

KARAYOLLARI ENERJİ TÜKETİMİNİN YAPAY ZEKA VE REGRESYON YÖNTEMLERİ İLE MODELLENMESİ

Ömer Faruk CANSIZ *^{ID}

Kevser ÜNSALAN *^{ID}

İbrahim ERGİNER *^{ID}

Alınma: 13.04.2020; düzeltme: 06.05.2020; kabul: 09.09.2020

Öz: Doğanın temelini dayanan iş-enerji ilişkisi, bize hareketin ve iş yapabilmenin şartının enerji olduğunu göstermektedir. Gelişen teknoloji, sanayileşme gibi faktörler üretimi arttırırken enerji tüketiminin de artmasına neden olmaktadır. Sanayileşmenin bir kolu olan ulaştırma sektörü enerji tüketiminde sektör bazında önemli bir yere sahiptir. Bu çalışmada, ulaştırma sektöründe özellikle Türkiye’de taşımacılık potansiyelinin yüksek olduğu karayolu taşımacılığında enerji tüketimi incelenmektedir. Çalışma kapsamında yapay zekâ tekniklerinden yapay sinir ağları (YSA) ve uyarlamalı sinirsel bulanık çıkarım sistemi (ANFIS), regresyon tekniklerinden ise çok değişkenli lineer regresyon (MLR) yöntemleri kullanılarak karayolu taşımacılığında enerji tüketimi tahmin modellenmesi yapılmaktadır. Modellemede karayolu yol ağı uzunluğu, taşıt-km, ağırlıklı ortalama günlük trafik (AOGT), motorlu taşıt sayısı ve nüfus parametreleri bağımsız değişken olarak incelenmektedir. Tahmin modellerinin karşılaştırılmasında determinasyon katsayısı (R^2), hataların karesinin ortalaması (HKO) ve ortalama yüzde hata (OYH) performans kriterleri dikkate alınmaktadır. Performans kriterlerine göre en iyi model lineer regresyon yöntemi ile elde edilmektedir. En iyi modelin R^2 , HKO, OYH değerleri sırasıyla 0,9474, 54084 ve % 4,86 çıkmaktadır. Geliştirilen model ile ulaşım politikalarına yön verilmesi hedeflenmektedir.

Anahtar Kelimeler: Karayolları Enerji Tüketimi, YSA, ANFIS, MLR

Modeling of Highway Energy Consumption by Artificial Intelligence and Regression Methods

Abstract: The relationship between work and energy, which is based on nature, shows us that energy is the condition for action and ability for doing work. While developing technology and industrialization factors increase production, it also causes an increase in energy consumption. The transportation sector, which is a branch of industrialization, has an important place on the basis of sector in energy consumption. In this study, energy consumption are studied in transportation sector especially road transportation of freight is high potential in Turkey. Within the scope of the study, energy consumption prediction modeling is made by using artificial neural networks (ANN) and adaptive neural fuzzy inference system (ANFIS) from artificial intelligence techniques, and multivariate linear regression (MLR) methods from regression techniques. In modeling, highway road network length, vehicle-km, weighted average daily traffic (WADT), number of motor vehicles and population parameters are examined as independent variables. When comparing the prediction models, the determination coefficient (R^2), the mean square error (MSE) and the average percentage error (APE) performance criteria are taken into consideration. According to

* İskenderun Teknik Üniversitesi Mühendislik ve Doğa Bilimleri Fakültesi İnşaat Mühendisliği Bölümü 31200 İskenderun/HATAY

İletişim Yazarı: İbrahim ERGİNER (ibrahim.erginer.mfbe18@iste.edu.tr)

performance criteria, the best model is obtained by linear regression method. R^2 , HKO, OYH values of the best model are 0.9474, 54084 and 4.86%, respectively. With the developed model, it is aimed to direct transportation policies.

Keywords: Road Transportation Energy Consumption, ANN, ANFIS, MLR

1. GİRİŞ

Dünya ekonomisinde önemli yere sahip olan enerji, hem üretimi hem tüketimi açısından ülke politikalarını etkilemektedir. Ülkelerin iktisadi ve sosyal kalkınmaları bakımından enerji tüketimi, çağımızın önemli konuları arasında yer almaktadır. Nitekim nüfus ve sanayileşme ile birlikte enerji tüketim seviyesi gelecekte üst seviyelere ulaşacak ve popülerliğini korumaya devam edecek ve Türkiye’de de benzer durum söz konusu olacaktır (Mucuk ve Uysal, 2009).

Gelişmekte olan ülkelerde hızlı nüfus artışı ile birlikte yerleşim alanı sınırları, geniş alanlara yayılmaktadır. Bu durum kent içi trafiğinin önemli bir unsuru olan karayolları ağının gelişmesi gerekliliğini göstermektedir. Türkiye’de bu gibi sebeplerle karayolları ağı yıldan yıla genişlemektedir. Karayolu ağının genişlemesi ve nüfusun artışına paralel olarak teknoloji çağının gelişmesiyle birlikte taşıt sahipliğinin de artması ile yenilenemez kaynakların kullanımındaki ani artışlar, karayolları enerji tüketimini incelenmesi gereken önemli bir konu haline getirmektedir (Cansız vd. 2009).

Ulaşım sektöründe enerji tüketimini optimize etmeye yönelik birçok çalışma bulunmaktadır (Haldenbilen, 2003- Solak, 2013). Karayolu ulaştırmasında yenilenemez enerji tüketiminin azaltılmasına yönelik çalışmalar, alternatif enerji kaynaklarının kullanımı üzerine yoğunlaşmaktadır (Öztürk, 2018- Gündüz vd, 2011).

Enerji tüketiminin optimize edilmesi hususunda, ülkelerin gelecek enerji tüketim miktarları konusunda bilgi sahibi olması gereklidir. Bu nedenle enerji tüketimi konusunda performansı yüksek tahmin modellerinin geliştirilmesi oldukça büyük önem arz etmektedir. Bu konu üzerine farklı değişkenlerin kullanıldığı ve sektör bazlı enerji tüketim tahmin modellerinin geliştirildiği birçok çalışma mevcuttur (Murat ve Ceylan, 2006). Cansız (2007), YSA metodunu kullanarak Türkiye için 1970 ve 2020 yılları arasındaki ulaştırma sektöründeki enerji tüketiminin mevcut durumu ve gelişimi üzerine analiz yapmaktadır. Bir başka çalışmada YSA metodu kullanılarak Türkiye’nin enerji talebi için tahmin modelleri geliştirilmektedir. Tahmin modellerinde gayri safi yurtiçi hâsıla (GSYH), nüfus, ithalat-ihracat miktarı, bina yüz ölçümü ve taşıt sayısı bağımsız değişken olarak irdelenmektedir (Es vd, 2014).

Bu çalışmada enerji tüketiminde önemli bir yere sahip karayolu taşımacılığı için tahmin modelleri geliştirilmektedir. Tahmin modellerinde karayolu ağı uzunluğu, taşıt-km, ağırlıklı ortalama günlük trafik (AOGT), motorlu taşıt sayısı ve nüfus sayısı bağımsız değişken olarak dikkate alınmaktadır. Modellemede yapay zekâ tekniklerinden nonlineer problemlerin çözümünde üstün performansa sahip YSA ve bulanık çıkarım sistemlerinin YSA tekniği ile geliştirilmiş hali olan ANFIS metodu kullanılmaktadır. İstatistiksel metotlardan ise çok değişkenli lineer regresyon (MLR) yöntemi kullanılmaktadır.

2. MATERYAL VE METOD

2.1. Materyal

Çalışmanın bu bölümünde tahmin modelleri için belirlenen bağımlı ve bağımsız değişkenler incelenmektedir. Tahmin modellerinde karayolu ağı uzunluğu (KU), taşıt-km (TK), ağırlıklı ortalama günlük trafik (AOGT), motorlu taşıt sayısı (TS) ve nüfus (N) bağımsız değişken olarak incelenirken karayollarında tüketilen enerji (ET) miktarı bağımlı değişken olarak incelenmektedir. Değişkenlerin 2000-2018 yılları arasındaki değişimi Tablo 1’de verilmektedir.

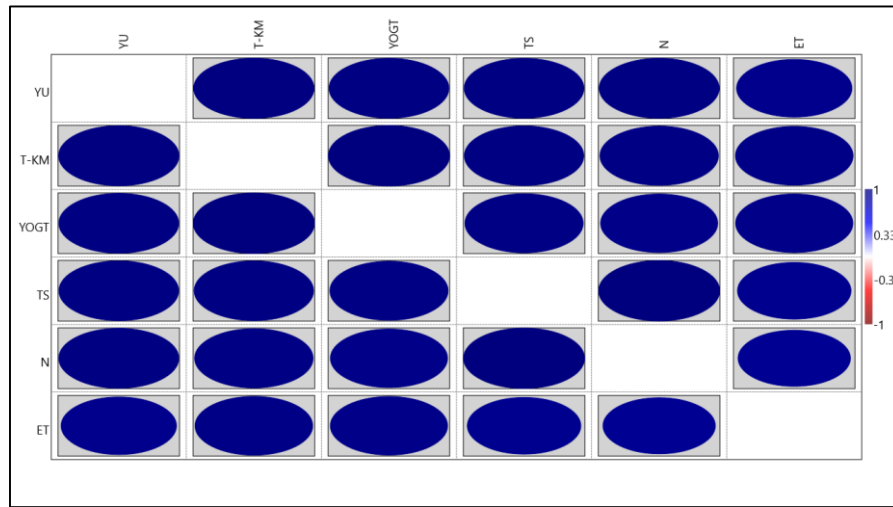
Tablo 1. Değişkenlerin 2000-2018 yılları arasındaki değişimi

Yıllar	KU	TK	AOGT	TS	N	ET
2000	417406	56151000000	3894	8320449	63240194	10505
2001	426249	52631000000	3706	8521956	64192243	9801
2002	427411	51664000000	3581	8655170	65145367	9967
2003	428415	52349000000	3590	8903843	66089402	10945
2004	349215	57767000000	3907	10236357	67010930	11510
2005	349238	61129000000	4032	11145826	67903469	11781
2006	349304	64577000000	4147	12227393	68756810	12599
2007	350708	69609000000	4442	13022945	69581848	13594
2008	351958	69771000000	4416	13765395	70418604	12624
2009	362660	72432000000	4568	14316700	71321399	12516
2010	367263	80124000000	4917	15095603	72326988	13535
2011	370276	85495000000	5426	16089528	73443863	15451
2012	385748	93989000000	5647	17033413	74653016	16303
2013	388783	99431000000	5937	17939447	75928564	17677
2014	236794	102988000000	6152	18828721	77231907	18497
2015	238899	113274000000	6783	19994472	78529409	21525
2016	242590	119671000000	7152	21090424	79821724	23478
2017	247514	127997000000	7670	22218945	81101892	24731
2018	247553	131625000000	7938	22865921	82319724	24795

Değişkenler arasındaki çapraz korelasyon ilişkisi ve değerleri Şekil 1 ve Tablo 2’de görülmektedir. $\alpha=0,05$ anlamlılık düzeyinde ET değişkeni ile diğer değişkenler arasında anlamlı bir ilişki olduğu görülmektedir. Bu nedenle ET tahmin modellerinin oluşturulmasında belirlenen tüm değişkenler dikkate alınmaktadır. ET ile en güçlü ilişki TK arasında çıkmaktadır. Bu da km başına düşen taşıt miktarının karayolu enerji tüketimini önemli ölçüde etkilediğini göstermektedir.

Tablo 2. Değişkenler arasındaki çapraz korelasyon değerleri

	KU	TK	AOGT	TS	N	ET
KU		0,99084	0,98636	0,98523	0,98779	0,94916
TK	0,99084		0,99774	0,98289	0,97521	0,96828
AOGT	0,98636	0,99774		0,97291	0,96386	0,96475
TS	0,98523	0,98289	0,97291		0,9964	0,93416
N	0,98779	0,97521	0,96386	0,9964		0,92445
ET	0,94916	0,96828	0,96475	0,93416	0,92445	

**Şekil 1:**

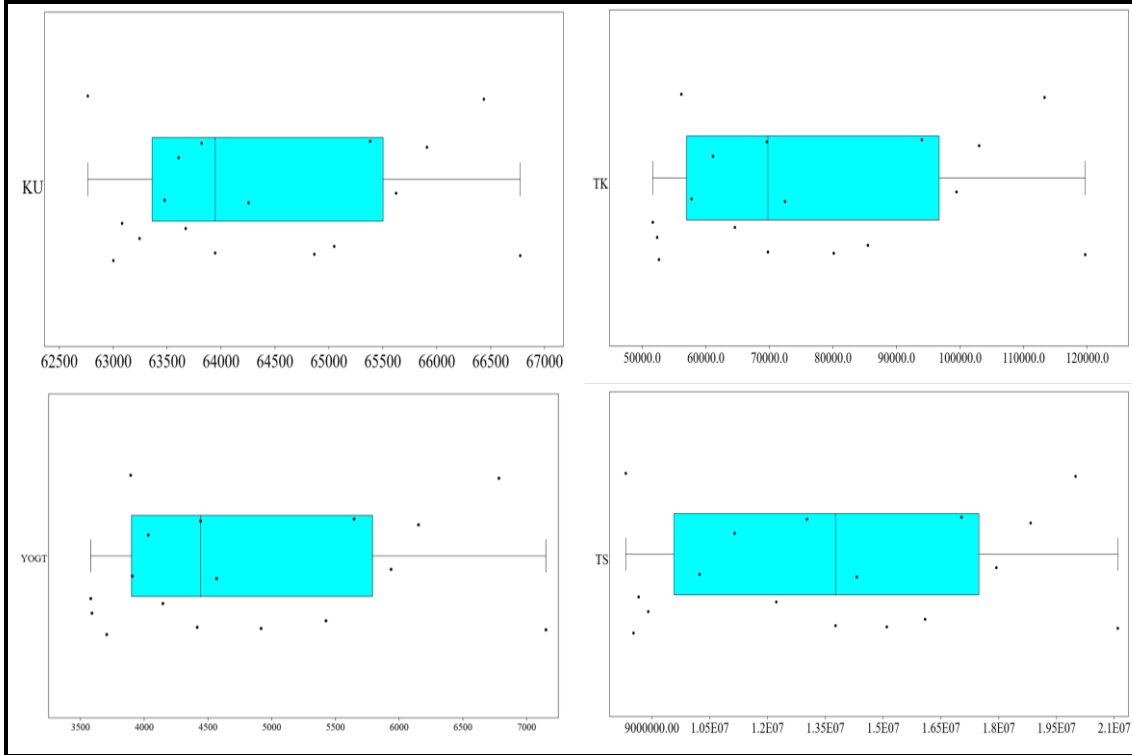
Değişkenler arasındaki çapraz korelasyon değerleri

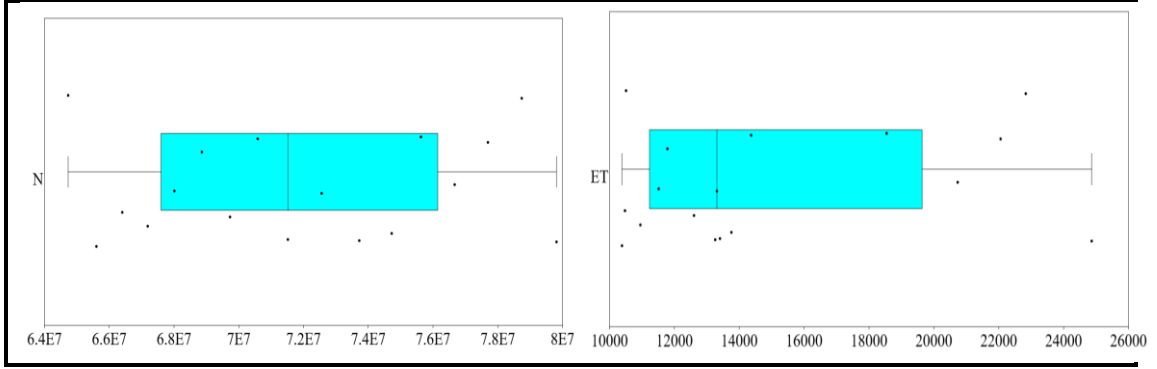
Değişkenlerin dağılım özelliklerini incelemek amacıyla Tablo 3'te özet istatistikleri verilmektedir. Verilerin çarpıklık katsayılarına bakıldığında ± 1 aralığında kaldığı görülmektedir. Bu da değişkenlerin normal dağılıma uygun olduğu ve tahmin modellemesi yapılmadan önce herhangi bir transformasyona gerek olmadığını göstermektedir.

Tablo 3. Değişkenlerin özet istatistikleri

	KU	TK	AOGT	TS	N	ET
N	17	17	17	17	17	17
Minimum	62764	51664	3581	8320449	6,473E07	10383
Maksimum	66774	119671	7152	2,109042E07	7,9815E07	24866
Ortalama	64406,12	76650,12	4841	1,383457E07	7,1893E07	15020,53
Standart Hata	305,8557	5377,515	276,4675	1019505	1159757	1167,768
Varyans	1590311	4,916004E08	1299383	1,766963E13	2,286563E13	2,318259E07
Standart Sapma	1261,075	22172,06	1139,905	4203526	4781802	4814,831
%25	63360	56959	3900,5	9570100	6,75985E07	11232
%75	65502,5	96710	5792	1,748643E07	7,61475E07	19638,5
Çarpıklık Katsayısı	0,530932	0,6458525	0,7580661	0,1989828	0,1417364	0,9828822
Basıklık Katsayısı	-0,9728232	-0,8191243	-0,6041092	-1,174981	-1,219265	-0,4891555
Varyasyon Katsayısı	1,958005	28,92633	23,54688	30,38423	6,651277	32,055

Şekil 2'de verilerin dağılımı box-jitter grafikleri ile incelendiğinde, çarpıklık katsayısı sıfıra yakın olan TS ve N değişkenlerinin kutu grafiğinin ikiye bölünme durumunun daha eşit olduğu görülmektedir. Bu da TS ve N değişkenlerinin normal dağılıma daha uygun olduğunu göstermektedir. KU, TK, AOGT ve ET değişkenleri her ne kadar normal dağılıma uygun olsa da sağa çarpıktır.





Şekil 2:
Değişkenlerin box-jitter grafikleri

2.2. Metod

2.2.1. MLR

Mühendislik problemlerinde iki ya da daha fazla değişkenin aldıkları değerlerin birbirinden bağımsız olmadığı görülmektedir. Dolayısıyla bu değişkenler arasında bir ilişki olduğu kabul edilmektedir (Bayazıt ve Oğuz, 2005). Bu değişkenler arasındaki ilişkiyi incelemek amacıyla regresyon yöntemleri kullanılmaktadır. Çalışmada değişken sayısı 2’den fazla olduğu için çok değişkenli lineer regresyon (MLR) yöntemi kullanılmaktadır. MLR yönteminde bağımsız değişkenlerin bağımlı değişkenler üzerindeki etkisi regresyon katsayısı ile ifade edilmektedir. Regresyon denklemindeki bu katsayı bağımsız değişkenin bağımlı değişkene olan bağımlılık derecesini ifade etmektedir. İki regresyon yönteminin genel formu denklem 1 ve 2’de verilmektedir.

$$Y_i = (B_0 + B_1X_1 + B_2X_2 + \dots + B_nX_n) + \varepsilon_i \quad (1)$$

$$Y_i = (B_0 + B_1X_1 + B_2X_2 + \dots + B_nX_n + B_{n+1}X_1^2 + B_{n+2}X_2^2 + \dots + B_mX_n^2) + \varepsilon_i \quad (2)$$

Denklemlerde “Y” bağımlı değişkeni, “B” regresyon katsayılarını ve “ε” hata bileşenini ifade etmektedir.

2.2.2. YSA

YSA problemlerin çözümündeki gücünü paralel dağılmış yapısından, öğrenme ve genelleştirme kabiliyetinden almaktadır. Genelleme özelliği eğitim ve öğretim sürecinde ağır, karşılaşılmayan verilere karşı uygun tepkileri vermesidir. Yöntemin bu özelliği karmaşık problemlerin çözümünde başarıya ulaşmasına katkı sağlamaktadır (Ergezer vd, 2003). Her ne kadar problemlerin çözümünde başarılı sonuçlara ulaşılsa da, modelde ağırlıkların atanması kara kutu mantığına dayanmaktadır. Yani ağırlıkların nasıl atandığı bilinmemektedir. YSA matematiksel olarak ifade edilemeyen problemlerin çözümünde kullanılan bir yöntem olup kara kutu model olarak tanımlanmaktadır (Cansız ve Easa, 2011).

Bu çalışmada tahmin modelleri oluşturulurken epoch sayısı 10 ile sınırlandırılması uygun görülmektedir. Modellerin eğitim algoritmasında Gradient Descent, gradient descent with momentum, levenberg-marquardt, resilient backpropagation öğrenim algoritmaları kullanılmaktadır. Analiz sonuçları hataların karesinin ortalaması (HKO) performans kriterine göre sonlandırılmakta ve yeniden ağır eğitilmesi için ağırlıklar yenilenmektedir. Bu döngüye göre toplamda her eğitim algoritması için 3000 YSA modeli oluşturulmaktadır. Toplamda 12000

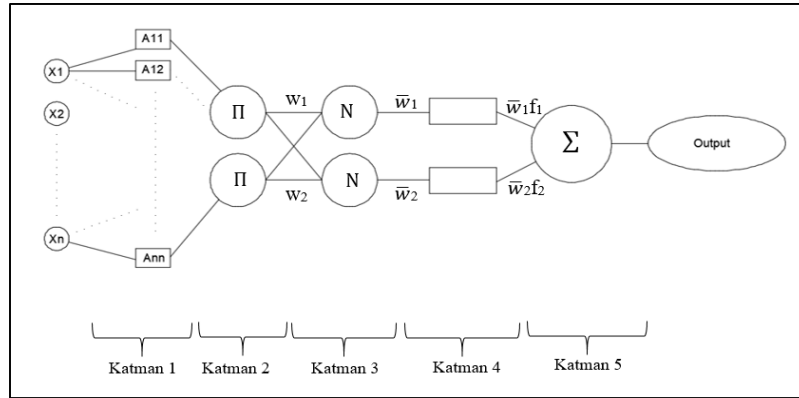
modelin HKO değerleri karşılaştırılarak en iyi YSA modeli belirlenmektedir. YSA modelinin genel formülü Denklem 3’te verilmektedir.

$$Y = f \left(\sum_{i=1}^n X_i * W_i + b_i \right) \quad (3)$$

Denklem 3’te Y bağımlı değişkeni, X değerleri bağımsız değişkenleri, W modelin katman ağırlıklarını ve b modelin bias değerini göstermektedir.

2.2.3. ANFIS

Uyarlamalı Sinirsel bulanık çıkarım sisteminin (ANFIS) yapısı, bulanık çıkarım sistemini (FIS) esas alan YSA modelinden meydana gelmektedir. ANFIS, YSA ve bulanık çıkarım (FIS) yöntemlerin birleşimi olduğu için her iki yöntemde avantajları ve dezavantajlarından yararlanan hibrit bir yöntemdir. Bulanık çıkarım modellemesi deneme-yanılma tekniğine dayanmaktadır. Fakat sinir ağı modellemesi ile birleşince üyelik fonksiyonlarının ve bulanık kuralların oluşturulmasında kolaylık sağlamaktadır (Murat, 2006). Yöntem nonlinear problemlerin çözümünde ve karmaşık zaman serilerinin modellenmesinde kullanılmaktadır (Jang, 1991). FIS temel yapısı olan üyelik fonksiyonlarının tasarımı ANFIS modelinin performansını belirleyici önemli bir etkidir. Öğrenme yeteneği ile YSA metodu ise FIS anatomisinde kuralların belirlenmesi için sistematik bir süreç sunabilmektedir (Jang vd, 1997).



Şekil 3:
ANFIS genel yapısı

Genel ANFIS yapısı Şekil 3’te görülmektedir. Burada “Xi”; “Aii”; “Π”; “N” ve “wi” sırasıyla girdi parametrelerini, üyelik fonksiyonlarını, kuralları, bulanık kurallardan sonraki ağırlıkları temsil etmektedir. Yapıdaki katmanlarda geçen süreçler aşağıdaki gibi açıklanmaktadır.

Katman 1: Bu katmanda her bir düğüm üyelik fonksiyonuna karşılık gelmektedir. Yani her bir değişken için üyelik fonksiyonlarının oluşturulduğu katmandır.

Katman 2: Bu katman bir önceki katmandan gelen sinyalleri çoğaltan ve bir sonraki katmana ileten düğümlerden meydana gelmektedir. Yani bulanık kuralların oluşturulduğu katmandır.

Katman 3: Bu aşamada bir önceki katmandan gelen sinyaller, uygun ağırlıkların belirlenmesi amacıyla normalize edilmektedir.

Katman 4: Üçüncü katmandan gelen sinyaller bu katmanda, polinomal eşitlikte çarpılmakta ve bulanık kurallar ile işlenmektedir.

Katman 5: Dördüncü katmandan gelen tüm sinyaller, beşinci katmanda toplanmakta ve ağırlık çıkışına ulaşılmaktadır.

2.2.4. Veri Setlerinin Ayrılması

MLR analizinde eğitim ve test data seçiminde 19 datanın 15'inin eğitim datası olarak alınmaktadır. Toplam 19 yıllık data seti son 3 yıl 3'lü olmak üzere, 4'er yıllık veri seti gruplarına ayrılmaktadır (Tablo 4). Toplamda dört farklı veri seti seçim kombinasyonu oluşmaktadır.

Tablo 4. MLR için k-fold veri ayrımı

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
1. k-fold	TEST			EĞİTİM			EĞİTİM			EĞİTİM			EĞİTİM						
2. k-fold	EĞİTİM			TEST			EĞİTİM			EĞİTİM			EĞİTİM						
3. k-fold	EĞİTİM			EĞİTİM			TEST			EĞİTİM			EĞİTİM						
4. k-fold	EĞİTİM			EĞİTİM			EĞİTİM			TEST			EĞİTİM						

Tablo 5. MLR analizinde bağımlı ve bağımsız değişkenler arasındaki ilişki

	β katsayıları ve t testi sonuçları ($\alpha=0,05$ için p değerleri)															
	Linear								Purequadratic							
	1. k-fold		2. k-fold		3. k-fold		4. k-fold		1. k-fold		2. k-fold		3. k-fold		4. k-fold	
	standardize edilmiş β	p	standardize edilmiş β	p	standardize edilmiş β	p	standardize edilmiş β	p	standardize edilmiş β	p	standardize edilmiş β	p	standardize edilmiş β	p	standardize edilmiş β	p
Sabit		0,005765		0,067445		0,310729		0,193812		0,012767		0,869498		0,964737		0,543561
KU	-0,14045	0,002321	-0,0884	0,176442	-0,07994	0,258963	-0,19009	0,029163	-	-	-0,053	0,980662	0,092	0,896454	-1,378	0,166088
TK	1,302926	0,030996	0,462402	0,528188	-0,56793	0,710347	1,153796	0,279226	1,426862	0,042089	0,648	0,578968	0,298	0,7713	0,230	0,857293
AOGT	0,844021	0,047724	0,786506	0,182287	1,307567	0,234525	0,106026	0,896605	0,231233	0,778308	0,297	0,842416	-	-	1,101	0,418769
TS	0,802015	0,114967	-1,13855	0,022015	-0,26014	0,745293	-1,09213	0,022446	0,568186	0,243705	-0,880	0,656706	-0,147	0,899376	-0,638	0,432848
N	-2,06774	0,006683	0,79844	0,057228	0,450609	0,321919	0,642389	0,081305	-	-	-	-	-	-	-	-
KUxKU,	-	-	-	-	-	-	-	-	-0,13644	0,007176	-0,026	0,990916	-0,179	0,803357	-1,082	0,225692
TKxTK	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

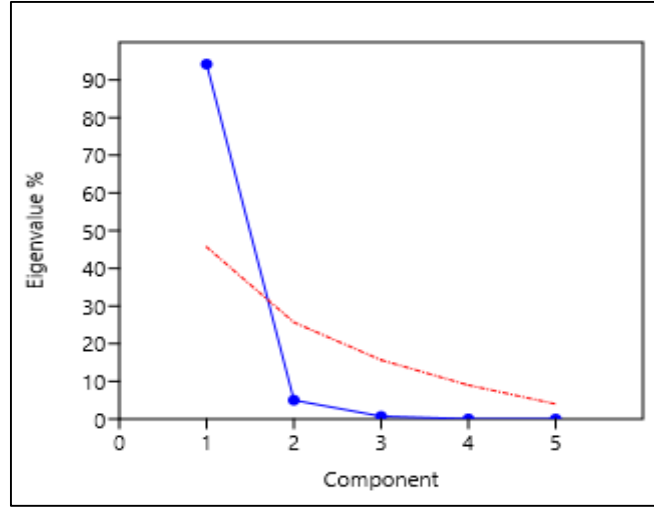
AOGTxAOG T	0,52894	0,22945	0,190	0,833532	0,512	0,281226	-0,570	0,48562
TSxTS	-	-	-	-	-	-	-	-
NxN	-1,86252	0,01649	0,661	0,5946	0,264	0,684636	0,563	0,311818

Modelde bağımlı ve bağımsız değişkenler arasındaki ilişki için t testi yapılmaktadır. T testi sonuçlarına göre $\alpha=0,05$ anlamlılık düzeyinde Tablo 5'te $0,05$ 'ten küçük değere sahip girdilerin bağımlı değişken üzerinde etkili olduğu söylenebilmektedir. Eğer parametre ile bağımlı değişken arasında anlamlı ilişki var ise ilişkinin negatif veya pozitif yönlü olup olmadığına β katsayılarına bakılarak karar verilmektedir. 1. k-fold lineer regresyon analizinde $\alpha=0,05$ anlamlılık düzeyinde KU ve N ile ET arasında negatif anlamlı ilişki mevcuttur. TK ve AOGT ile ET arasında ise pozitif anlamlı ilişki görülmektedir. TS ile ET arasında ise anlamlı bir ilişki bulunmamaktadır.

Tablo 6. MLR için parametrelerin VIF değerleri

	VIF							
	Linear				Purequadratic			
	1. k-fold	2. k-fold	3. k-fold	4. k-fold	1. k-fold	2. k-fold	3. k-fold	4. k-fold
KU	2,696	4,747	4,994	7,371	-	4454,601	534,157	1031,599
TK	626,309	649,289	2489,938	1375,084	704,2712	1241,390	1115,479	1964,122
AOGT	326,695	387,098	1194,828	861,578	1273,576	2074,442	-	2129,101
TS	508,383	221,998	684,074	215,858	411,7106	3597,337	1432,863	762,398
N	838,030	175,434	209,704	146,717	-	-	-	-
KUxKU	-	-	-	-	2,931714	4904,335	546,297	860,821
TKxTK	-	-	-	-	-	-	-	-
AOGTxAOG T	-	-	-	-	333,8113	763,313	222,244	777,275
TSxTS	-	-	-	-	-	-	-	-
NxN	-	-	-	-	766,82	1408,207	446,133	346,001

Çoklu doğrusal regresyon modellerinde çoklu doğrusallık problemine karşı varyans şişirme (VIF) değerine bakıldığında literatürdeki sınır değerlerden oldukça yüksek olduğu dikkat çekmektedir (Allison, 1999; Hair vd, 2010)(Tablo 6). Bu nedenle değişkenler arasındaki çoklu doğrusallık probleminin giderilmesi amacıyla temel bileşen analizi (TBA) yapılmaktadır. TBA tekniğinde bileşenler en büyük varyanstan en küçük varyansa doğru sıralanmaktadır. Modelde hangi bileşenlerin dikkate alınacağına ise eigen grafiğinde kırılma çubuğu üzerinde kalan bileşenler belirlenerek karar verilmektedir. Şekil 4'te çalışmada kullanılan bağımsız değişkenler için TBA sonuçlarına göre eigen değerlerinin değişimi görülmektedir. Kırılma çizgisi referans alındığında sadece PC1 bileşeni dikkate alınmaktadır. Tablo 7'de PC1 bileşenini oluşturan bağımsız değişkenler ve katkıları görülmektedir. PC1 bileşenine göre oluşturulan regresyon modellerindeki bağımlı ve bağımsız değişkenler arasındaki ilişki tablo 8 ve 9 da açıklanmaktadır. VIF değerleri tüm modeller için 10 'un altında olduğu için modellerdeki çoklu doğrusallık probleminin üstesinden gelinmektedir (Tablo 9). PC1, ve PC1xPC1 parametreleri ile ET arasında tüm k-foldlar için pozitif anlamlı ilişki olduğu göze çarpmaktadır (Tablo 8).



Şekil 4:
TBA eigen değerlerinin değişimi

Tablo 7. Temel bileşenlerdeki parametrelerin katılım payları

	PC1	PC2	PC3	PC4	PC5
KU	-0,41001	0,91183	0,016293	0,010928	0,0082074
TK	0,45682	0,22015	-0,35432	-0,1186	-0,77669
AOGT	0,45495	0,2089	-0,62053	0,11982	0,59158
TS	0,45619	0,20301	0,47429	-0,69239	0,21522
N	0,4562	0,18773	0,51399	0,70146	-0,020054

Tablo 8. PCA uygulanmış MLR analizinde bağımlı ve bağımsız değişkenler arasındaki ilişki

		β katsayıları ve t testi sonuçları ($\alpha=0,05$ için p değerleri)															
		Linear								Purequadratic							
		1. k-fold		2. k-fold		3. k-fold		4. k-fold		1. k-fold		2. k-fold		3. k-fold		4. k-fold	
		standardize edilmiş β	p	standardize edilmiş β	p	standardize edilmiş β	p	standardize edilmiş β	p	standardize edilmiş β	p	standardize edilmiş β	p	standardize edilmiş β	p	standardize edilmiş β	p
sabit			0,000		0,000		0,000		0,000		0,000		0,000		0,000		0,000
PC1		0,987	0,000	0,981	0,000	0,990	0,000	0,985	0,000	0,862	0,000	0,916	0,000	0,939	0,000	0,891	0,000
PC1xPC1										0,147	0,083	0,159	0,002	0,104	0,011	0,173	0,000

Tablo 9. PCA uygulanmış MLR için parametrelerin VIF değerleri

		VIF							
		Linear				Purequadratic			
		1. k-fold	2. k-fold	3. k-fold	4. k-fold	1. k-fold	2. k-fold	3. k-fold	4. k-fold
PC1		1,000	1,000	1,000	1,000	3,718	1,204	1,322	1,419
PC1xPC1						3,718	1,204	1,322	1,419

PC1 bileşenine göre regresyon analizlerinin performans sonuçları Tablo 10’da verilmektedir. HKO test değerleri dikkate alındığında en iyi model 2. k-fold da çıkmaktadır.

Tablo 10. MLR analizi performans sonuçları

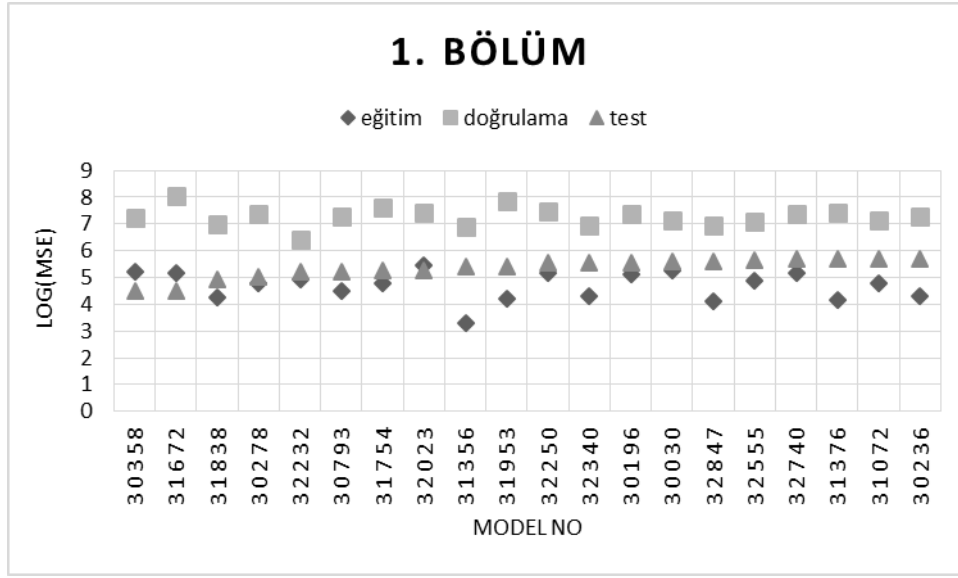
	Reğitim	Rtest	HKOeğitim	HKOtest	OYHeğitim (%)	OYHtest (%)
1. k-fold (Lineer)	0,9873	0,9994	552334	3633581	3,93	18,12
1. k-fold (Purequadratic)	0,9902	0,9995	425867	308426	3,37	3,85
2. k-fold (Lineer)	0,9813	0,9823	1012173	54084	6,08	1,63
2. k-fold (Purequadratic)	0,9920	0,9864	437106	154001	3,75	2,16
3. k-fold (Lineer)	0,9904	0,9961	553531	2205242	3,86	10,78
3. k-fold (Purequadratic)	0,9945	0,9699	316743	828734	2,85	6,10
4. k-fold (Lineer)	0,9846	0,9247	802153	978305	5,73	3,69
4. k-fold (Purequadratic)	0,9952	0,9314	250083	874360	3,15	4,44

YSA ve ANFIS analizinde veri seti 11+4+4 olmak üzere sırasıyla eğitim, test, doğrulama olarak ayrılmaktadır. Veri setinde hangi veri yılının hangi grupta olacağına k-fold analizi ile karar verilmektedir. k-fold analizinde üç bölüm için 6 farklı veri seti dizilimi Tablo11’de verilmektedir (Murat ve Ceylan, 2006).

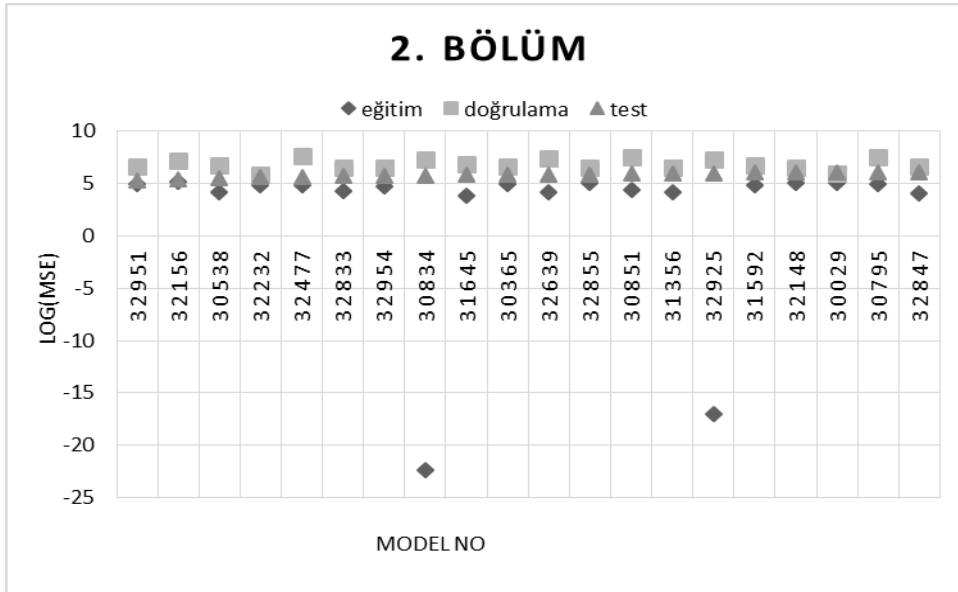
Tablo 11. K-fold analizi için verilerin muhtemel fold dizilimleri

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	
1. k-fold	EĞİTİM										TEST			DOĞRULAMA						
2. k-fold	EĞİTİM										DOĞRULAMA			TEST						
3. k-fold	TEST			DOĞRULAMA			EĞİTİM													
4. k-fold	TEST			EĞİTİM														DOĞRULAMA		
5. k-fold	DOĞRULAMA			EĞİTİM														TEST		
6. k-fold	DOĞRULAMA			TEST			EĞİTİM													

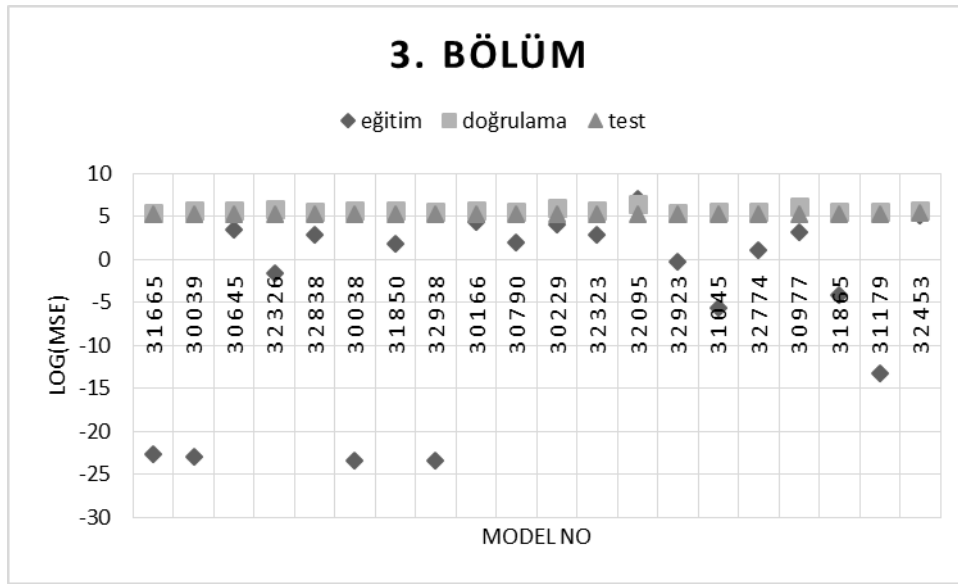
YSA analizinde her bir fold için 1’den 20’ye kadar nöron sayısı ağda denenmekte ve ağdaki ağırlıklar yenilenecek her bir fold için 3000 ağ modeli oluşturulmaktadır. Model performansları HKO kriteri baz alınarak incelenmektedir. Test performansına göre en iyi 20 YSA modelinin karşılaştırılması Şekil 5-10’da her bir bölüm için verilmektedir. Ayrıca HKO değerlerinin logaritması alınarak LOG(MSE) şeklinde grafiklerde ifade edilmektedir. Bu değerlerin birbirine yakın olması modelin performansının yüksek olduğunu göstermektedir.



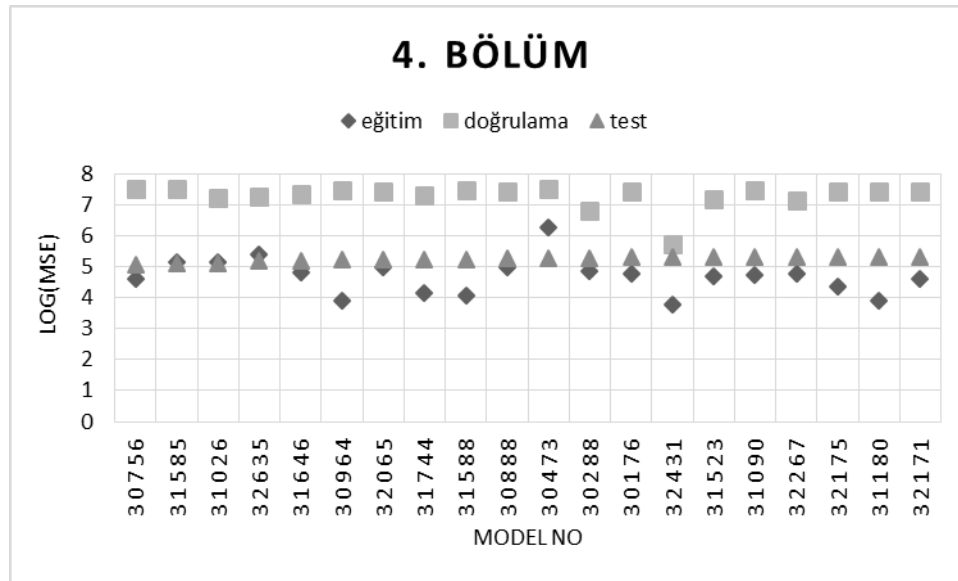
Şekil 5:
1. k-fold için en iyi YSA modellerin karşılaştırılması



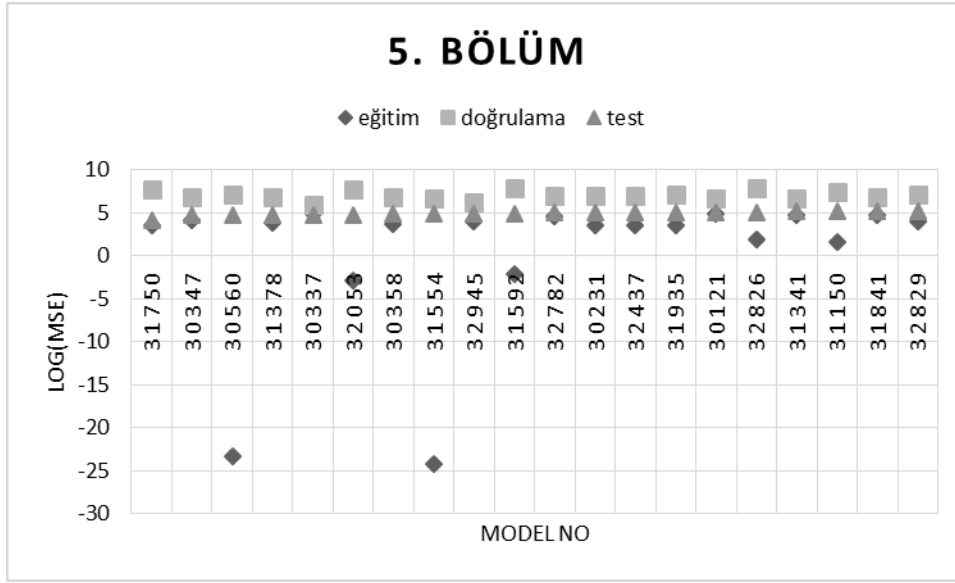
Şekil 6:
2. k-fold için en iyi YSA modellerin karşılaştırılması



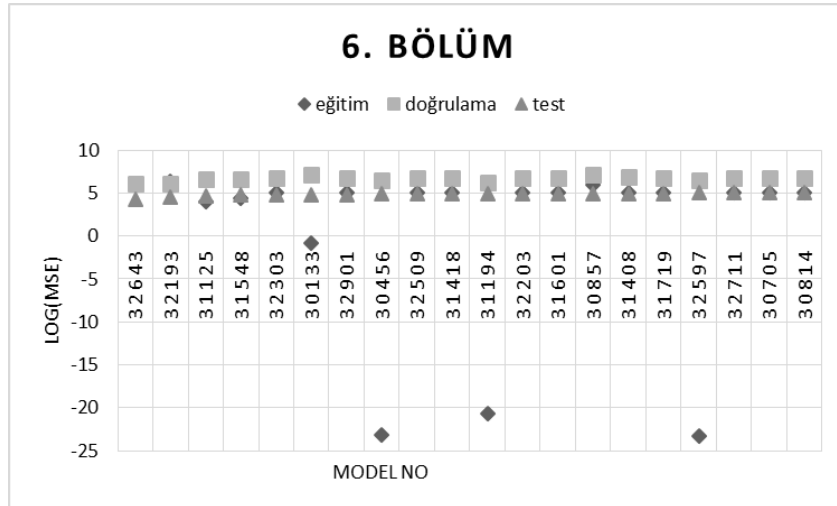
Şekil 7:
3. k-fold için en iyi YSA modellerin karşılaştırılması



Şekil 8:
4. k-fold için en iyi YSA modellerin karşılaştırılması



Şekil 9:
5. k-fold için en iyi YSA modellerin karşılaştırılması

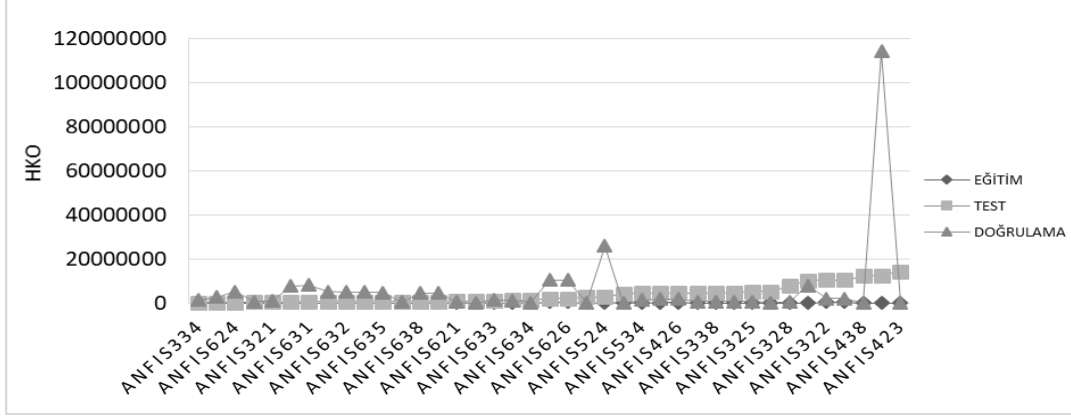


Şekil 10:
6. k-fold için en iyi YSA modellerin karşılaştırılması

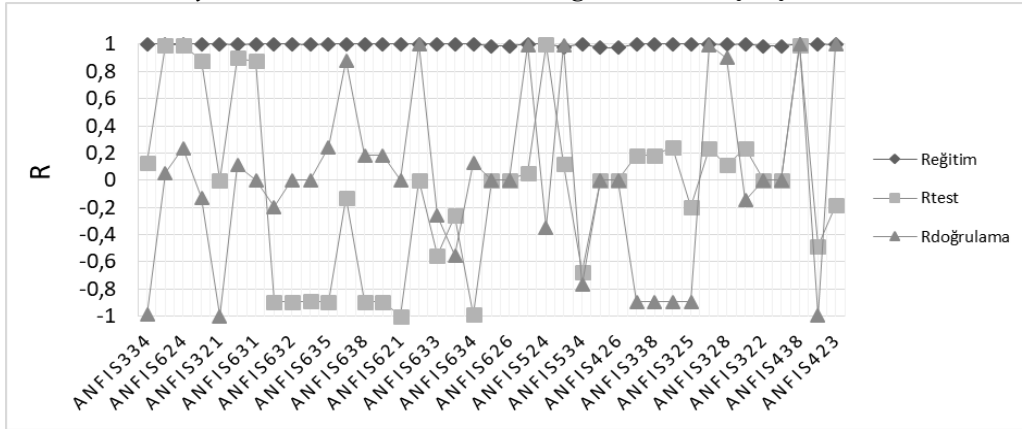
Şekil 5-10 grafikler incelendiğinde, hem test HKO değeri düşük hem de eğitim ve test HKO performansları birbirine en yakın model arandığında en iyi model 1. k-fold da 31838 çıkmaktadır. En iyi performanslı ağ 18 nöronlu, iki gizli katmana sahip ağ modelidir. Böylelikle en iyi veri seti dizilimi ilk 11 yılın eğitim, 2011-2014 yılları arasının test, 2015- 2018 yılları arasının doğrulama datası olan 2. k-fold seçilmektedir.

ANFIS modellerinde 6 k-fold için analiz yapılmaktadır. 5 bağımsız değişken için 2 ve 3 üyelik fonksiyonu kullanılmaktadır. Üyelik fonksiyonu tipi trimf, trapmf, gbellmf, gaussmf, gauss2mf, pimf, dsigmf ve psigmf kullanılmaktadır. Modellerde üç haneli isimlendirmede birinci hane k-fold numarasını, ikinci hane üyelik fonksiyonu sayısını, üçüncü hane ise üyelik fonksiyonu tipini temsil etmektedir. Modellerin HKO değerlerine göre performans değerlendirmesi şekil 11’de, R değerlerinin ise karşılaştırılması şekil 12’de verilmektedir. R değerlerine göre eğitim,

test ve doğrulama bölümlerine göre R değerleri incelendiğinde en iyi model ANFIS438 nolu model çıkmaktadır. Modelde 4. k-fold, her bağımsız değişken için üç üyelik fonksiyonu ve psigmf üyelik fonksiyonu tipi kullanılmaktadır. ANFIS438 modelinin eğitim, test ve doğrulama için HKO değerleri sırasıyla 608, 12352001, 7 dir. Modelin eğitim, test ve doğrulama için R değerleri sırasıyla %99,99, %98,84 ve %99,99'dur.



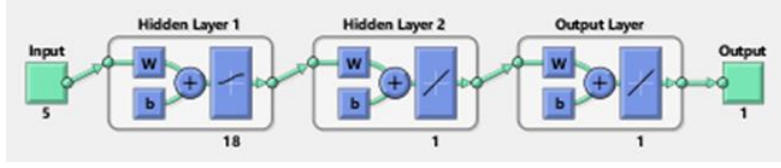
Şekil 11:
En iyi 40 ANFIS modelinin HKO değerlerinin karşılaştırılması



Şekil 12:
En iyi 40 ANFIS modelinin R değerlerinin karşılaştırılması

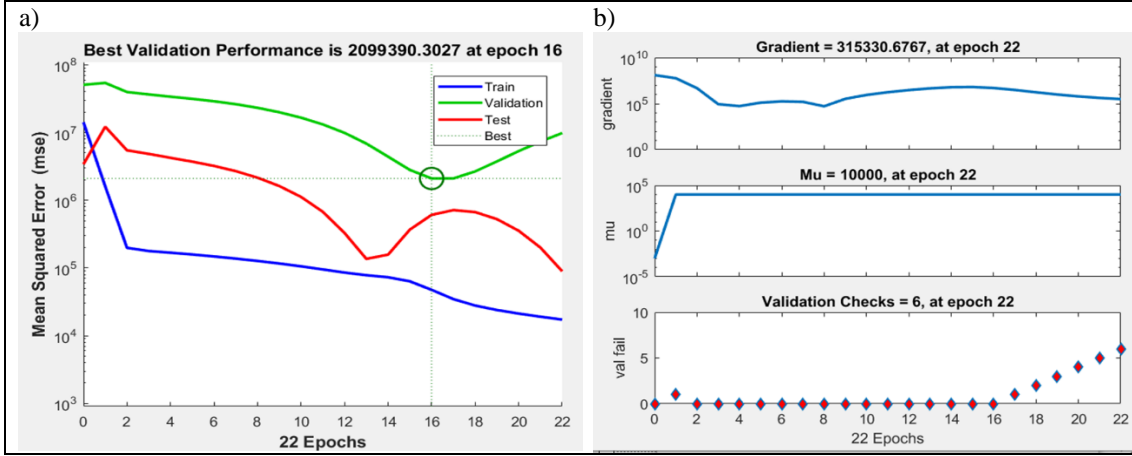
3. BULGULAR VE TARTIŞMA

YSA analizleri sonucunda iki gizli katmana sahip, birinci katmanda 18 adet nöronu, ikinci katmanda ise 1 adet nöronu bulunan model en iyi sonuçları vermektedir. Bu modelin birinci ve ikinci gizli katmanında sırasıyla logsig ve purelin transfer fonksiyonu kullanılmaktadır. Ayrıca bu modelin öğrenim algoritması Levenberg-Marquardt olarak belirlenmektedir.

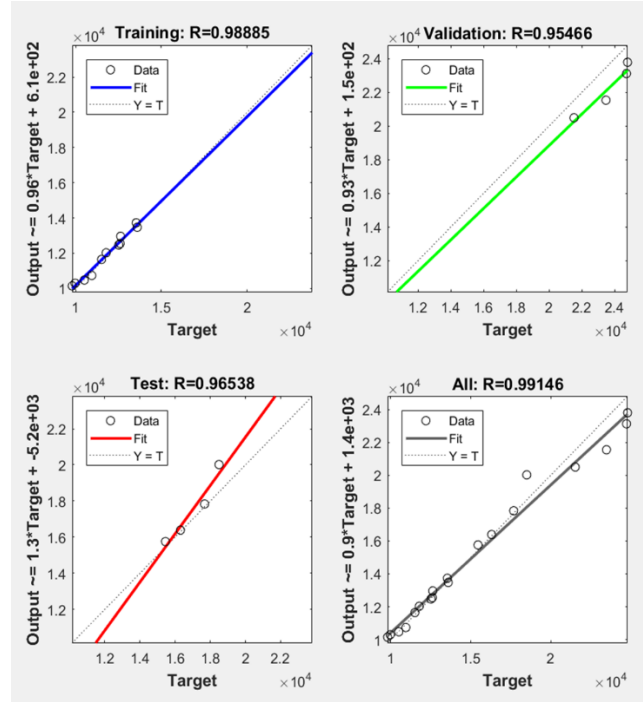


Şekil 13:
En iyi performansa sahip YSA ağ modeli

En iyi performansa sahip YSA modelinin ağ yapısı, HKO, eğitim süreci ve korelasyon grafikleri sırasıyla Şekil 13, Şekil 14, Şekil 15’de verilmektedir.



Şekil 14:
a. YSA modelinin HKO değerinin değişimi b. YSA modelinin eğitim süreci



Şekil 15:
Modelin korelasyon katsayı grafiği

Çok değişkenli regresyon analizleri sonucunda ET, KU, TK, AOGT, TS, ve N değişkenleri ile lineer regresyon ve purequadratic regresyon denklemleri oluşturulmaktadır. Fakat modellerde çoklu doğrusallık problemi görülmesi üzerine TBA analizi yapılmaktadır. TBA analizine göre yapılan lineer ve purequadratic regresyon modellemesinde denklemler sırasıyla Denklem 5 ve Denklem 6'da verilmektedir.

$$ET = 15390,133 + 2272,881 * PC1 \quad (5)$$

$$ET = 14519,021 + 2120,968 * PC1 + 177,387 * (PC1)^2 \quad (6)$$

Denklem 5 ve 6'ya bakıldığında PC1 ve ET arasında pozitif ilişki olduğu görülmektedir. Yapılan analizler sonucunda en iyi YSA, ANFIS ve regresyon modellerinin eğitim, test ve doğrulama performans karşılaştırması Tablo 12'de sunulmaktadır. Dört tekniğe göre oluşturulan modellerde en iyi performans 54084 HKO değeri ile lineer regresyon modelinde görülmektedir. En düşük performans ise 12352001 HKO ile ANFIS tekniğinde gözlenmektedir.

Tablo 12. Modellerin Eğitim Performansı Karşılaştırmaları

Model	HKO			OYH (%)			Düzeltilmiş R ²		
	Eğitim	Test	Doğrulama	Eğitim	Test	Doğrulama	Eğitim	Test	Doğrulama
YSA	47455	608475	2099390	1,64	2,89	5,89	0,9559	0,8980	0,8672
ANFIS	608	12352001	7	0,10	33,70	0,01	0,9996	0,9654	0,9997
Lineer Regresyon	1012173	54084	-	2,07	4,86	-	0,9601	0,9474	-
Purequadratic Regresyon	437106	154001	-	1,74	3,73	-	0,9814	0,9190	-

4. SONUÇ

Yapılan bu çalışmada karayolu uzunluğu, taşıt kilometre, AOGT, taşıt sayısı ve nüfus değişkenleri kullanılarak karayollarında enerji tüketim miktarı tahmin modelleri geliştirilmektedir. Modeller YSA, ANFIS, çok değişkenli lineer regresyon ve çok değişkenli purequadratic regresyon metotları ile oluşturulmaktadır.

Çalışmada 19 yıllık veri seti kullanılmaktadır. Analizlerde eğitim ve test verileri dışında az sayıda da olsa doğrulama verileri de kullanılmaktadır. Bu sayede modelin performansının arttığı gözlemlenmiştir.

Çalışmada YSA ve ANFIS modellerinin performansı, regresyon modellerinin gerisinde kalmaktadır. ANFIS modelinin çıktılar gerçek ET değerleri ile izlediği uyum her ne kadar regresyon modellerine göre üstün olsa da gerçek değerler ile hata payı yüksektir. Bu da ANFIS modelini performans bakımından geri plana atmaktadır.

Dört farklı tahmin tekniği kullanılarak geliştirilen ET tahmin modelleri arasında en iyi model lineer regresyon çıkmaktadır. Lineer regresyon modelinin HKO, OYH ve R performans değerleri sırasıyla 54084, %4,86 ve 0,9474 çıkmaktadır. En iyi ET modeli olan lineer regresyon modeli düzeltilmiş belirlilik katsayısı 0,9474 ile lineer doğrusallık göstermektedir. Geliştirilen regresyon modeli ile gelecekte karayolu enerji tüketiminin tahmini ve ulaşım planlamalarına ışık tutacağı açık bir şekilde görülmektedir.

KAYNAKLAR

1. Allison, P. D. (1999). Multiple regression: A primer. Pine Forge Press.
2. Bayazıt, M. and Oğuz, E. B. Y. (2005) *Mühendisler için istatistik*, Birsen Yayınevi, Türkiye.

3. Cansız, Ö. F. (2007) Enerji politikalarının ulaştırma sistemlerinin optimizasyonu ile geliştirilmesi ve uygulamadan elde edilen getirilerin ortaya konması, *Doktora Tezi*, Gazi Üniversitesi Fen Bilimleri Enstitüsü, Ankara.
4. Cansız, Ö. F., Erginer, İ. ve Genç, G. G. (2018) Ulaştırma sektöründe karayollarının payına düşen enerji tüketiminin yapay sinir ağları ve çok değişkenli lineer regresyon yöntemleri ile tahmini, *International Eurasian Conference on Science, Engineering and Technology*, Ankara, 627-633.
5. Cansiz, O. F. and Easa, S. M. (2011) Using artificial neural network to predict collisions on horizontal tangents of 3D two-lane highways, *International Journal of Engineering and Applied Sciences*, 7(1), 47-56.
6. Ergezer, H., Dikmen, M. and Özdemir, E. (2003) Yapay sinir ağları ve tanıma sistemleri, *Pivolka*, 2(6), 14-17.
7. Es, H., Kalender, F. and Hamzaçebi, C. (2014) Yapay sinir ağları ile Türkiye net enerji talep tahmini, *Gazi Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi*, 29(3), 495-504.
8. Generalized bell-shaped membership function . Erişim Adresi: <https://www.mathworks.com/help/fuzzy/gbellmf.html> (Erişim Tarihi: 05/01/2020)
9. Gündüz, A., Kaya, M. and Aydemir, C. (2011) Kentiçi ulaşımında karayolu ulaşımına alternatif sistem: raylı ulaşım sistemi, *Akademik Yaklaşımlar Dergisi*, 2(1), 134-151.
10. Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E. (2010) *Multivariate Data Analysis*, 7th Edition, NJ: Prentice Hall.
11. Haldenbilen, S. (2003) Genetik algoritma yaklaşımı ile Türkiye için sürdürülebilir ulaştırma göstergelerinin analizi ve planlaması, *Doktora Tezi*, Pamukkale Üniversitesi Fen Bilimleri Enstitüsü, Denizli.
12. Jang JSR, Sun CT, Mizutani E. (1997). Neurofuzzy and soft computing: a computational approach to learning and machine intelligence [Book Review], *IEEE Transactions on automatic control*, 42(10), 1482-1484
13. Jang, J. S. R. (1991), Fuzzy modeling using generalized neural networks and kalman filter algorithm, In *AAAI*, 91, 762-767.
14. Mucuk, M. and Uysal, D. (2009) Türkiye ekonomisinde enerji tüketimi ve ekonomik büyüme, *Maliye Dergisi*, 157, 105-115.
15. Murat, Y.S. (2006). Comparison of fuzzy logic and artificial neural networks approaches in vehicle delay modeling. *Transportation Research Part C: Emerging Technologies*, 14(5), 316-334.
16. Murat, Y.S., Ceylan, H. (2006) Use of artificial neural networks for transport energy demand modeling. *Energy policy*, 34(17), 3165-3172.
17. Öztürk, Z. (2018) İstanbul'da karayolu yolcu taşımacılığında elektrikli araç kullanımının incelenmesi, *El-Cezeri Journal of Science and Engineering*, 5(2), 367-386. Doi: 10.31202/ecjse.382348
18. Solak, A. O. (2013) Türkiye'de ulaştırma sektörü enerji tüketiminin azaltılması: bir senaryo yaklaşımı, *International Journal of Economic & Social Research*, 9(1), 125-141.

