA(0,d,2)	0.836*			0.382*	0.411*	0.342*
ARIMA(1,1,0)	0.007	-0.184				
ARIMA(0,1,1)	0.008			-0.196		
ARIMA(1,1,1)	0.010	0.345		-0.537		
ARIMA(2,1,0)	0.007	-0.182	0.013			
ARIMA(0,1,2)	0.008			-0.202	-0.031	

Hindistan serisi için iki model arasından en iyiye karar verebilmek amacıyla Tablo 7'de yer alan AIC ve BIC bilgi kriterlerinden yararlanılmıştır. AIC ve BIC bilgi kriterlerinden en küçük olan ARFIMA(0,0.342,2) modelinin bu seri için en iyi model olduğu açık bir şekilde görülmektedir.

Tablo 7. HİNDİSTAN serisinin bilgi kriteri sonuçları

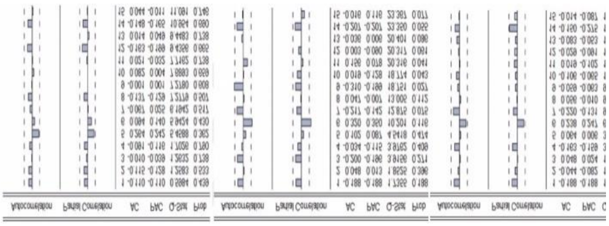
	ARFIMA(0,0.442,1)	ARFIMA(0,0.342,2)
AIC	46.956	43.928
BIC	54.356	53.179

Aynı şekilde, Brezilya serisine ait ARFIMA ve ARIMA modellerinin parametre tahmin sonuçları Tablo 8'de sunulmuştur. Bu tabloya göre, ARIMA modeller ile serideki ilişkiyi açıklamak yetersiz kalırken, ARFIMA modellerde ise sadece ARFIMA(0,0.462,1) modeli seriyi açıklamada yeterli olmaktadır.

Tablo 8. BREZİLYA serisi için ARFIMA-ARIMA model sonuçları

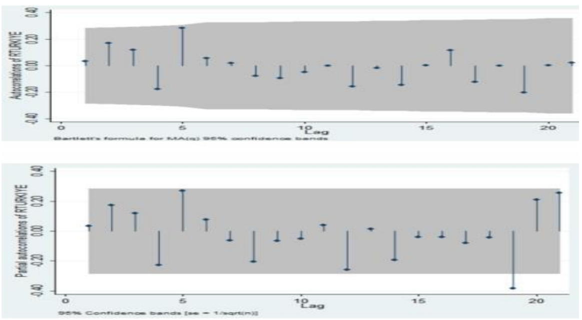
	Sabit	AR(1)	AR(2)	MA(1)	MA(2)	d
ARFIMA(1,d,0)	0.761*	0.927*				-0.136
ARFIMA(0,d,1)	0.804			0.322*		0.462*
ARFIMA(1,d,1)	0.751*	0.905		-0.133		-0.013
ARFIMA(2,d,0)	0.752*	0.785*	0.1097			-0.028
ARFIMA(0,d,2)	0.788			0.274	0.266	-0.028
ARIMA(1,1,0)	0.021	-0.184				
ARIMA(0,1,1)	0.022			-0.210		
ARIMA(1,1,1)	0.024	0.282		-0.484		
ARIMA(2,1,0)	0.022	-0.110	-0.088			
ARIMA(0,1,2)	0.023			-0.201	-0.065	

Şekil 2'de birinci farkı alınan Türkiye, Hindistan ve Brezilya serilerinin korelogramları sırasıyla sunulmuştur. Bu korelogramlar incelendiğinde, birinci farkı alınan serilerin sırasıyla otokorelasyon ve kısmi otokorelasyon değerlerinin güven aralığı içinde olduğu görülmektedir. Bir başka deyişle, bu seriler için AR ve MA terimlerinin etkisi devreden çıktığından dolayı ARIMA modeller anlamlı değildir. Böylece, her üç seri hem uzun hafızaya sahip hem de ilişkiyi açıklamada ARFIMA modelinin en iyi olduğu yorumu yapılabilir.

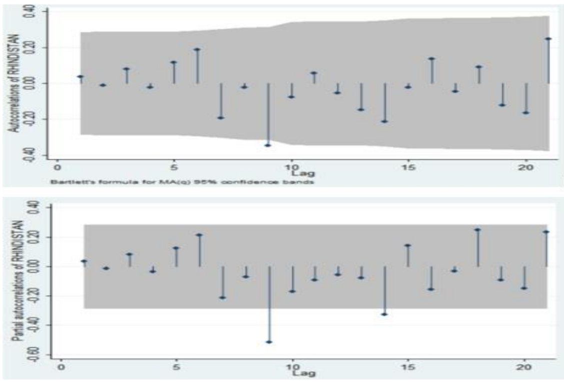


Şekil 2. TÜRKİYE, HİNDİSTAN ve BREZİLYA serilerinin korelogramları

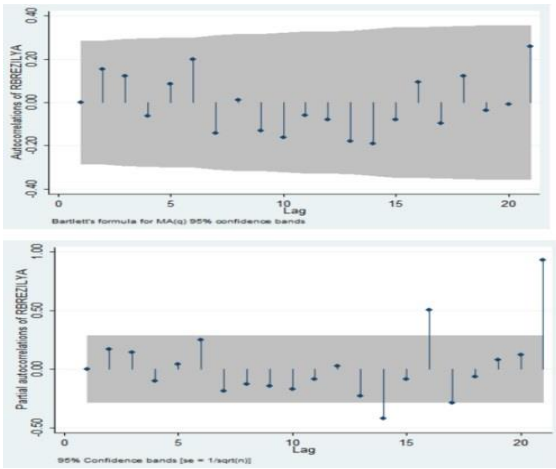
En iyi model belirlendikten sonra, bu modellerin artıkları için varsayımlar kontrol edilmelidir. Başka bir ifadeyle, modellere ait artıkların akgürültü (white-noise) özelliğine sahip olması gerekmektedir. Şekil 3'te Türkiye serisi için ARFIMA(0,0.423,1) modelinin artıklarının otokorelasyon fonksiyonu (ACF) ve kısmi otokorelasyon fonksiyonu (PACF) grafikleri verilmiştir. Bu grafikler incelendiğinde, modelden elde edilen artıkların gecikme değerleri güven bölgesinin içinde yer aldığından ilişkisiz olduğu görülmektedir. Aynı şekilde, Şekil 4'te verilen Hindistan serisi için ARFIMA(0,0.342,2) modelinin artıkları incelendiğinde, artıklar arasında PACF grafiğinde dokuzuncu ve on dördüncü gecikmelerde güven bölgesinin dışına çıkıldığı görülmüştür. Ancak, bu durum ihmal edilerek artıkların ilişkisiz olduğu yorumu yapılabilir. Brezilya serisi için ARFIMA(0,0.462,1) modelinin artık ilişki grafikleri ise Şekil 5'te sunulmuştur. Burada da PACF grafiğinde birkaç gecikme güven aralığının dışındadır. Fakat, ilk on gecikmenin ötesi ihmal edilerek artıkların ilişkisiz olduğu sonucuna varılabilir.



Şekil 3. TÜRKİYE serisi için ACF (üstteki) ve PACF (alttaki) grafikleri



Şekil 4. HİNDİSTAN serisi için ACF (üstteki) ve PACF (alttaki) grafikleri



Şekil 5. BREZİLYA serisi için ACF (üstteki) ve PACF (alttaki) grafikleri

Artıklar için diğer önemli varsayımlardan biri de otoregresif koşullu değişen varyans (ARCH) testi ile değişen varyanslılığın olup olmadığının belirlenmesidir. Tablo 9'da sırasıyla Türkiye, Hindistan ve Brezilya serileri için ARCH testi p değerlerinin sonuçları yer almaktadır. Sonuçlar incelendiğinde, ilk üç gecikmenin olasılık değerlerinin %5 anlam düzeyinden büyük olduğu görülmektedir. Bu nedenle, her üç serinin artıkları değişen varyansa sahip değildir.

Tablo 9. Seriler için ARCH testi p değerleri sonuçları

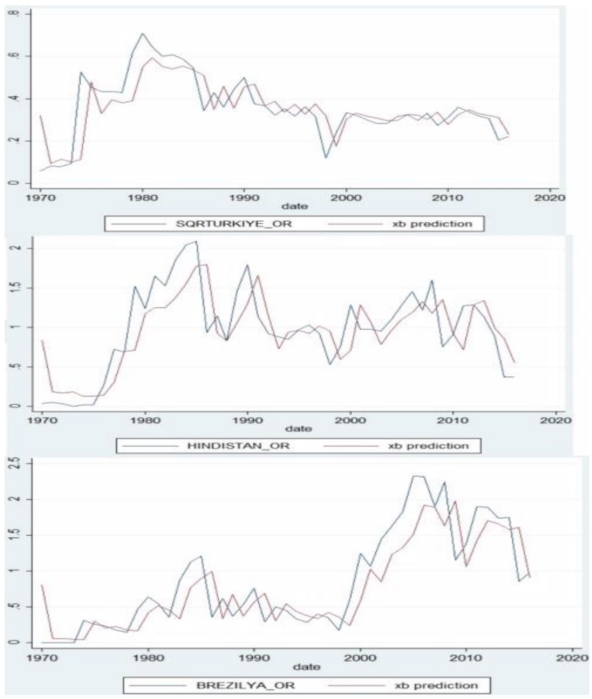
	1	2	3
TÜRKİYE	0.286	0.627	0.665
HİNDİSTAN	0.286	0.627	0.665
BREZİLYA	0.743	0.950	0.998

Aslında akgürültü, modelin geçerliliğinde ve tahmin edilmesinde önemli rol oynamaktadır. Artıklar için son varsayımlardan biri de akgürültü özelliğine sahip olması gerekliliğidir. Bunun için Bartlett ve Portmanteau testlerinden yararlanılmıştır. Tablo 10'da yer alan Bartlett ve Portmanteau testlerinin sonuçları incelendiğinde, her iki testin parantez içinde yer alan olasılık değerlerinin 0.05 anlam düzeyinden büyük olduğu görülmektedir. Bu nedenle, her üç serinin artıkları akgürültü özelliğine sahiptir. Başka bir deyişle, Türkiye, Hindistan ve Brezilya serileri için kurulan modeller geçerlidir ve ileriye yönelik tahmin yapılabilmektedir.

Tablo 10. Seriler için Bartlett ve Portmanteau istatistik sonuçları

	Bartlett'in (B) istatistiği	Portmanteau (Q) istatistiği
TÜRKİYE	0.719(0.679)	8.402(0.135)
HİNDİSTAN	0.698(0.715)	1.170(0.948)
BREZİLYA	0.864(0.445)	2.641(0.755)

Bütün varsayımlar incelendikten sonra, Türkiye, Hindistan ve Brezilya serilerinin gerçek ve kestirim değer grafikleri sırasıyla Şekil 6'da yer almaktadır. Grafiklere göre her üç seri için gerçek değerler ile kestirim değerlerinin birbirine yakın olması sebebiyle ARFIMA modellerinin seriler için uygun olduğu söylenebilmektedir.



Şekil 6. TÜRKİYE, HİNDİSTAN ve BREZİLYA serilerinin gerçek ve kestirim değer grafikleri

Tablo 11'de her üç seri için gelecek 5 yılın tahmin sonuçları ve parantez içinde gerçek değerleri yer almaktadır. ARFIMA modeli ile tahmin sonuçları incelendiğinde, Türkiye serisinde bir artışın olduğu ve bu artışın belli bir sürede kalıcı olabileceği görülmektedir. Hindistan serisi için de yine bir artış ve belli bir sürede kalıcılık söz konusudur. Brezilya serisi için ise belli bir dönem artış sonrasında bir azalış olduğu görülmektedir. Hata Kareler Ortalamasının Karekökü (RMSE) değerleri incelendiğinde, ARFIMA modeli Türkiye için çok iyi, Hindistan için makul ve Brezilya için nispeten düşük bir tahmin performansı göstermektedir. Modelin farklı ülkelerde farklı performans sergilemesi, her ülkenin petrol kiraları dinamiklerinin ve değişkenliklerinin farklı olmasından kaynaklanabilir. Türkiye için ARFIMA modelinin oldukça iyi bir tahmin yapması, modelin uzun hafızalı verilerde başarılı olduğunu ve Türkiye'nin petrol kiralarının belirgin bir uzun hafıza varlığı olduğunu düşündürebilir. Hindistan ve Brezilya için RMSE değerlerinin daha yüksek olması, bu ülkelerin petrol kiralarında daha fazla kısa dönemli dalgalanmalar veya modelin yakalayamadığı belirli olaylar olabileceğini gösterebilir.

Tablo 11. Seriler için 2017-2021 tahmini değer sonuçları

	TÜRKİYE	HİNDİSTAN	BREZİLYA
2017	0.070 (0.054)	0.4351 (0.313)	1.120 (1.128)
2018	0.078 (0.108)	0.608 (0.447)	1.1261 (1.901)
2019	0.082 (0.103)	0.719 (0.305)	1.123 (1.676)
2020	0.085 (0.061)	0.7593 (0.144)	1.115 (1.079)
2021	0.087 (0.138)	0.783 (0.326)	1.106 (2.605)
RMSE	0.031	0.310	0.794

IV.TARTIŞMA VE SONUÇLAR

Uzun hafızaya sahip bir zaman serisinin önceki dönemlerdeki davranışlarını modele yansıtmak, FI modellerinde öngörü başarısını artırmaktadır. ARIMA gibi modeller ise kısa hafızalıdır ve genellikle serinin son dönemlerdeki davranışlarına dayanarak modelleme yaparlar. Ayrıca, AR(10) ve AR(20) gibi birçok parametre içeren modeller, basitlik (parsimoni) ilkesine aykırı oldukları ve başarı seviyeleri düşük olduğu için uygun olmayabilir. Bu sebeple, ARFIMA modelleri, uzun hafızalı olmaları ve genellikle az sayıda parametre ile tahmin edilebilmeleri nedeniyle daha avantajlıdır [26].

Bu çalışmada, ekonomik seri olan petrol kiraları üzerinden verimli bir tahmin yapabilmek için ARFIMA modelini sunuyoruz. Bölüm 2'de tartışılan teorik çalışma, petrol kiralarının tahmin edilmesinde klasik modele kıyasla parametreleri tahmin etmede ve ileriye yönelik tahminlerde daha iyi sonuçlar vermiştir. Klasik ARIMA modelleri parametre tahmininde düşük bir performans sergilemiştir. Sayısal sonuçlara ve grafiğe baktığımızda da ARIMA modeli ile parametre tahmin etmede ARFIMA modeli alternatif olarak düşünülebilir.

Bu çalışmada sunulan ARFIMA modeli, diğer yarı-parametrik veya parametrik yöntemlerle genişletilebilir. Örneğin, model performansı ve belirlenmesi, Whittle metodu gibi tekniklerle sağlanabilir. Yarı-parametrik veya parametrik yöntemlerle yapılan karşılaştırmalı analizler, modelin avantajlarını ve sınırlamalarını daha iyi anlamamıza olanak tanır. Bu karşılaştırmalar, ARFIMA modelinin farklı veri setlerinde ve koşullarda nasıl davrandığını incelemek için faydalı olabilir.

ARFIMA modeli çalışmasında, modelin belirlenmesi ve optimizasyonu için bilgi kriterlerinden yararlanılmaktadır. Modelin, en iyi bilgi kriterine göre değerlendirilmesi, farklı yapılandırmalarda nasıl performans gösterdiğini ve en uygun parametrelerin nasıl belirlendiğini ortaya koyabilir. Ayrıca, ARFIMA modelinin çeşitli veri setleri ve uygulama alanlarında test edilmesi, modelin genelleştirilebilirliğini ve pratik faydasını artırabilir. Örneğin, finansal zaman serileri, iklim verileri veya ekonomik göstergeler gibi farklı alanlarda modelin performansı incelenebilir. Bu, modelin doğruluğunu ve genel performansını artırma potansiyeline sahiptir.

TEŞEKKÜR

Bu çalışma, Mimar Sinan Güzel Sanatlar Üniversitesi Fen Bilimleri Enstitüsü İstatistik Bölümünde hazırlanmış “Otoregresif Kesirli Bütünleşik Hareketli Ortalama (Arfima) Modelinin Belirlenmesi: Türkiye, Hindistan ve Brezilya Ülkeleri Üzerine Uygulamalar” adlı yüksek lisans tezinden üretilmiştir.

KAYNAKLAR

[1] Kutlar, A.ve Turgut, T. (2006). Türkiye’deki Başlıca Ekonomi Serilerinin ARFIMA Modelleri ile Tahmini ve Öngörülebilirliği. Kocaeli Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 11(1), 120-149.

[2] Barışık, S. ve Çevik, E. İ. (2008). İşsizlikte Histeri Etkisi: Uzun Hafıza Modelleri. Kamu-İş, 9(4), 1-36.

[3] Çevik, E. İ. ve Erdoğan, S. (2009). Bankacılık Sektörü Hisse Senedi Piyasasının Etkinliği: Yapısal Kırılma ve Güçlü Hafıza. Doğuş Üniversitesi Dergisi, 10 (1), 26- 40.

[4] Pekkaya, M. (2011). Arfima ve Figarch Yöntemlerinin Markowitz Ortalama Varyans Portföy Optimizasyonunda Kullanılması: İMKB-30 Endeks Hisseleri Üzerine Bir Uygulama. Doktora Tezi, Zonguldak Bülent Ecevit Üniversitesi, Türkiye.

[5] Cevik, E. İ. (2012). İstanbul Menkul Kıymetler Borsası’nda Etkin Piyasa Hipotezinin Uzun Hafıza Modelleri İle Analizi: Sektörel Bazda Bir İnceleme. Yaşar Üniversitesi E-Dergisi, 7(26), 4437-4454.

[6] Karia, A. A., Bujanga, I. ve Ahmad, İ. (2013). Fractionally Integrated ARMA for Crude Palm Oil Prices Prediction: Case of Potentially Overdifference. Journal of Applied Statistics, 40(12), 2735–2748.

[7] Pekkaya, M., ve Albayrak, A. S. (2013). Arfima ve Figarch Yöntemlerinin Markowitz Ortalama Varyans Portföy Optimizasyonunda Kullanılması: İMKB-30 Endeks Hisseleri Üzerine Bir Uygulama. İstanbul Üniversitesi İşletme Fakültesi Dergisi, 42(1), 1303-1732.

[8] Yılmaz, Ç. (2015). Hisse Senedi Getirilerindeki Uzun Hafıza Etkisinin ve Volatilitenin Belirlenmesi: BRIC Ülkeleri Üzerine Bir Uygulama. Yüksek Lisans Tezi, İstanbul Üniversitesi, Türkiye.

[9] Boateng, A., Gil-Alana, L. A., Maseka, L., Siweya, H. ve Belete, A. (2016). Long Memory and Arfima Modelling: The Case of CPI Inflation Rate in Ghana. The Journal of Developing Areas, 50(3), 287-304.

[10] Omekara C. O., Okereke O. E. ve Ukaegeu L. U. (2016). Forecasting Liquidity Ratio of Commercial Banks in Nigeria. Microeconomics and Macroeconomics, 4(1), 28-36.

[11] Fahreddinoğlu, H. G. (2017). Enflasyon’daki Hareketliliğin ARIMA ve ARFIMA Modelleriyle Belirlenmesi ve Tahmin Edilmesi: Azerbaycan Örneğinde. 3. Uluslararası Öğrenciler Sosyal Bilimler Kongresi, İstanbul, Türkiye.

[12] Joshi, P. (2016). What is Long Memory in Time Series Analysis.https://prateekvjoshi.com/2016/08/27/what-is-long-memory-in-time-series-analysis/.

[13] Palma, W. (2007). Long Memory Time Series Theory and Methods. John Wiley & Sons Inc., Canada.

[14] Granger, C. W. J. ve Joyeux, R. (1980). An Introduction to Long-Memory Time Series Models and Fractional Differencing. Journal of Time Series Analysis, 1(1), 15-29.

[15] Hosking, J. R. M. (1981). Fractional Differencing. Biometrika, 68(1), 165-176.

[16] Maddala, G. S. ve Kim, In-M. (2003). Unit Roots, Cointegration and Structural Changes. Cambridge University Press, New York.

[17] Sowell, F. (1992). Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models. Journal of Econometrics, 53(1-3), 165-188.

[18] Geweke, J. ve Porter-Hudak, S. (1983). The Estimation and Application of Long Memory Time Series Models. Journal of Time Series Analysis, 4(4), 221-238.

[19] Beran, J., Feng, Y., Ghosh, S. ve Kulik, R. (2010). Long-Memory Processes Probabilistic Properties and Statistical Methods, Springer, London.

[20] Bisaglia, L. (2002). Model Selection for Long-Memory Models. Quaderni di Statistica, 4.

[21] Hosking, J. R. M. (1984). Modeling Persistence in Hydrological Time Series Using Fractional Differencing. Water Resources Research, 20(12),1898-1908.

[22] Cheung, Y. W. (1993). Long-Memory in Foreign Exchange Rates. Journal of Business and Economic Statistics, 11(1), 93-101.

[23] Crato, N. ve Ray, B. K. (1996). Model Selection and Forecasting for Long-Range Dependent Processes. Journal of Forecasting, 15(2), 107-125.

[24] Schmidt, C. M. ve Tschernig, R. (1993). Identification of Fractional ARIMA Models in The Presence of Long Memory. Discussion Paper, University of Munich.

[25] World Bank Databank, World Development Indicators, Oil rents. (2019). https://databank.worldbank.org/data/reports.aspx?dsid=2&series=NY.GDP.PETR.RT.ZS.

[26] Caballero, R., Jewson, S. ve Brix, A. (2002). Long Memory in Surface Air Temprature: Detection, Modeling, and Aplication to Weather Derivative Valuation. Climate Research, 21(2), 127-140.