



Sezgisel Regresyon Teknikleri ile Sayısal Yükseklik Modellenmesi

Vahdettin Demir^{1*}, Esra Aslı Çubukçu²

^{1*} KTO Karatay Üniversitesi, Mühendislik Fakültesi, İnşaat Bölümü, Konya, Türkiye, (ORCID: 0000-0002-6590-5658), vahdettin.demir@karatay.edu.tr

² KTO Karatay Üniversitesi, Mühendislik Fakültesi, İnşaat Bölümü, Konya, Türkiye, (ORCID: 0000-0003-4159-205X), cubukcuasli@gmail.com

(2nd International Conference on Access to Recent Advances in Engineering and Digitalization (ARACONF)-10–12 March 2021)

(DOI: 10.31590/ejosat.916012)

ATIF/REFERENCE: Demir, V. & Çubukçu, E. A. (2021) Sezgisel Regresyon Teknikleri ile Sayısal Yükseklik Modellenmesi. Avrupa Bilim ve Teknoloji Dergisi, (24), 484-488.

Öz

Bu çalışmada, Samsun Mert Irmağı Havzası'nda, hâlihazır haritalardan temin edilmiş yatay ve düşey koordinat bilgilerinin yer aldığı noktalardan yükseklik değerleri, M5 model ağacı (M5-tree) ve çok değişkenli uyarlamalı regresyon eğrileri (MARS) sezgisel regresyon yöntemleri kullanılarak tahmin edilmeye çalışılmıştır. Sonuçlar Tek ve Çok Değişkenli Regresyon (TDR-ÇDR) yöntemiyle karşılaştırılmıştır. Çalışmada 3 farklı giriş senaryosu incelenmiştir. Bunlar: X yönündeki koordinat bilgisiyle yükseklik tahmini (i); Y yönündeki koordinat bilgisiyle yükseklik tahmini (ii); X ve Y koordinat bilgisiyle yükseklik tahmini şeklindedir (iii). Karşılaştırma kriterleri olarak determinasyon katsayısı (R^2), Ortalama Mutlak Hata (OMH) ve Karekök Ortalama Karesel Hata (KOKH) kullanılmıştır. Modelleme sonuçları incelendiğinde; (1) M5-tree regresyon yönteminin en iyi sonucu verdiği, (2) MARS yöntemi ÇDR ve TDR yöntemlerine göre daha iyi olduğu, (3) En kötü sonuç TDR yöntemi kullanılarak yapılan doğrusal regresyon modellemesinde tespit edilmiştir. (4) Özellikle, M5-tree sezgisel regresyon yönteminin yükseklik modellemesinde oldukça başarılı bir metot olduğu sonucuna ulaşılmıştır.

Anahtar Kelimeler: Sayısal Yükseklik Modeli, M5-tree, MARS, Samsun.

Digital Elevation Modeling with Heuristic Regression Techniques

Abstract

In this study, elevation values of the Mert River Basin of Samsun were estimated by M5 model tree (M5-tree) and Multivariate Adaptive Regression Splines (MARS) using base map which contains horizontal and vertical location informations. Results were compared with univariate and multivariate linear regression methods (MLR). In the study, three different input scenarios were tried as: (i) elevation forecast with X coordinate information, (ii) elevation forecast with Y coordinate information, (iii) elevation forecast with X and Y coordinate information. Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and coefficient of determination (R^2) were utilized as comparison criteria. According to modeling results: (1) M5-tree regression method provided the best results, (2) MARS method was more suitable than the univariate and multivariate linear regression methods, (3) Single variable linear regression method provided the worst estimate, (4) M5-tree regression method can be successfully used in elevation modeling.

Keywords: Digital Elevation Models, M5-tree, MARS, Samsun.

* Sorumlu Yazar: vahdettin.demir@karatay.edu.tr

1. Giriş

Birçok mühendislik uygulamasında, topoğrafik yüzeye ihtiyaç duyulmakta ve topografyanın uygun doğrulukta belirlenmesi gerekmektedir. Bu amaçla Sayısal Yükseklik Modeline (SYM) dayalı uygulamalarda olduğu gibi, yatay ve düşey konum bilgileri kullanılarak yükseklik değerleri uygun yöntemlerle belirlenebilir (Akçın, Kutoğlu, & Terlemezoğlu, 2005). İster arazi üzerinden alım isterse uzaktan algılama yöntemlerine göre olsun yükseklik bilgisinin belirlenmesi mühendislikte her zaman güç bir konu olmuştur (Demir & Ülke, 2020). Bu güçlü bilinen fonksiyonların dışında son zamanlarda sıklıkla kullanılan sezgisel regresyon yaklaşımlarıyla modellenen birliği bu çalışmanın amacını oluşturmaktadır.

Su kaynaklarının yönetimi, kentsel ve kırsal planlama, ulaştırma planlaması, tarım, ormancılık, havza yönetimi, afet risk değerlendirmesi vb. birçok alanda arazi noktalarından üretilmiş sayısal modeller kullanılmaktadır. Bu modeller özellikle mühendislik çalışmalarının temel proje altlığını oluşturmaktadır. Sayısal yüzey modelleri ile yer altı ve yüzey hakkında bilgiler işlenebilir, analiz edilebilir ve görsel olarak sunulabilir. Ancak, arazideki tüm noktaların yüksekliklerini ölçmek zor ve yüksek maliyetlidir. Bu nedenle arazi noktalarının matematiksel fonksiyonlar kullanılarak yoğunlaştırması veya ara noktaların tahmin edilmesi hızlı ve ücretsiz olarak sağlanabilir. Bu çalışmada yükseklik modellemesi, M5 model ağacı (M5-tree) ve çok değişkenli uyarlamalı regresyon eğrileri (MARS) sezgisel regresyon yöntemleri kullanılarak incelenmiştir. Sonuçlar regresyon yöntemiyle karşılaştırılmıştır. MARS, 1991 yılında J.H. Friedman tarafından tanıtılan bir regresyon analizi türüdür. Bu yöntem doğrusal modellerin bir uzantısı olan parametrik olmayan regresyon tekniklerindedir. M5-tree ise 1992 yılında J. R. Quinlan tarafından ilk kez kullanılan ve ikili karar ağaçlarına dayanan sezgisel regresyon yaklaşımlarındandır. Gerçek bir ağaca yapısal olarak benzeyen M5-tree yöntemi, "IF-THEN" türündeki kurallarla modellenen değişkeni tahmin etmektedir. MARS ve M5-tree birçok mekanizma ve sistemin çıktılarını hesaplamak için akademisyenler tarafından farklı disiplinlerde (meteoroloji, hidroloji, kıyı mühendisliği, finans, bankacılık, balıkçılık vb.) başarılı bir şekilde kullanılmaktadır (Abolfathi, Yeganeh-Bakhtiary, Hamze-Ziabari, & Borzooei, 2016; Adamowski, Chan, Prasher, & Sharda, 2012; Bhattacharya & Solomatine, 2005; De Andrés, Lorca, De Cos Juez, & Sánchez-Lasheras, 2011; Elith & Leathwick, 2007; Etemad-Shahidi & Mahjoobi, 2009; Leathwick, Rowe, Richardson, Elith, & Hastie, 2005; Lee, Chiu, Chou, & Lu, 2006; Put, Xu, Massart, & Vander Heyden, 2004; W. G. Zhang & Goh, 2013; W. Zhang & Goh, 2014). Fakat bu yöntemler ilk defa sayısal yükseklik modellemesinde kullanılacaktır. Bu çalışmada modelleme performansı 3 farklı giriş senaryosu için denenmiştir. Bunlar: (i) X koordinat bilgisiyyle yükseklik tahmini; (ii) Y koordinat bilgisiyyle yükseklik tahmini; (iii) X ve Y koordinat bilgisiyyle yükseklik tahmini şeklindedir.

2. Materyal ve Metot

2.1. Kullanılan Veri ve Çalışma Alanı

Topoğrafik yüzeyin tahmin edilmesi modelleme çalışmalarının ilk aşamasıdır. Bu modelleme ile yer altı ve yüzey hakkında bilgiler işlenebilir, analiz edilebilir ve görsel olarak sunulabilir. Bu nedenle çalışmamızda Samsun ili Mert Irmağı Havzası ve kent sınırlarının keşiştiği alanda, hâlihazır haritalardan

temin edilmiş yersel noktalardan kot değerleri tahmin edilmeye çalışılmıştır (Şekil 1).



Şekil 1. Çalışma Alanı

Çalışma verilerine ait istatistik bilgileri Tablo 1'de verilmiştir

Tablo 1. Tanımlayıcı istatistik bilgileri

Değişken (m)	Veri sayısı	Maksimum değeri (m)	Minimum değeri (m)	Ortalama (m)	Standart sapma	Çarpıklık
X koordinat	14362	4572051.15	4568584.58	4570318	889.7	0.008
Y koordinat	14362	530920.69	528296.48	52955	686.6	0.057
Z Kot	14362	238.13	0.07	43.68	52.35	1.429

2.2. Metot

2.2.1 M5-Tree

M5 model ağacı algoritması Quinlan tarafından 1992 yılında geliştirilen yeni bir regresyon yöntemidir (Quinlan, 1992). İki bileşenli karar ağacı bu modelin belkemiğidir. Son yaprak düğümlerine (terminal leaf nodes) uygulanan lineer regresyon fonksiyonu bağımsız ve bağımlı değişkenler arasındaki ilişkiyi tanımlar. M5 model ağacı, kategorik veri için kullanılan diğer karar ağaç modellerinden daha iyidir (Mitchell, 1997). Model aynı zamanda sayısal verilere odaklanır. Bu özellik, diğer modellerle kıyaslandığında modelin önemini artırır. M5 model ağacı iki aşamalı bir modeldir. İlk aşamasında, veri karar şemasını (ağacını) üretmek için alt setlere ayrılır. Bir düğüme ulaşılan sınıf değerinin standart sapması, kriterlere ayırmak için kullanılır. Bu düğüme etki eden elemanların test edilmesinin sonucu olarak ortaya çıkan hataya bağlı olarak beklenen indirgeme hesaplanır (Pal & Deswal, 2009; Solomatine & Xue, 2004). Standart sapma indirgemesinin (SDR) formülasyonu aşağıdaki gibidir.

$$SDR = sd(T) - \sum \frac{|T_i|}{|T|} sd(T_i) \quad (1)$$

Bu formülde sd standart sapma olarak ifade edilir. T düğüme etki eden bir dizi örneği simgeler. Potansiyel verilerin i inci sonuçlarına sahip olan alt küme örnekleri T_i tarafından simgelenir. Gruplara ayrılma sürecinde alt düğümler verilerin standart sapmasında üst düğümlerden daha azdır. Her bir alt grup kontrol edildikten sonra öngörülen hatayı en çok azaltan tercih edilir. Ancak bu aşama genellikle geniş ağaçsı bir tasarım sunar. İkinci aşamasında ise büyümüş olan ağaç budanır. Ve sonrasında budanmış alt ağaçlar genellikle lineer regresyon fonksiyonları ile

değiştirilir (Quinlan, 1992). Bu çalışmada M5 model ağacı sayısal yükseklik modellemesinden kullanılmıştır.

2.2.2. MARS

Türkçesi çok değişkenli uyarlanabilir regresyon eğrileri olan MARS, lineer olmayan sürekli sayısal sonuçları tahmin etmek için kullanılan bir modeldir. Model, tahmin yöntemi ve bağımlı değişkenler arasındaki karmaşık doğrusal olmayan ilişkiyi açıklar. MARS algoritması ileriye ve geriye doğru olmak üzere iki adımdan oluşmaktadır. İleriye doğru adım algoritması ile bir dizi uygun giriş değişkenlerini seçer (De Andrés et al., 2011). Geriye doğru adım algoritması ile de önceden seçilmiş sette var olan gereksiz değişkenleri elimine eder. Bu yöntem aynı zamanda tahminlerin doğruluğunu artırır. İki temel fonksiyon veya giriş aralığı boyunca sapma noktasında tanımlanan her iki değişken değeri tarafından X değişkeninden yeni değişken Y'ye fonksiyon çizilir (Sharda, Patel, Prasher, Ojasvi, & Prakash, 2006).

$$Y = \max(0, X - c) \quad (2)$$

$$Y = \max(0, c - x) \quad (3)$$

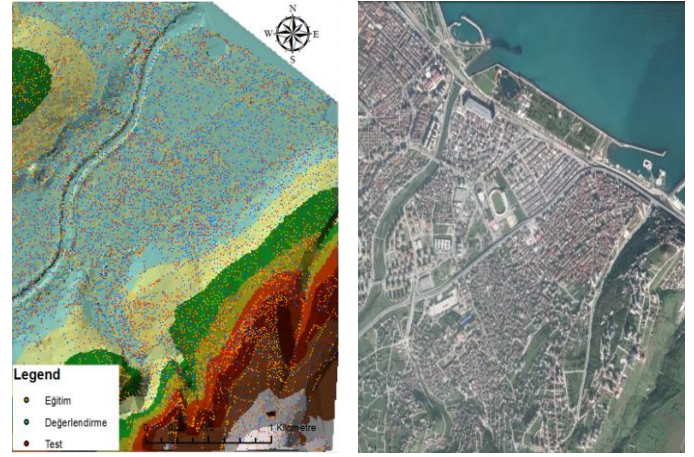
Burada c eşik (alt sınır) değerini temsil eder. MARS modeli özellikle mali işler yönetim sistemi, zaman serisi verisinde ve birçok alanda kullanılmaktadır (P. Bera et al., 2006; Sephton, 2001). Fakat yöntem topoğrafik amaçlı herhangi bir modelleme çalışmasında kullanılmamıştır.

2.2.3. Regresyon Analizi

Yapılan araştırmalar sonucunda elde edilen verilere göre değişkenler arasında ilişkiyi belirlemek için; bağımsız değişkenlere dayanarak bağımlı değişkeni tahmin etmek için regresyon analizi kullanılır (Dipova & Cangir, 2010). Birden fazla bağımsız değişken kullanılarak yapılan analize çok değişkenli regresyon analizi (ÇDR), tek bir bağımsız değişken kullanılarak yapılan analize ise tek değişkenli regresyon analizi denir (TDR) (Demir & Ülke, 2020).

3. Araştırma Sonuçları ve Tartışma

Çalışmada regresyon yöntemlerinden M5-tree ve MARS yöntemleri kullanılarak yükseklik değerleri modellenmiştir. Giriş verileri olarak X ve Y (ED-50) koordinat verileri kullanılmıştır. Modelleme üç aşamada (eğitim, değerlendirme ve test) gerçekleştirilmiştir. Eğitim aşamasında verilerin %50'si, değerlendirme ve test aşamasında ise verilerin %25'i kullanılmıştır. Modelleme çalışmalarda genellikle verilerin %50'si eğitimde kalan %25'i de değerlendirme ve test aşamasında kullanılmaktadır (Pourghasemi & Rossi, 2019). Eğitim, değerlendirme ve test verilerinin rastgele dağılımı Şekil 2'te, istatistik bilgileri ise Tablo 2'de yer almaktadır. Yükseklik değerlerinin çok değişken olduğu tablodan açıkça görülmektedir.



Şekil 2. Eğitim, değerlendirme ve test verilerinin dağılımı

Yükseklik modellemesinde (Z kot) üç farklı giriş durumu denenmiş olup bunlar: (i) X koordinat bilgileri; (ii) Y koordinat bilgileri; (iii) X ve Y koordinat bilgileri şeklindedir. M5-tree ve MARS modellemeleri ayrıca TDR ve ÇDR ile de karşılaştırılmıştır. Değerlendirme kriterleri olarak KOKH (Karekök Ortalama Karesel Hata), OMH (Ortalama Mutlak Hata) ve R² (Determinasyon katsayısı) kullanılmıştır (Demir, Bilge, & Bektaş, 2017; Demir, Çıtakoğlu, Geyikli & Kuyucu 2017). KOKH ve OMH formülleri aşağıdaki gibi ifade edilebilir (Denklem 4, 5).

$$KOKH = \sqrt{\frac{1}{N} \sum_{i=1}^N (Z_t - Z_g)^2} \quad (4)$$

$$OMH = \frac{1}{N} \sum_{i=1}^N |Z_t - Z_g| \quad (5)$$

$$R^2 = \left(\frac{N * (\sum Z_g * Z_t) - (\sum Z_g) * (\sum Z_t)}{\sqrt{((N * \sum Z_g^2) - (\sum Z_g)^2) * ((N * \sum Z_t^2) - (\sum Z_t)^2)}} \right)^2 \quad (6)$$

Yukardaki eşitliklerde Z_t ve Z_g tahmin edilen ve gözlenen Z değerlerini, N ise veri sayısını göstermektedir.

Tablo 2. Eğitim, değerlendirme ve test veri setlerine ait bilgiler

Veri Seti	Değişken	Veri sayısı	Maksimum değeri (m)	Minimum değeri (m)	Ortalama (m)	Standart sapma	Çarpıklık
Eğitim	X	7182	4572051.1	4568584.5	4570314.8	888.84	0.010
	Y	7182	530920.69	528298.59	529544.74	683.08	0.077
	Z	7182	237.38	0.07	43.41	52.04	1.436
Değerl.	X	3590	4572048.4	4568585.2	4570329.7	894.21	0.002
	Y	3590	530918.31	528298.77	529578.42	685.23	0.025
	Z	3590	238.13	0.07	43.80	52.43	1.423
Test	X	3590	4572049.8	4568591.2	4570314.5	892.78	0.006
	Y	3590	530919.13	528296.48	529561.80	689.89	0.040
	Z	3590	236.00	0.07	44.12	52.88	1.422

M5-tree yöntemiyle 3 farklı giriş durumu modellenmiş ve Tablo 3'te eğitim, değerlendirme ve test aşamalarında elde edilen sonuçlar verilmiştir. Tablodan açıkça görüldüğü gibi X ve Y girişini kullanan M5-tree modeli test aşamasında diğerlerine göre daha düşük KOKH (3.54 m), OMH (2.04 m) ve daha büyük R²

(0.995) değerleri vermiştir. En kötü sonuç 1'inci durumda (sadece X koordinat bilgisi kullanılarak) yapılan tahminde elde edilmiştir.

Tablo 3. M5-tree modellerinin eğitim, değerlendirme ve test aşamasındaki hataları

Veri Seti	Model	Giriş	KOKH (m)	OMH (m)	R ²
Eğitim	(i)	X	40.34	28.67	0.399
	(ii)	Y	28.78	19.51	0.694
	(iii)	X,Y	2.49	1.50	0.997
Değerlendirme	(i)	X	51.62	37.81	0.113
	(ii)	Y	36.58	25.31	0.523
	(iii)	X,Y	3.54	2.04	0.996
Test	(i)	X	51.62	37.82	0.101
	(ii)	Y	36.58	25.31	0.538
	(iii)	X,Y	3.54	2.04	0.995

Tablo 4. MARS modellerinin eğitim, değerlendirme ve test aşamasındaki hataları

Veri Seti	Model	Giriş	KOKH (m)	OMH (m)	R ²
Eğitim	(i)	X	46.23	35.28	0.210
	(ii)	Y	32.78	23.88	0.603
	(iii)	X,Y	24.67	19.65	0.775
Değerlendirme	(i)	X	46.56	35.60	0.211
	(ii)	Y	32.92	24.03	0.606
	(iii)	X,Y	24.60	19.40	0.779
Test	(i)	X	47.01	35.88	0.209
	(ii)	Y	33.34	24.26	0.602
	(iii)	X,Y	24.95	19.90	0.777

MARS yönteminden eğitim, değerlendirme ve test aşamasında elde edilen sonuçlar Tablo 4'te verilmiştir. Tablodan görüldüğü gibi tüm veri setlerinde X ve Y girişini kullanan MARS modeli (iii) diğerlerine göre daha düşük KOKH, OMH ve daha büyük R² değerleri vermiştir. En kötü sonuçlar X koordinat bilgilerinin kullanıldığı (i) giriş modelinden elde edilmiştir.

Tablo 5. TDR ve ÇDR modellerinin test aşamasındaki hataları

Giriş	Katsayılar a, b	KOKH (m)	OMH (m)	R ²
X (TDR)	-0.0398; 181948.2	48.98	38.57	0.142
Y (TDR)	0.0273;-14434.1	38.35	31.30	0.468
X,Y(ÇDR)	-0.0378; 0.0216; 161682.6	46.61	36.80	0.552

TDR ve ÇDR'ye ait test sonuçları Tablo 5'te verilmiştir. En küçük kareler ile elde edilen regresyon katsayıları da tabloda görülmektedir. M5-tree ve MARS sonucuna benzer şekilde en iyi tahmin her iki koordinat bilgisinin kullanıldığı (iii) giriş modelinde elde edilmiştir. Tablo 3-5 karşılaştırıldığında M5-tree modelinin, MARS ve TDR-ÇDR yöntemlerine göre Z değerlerini daha iyi tahmin ettiği görülmektedir.

4. Sonuç

Topografyanın oluşturulması, başka bir ifade ile yatay, düşey koordinatlarının yanı sıra yükseklik değerlerinin uygun doğrulukta belirlenmesi, hidrografik amaçlı mühendislik uygulamalarının çoğunda ihtiyaç duyulan temel işlevlerdendir. Bu çalışmada Samsun Mert Irmağı Karadeniz çıkış havzasının, kentsel sınırlarla kesiştiği bölgede üç farklı sezgisel regresyon yöntemine göre sayısal yüksekliklerin modellenen bilirliliği araştırılmıştır. Kullanılan yöntemler, M5 model tree ve çok değişkenli uyarlamalı regresyon eğrileri yöntemleri ve literatürde sıkça kullanılan doğrusal regresyon yöntemidir. Yöntemlerin performansları incelenmiş ve aşağıdaki sonuçlar elde edilmiştir:

(1) M5-tree, MARS ve doğrusal regresyon yöntemlerine göre Z değerlerini daha iyi tahmin etmektedir,

(2) M5-tree ile yapılan tahminde en iyi sonucu X ve Y koordinat bilgileri (iii) kullanılarak yapılan modelleme vermiştir,

(3) MARS modellemesi regresyon modellemelerine göre Z değerlerini daha iyi tahmin etmektedir,

(4) ÇDR yöntemi ise, TDR yöntemine göre daha başarılı sonuçlar vermiştir.

Bu çalışma ile yüzey modellemesinin modern regresyon yöntemleri yardımıyla elde edilebileceği görülmüştür. Elde edilen sonuçlara incelendiğinde modern regresyon modellemeleri ÇDR ve TDR gibi regresyon modellemesinden daha başarılıdır. Bu sonuçlar modern regresyon tekniklerinin hidrolik modellemeler gibi birçok mühendislik çalışmalarının temel altlığını oluşturan yükseklik modellemelerinde kullanılabilir olduğunu göstermektedir.

5. Teşekkür

Yazarlar maddi desteklerinden dolayı KTO Karatay Üniversitesine teşekkür etmektedir.

Kaynakça

- Abolfathi, S., Yeganeh-Bakhtiary, A., Hamze-Ziabari, S. M., & Borzooei, S. (2016). Wave runoff prediction using M5' model tree algorithm. *Ocean Engineering*, 112, 76–81. <https://doi.org/10.1016/j.oceaneng.2015.12.016>
- Adamowski, J., Chan, H. F., Prasher, S. O., & Sharda, V. N. (2012). Comparison of multivariate adaptive regression splines with coupled wavelet transform artificial neural networks for runoff forecasting in Himalayan micro-watersheds with limited data. *Journal of Hydroinformatics*, 14(3), 731–744. <https://doi.org/10.2166/hydro.2011.044>
- Akçın, H., Kutoğlu, H. Ş., & Terlemezoğlu, B. (2005). Deniz Dibi Topografyasının Yapay Sinir Ağlarıyla Modellenmesi (pp. 1–8). Ankara: TMMOB Harita ve Kadastro Mühendisleri Odası 10. Türkiye Harita Bilimsel ve Teknik Kurultayı.
- Bhattacharya, B., & Solomatine, D. P. (2005). Neural networks and M5 model trees in modelling water level-discharge relationship. *Neurocomputing*, 63(SPEC. ISS.), 381–396. <https://doi.org/10.1016/j.neucom.2004.04.016>

- De Andrés, J., Lorca, P., De Cos Juez, F. J., & Sánchez-Lasheras, F. (2011). Bankruptcy forecasting: A hybrid approach using fuzzy c-means clustering and multivariate adaptive regression splines (MARS). *Expert Systems with Applications*, 38(3), 1866–1875. <https://doi.org/10.1016/j.eswa.2010.07.117>
- Demir, V., & Ülke, A. (2020). Yapay Sinir Ağları Yardımıyla Yükseklik Modellemesi (Samsun -Mert Irmağı Havzası Örneği) Height Modeling with Artificial Neural Networks (Samsun-Mert River Basin). *Gazi Mühendislik Bilimleri Dergisi*, 6(1), 54–61. <https://doi.org/https://dx.doi.org/10.30855/gmbd.2020.01.05>
- Demir V, Bilge H, & Bektaş S. (2017) Yapay Sinir Ağları Yardımıyla Sayısal Yükseklik Modellemesi, IX. Ulusal Hidroloji Kongresi 04 – 06 Ekim 2017 Diyarbakır/Türkiye (Özet Bildiri)
- Demir V, Çıtakoğlu H, Geyikli M. S., & Kuyucu H. (2017). Estimation of Digital Elevation Model by Artificial Intelligence Methods. 1 st International Symposium on Multidisciplinary Studies and Innovative Technologies, 257, November 2-4 Tokat/Türkiye (Özet Bildiri)
- Dipova, N., & Cangir, B. (2010). Lagün kökenli kil-silt zeminde sikişabilirlik özelliklerinin regresyon ve yapay sinir ağları yöntemleri ile belirlenmesi. *Teknik Dergi/Technical Journal of Turkish Chamber of Civil Engineers*, 21(3), 5069–5086.
- Elith, J., & Leathwick, J. (2007). Predicting species distributions from museum and herbarium records using multiresponse models fitted with multivariate adaptive regression splines. *Diversity and Distributions*, 13(3), 265–275. <https://doi.org/10.1111/j.1472-4642.2007.00340.x>
- Etemad-Shahidi, A., & Mahjoobi, J. (2009). Comparison between M5 tree model tree and neural networks for prediction of significant wave height in Lake Superior. *Ocean Engineering*, 36(15–16), 1175–1181. <https://doi.org/10.1016/j.oceaneng.2009.08.008>
- Leathwick, J. R., Rowe, D., Richardson, J., Elith, J., & Hastie, T. (2005). Using multivariate adaptive regression splines to predict the distributions of New Zealand's freshwater diadromous fish. *Freshwater Biology*, 50(12), 2034–2052. <https://doi.org/10.1111/j.1365-2427.2005.01448.x>
- Lee, T.-S., Chiu, C.-C., Chou, Y.-C., & Lu, C.-J. (2006). Mining the customer credit using classification and regression tree and multivariate adaptive regression splines. *Computational Statistics & Data Analysis*, 50(4), 1113–1130. <https://doi.org/10.1016/j.csda.2004.11.006>
- Mitchell, T. M. (1997). *Machine learning*. New York: The McGraw-Hill Companies.
- P. Bera, S. O. Prasher, R. M. Patel, A. Madani, R. Lacroix, J. D. Gaynor, ... S. H. Kim. (2006). Application of MARS in Simulating Pesticide Concentrations in Soil. *Transactions of the ASABE*, 49(1), 297–307. <https://doi.org/10.13031/2013.20228>
- Pal, M., & Deswal, S. (2009). M5 model tree based modelling of reference evapotranspiration. *Hydrological Processes*, 23(10), 1437–1443. <https://doi.org/10.1002/hyp.7266>
- Pourghasemi, H. R., & Rossi, M. (Eds.). (2019). *Natural Hazards GIS-Based Spatial Modeling Using Data Mining Techniques (Vol. 48)*. Cham: Springer International Publishing. <https://doi.org/10.1007/978-3-319-73383-8>
- Put, R., Xu, Q. S., Massart, D. L., & Vander Heyden, Y. (2004). Multivariate adaptive regression splines (MARS) in chromatographic quantitative structure-retention relationship studies. *Journal of Chromatography A*, 1055(1–2), 11–19. <https://doi.org/10.1016/j.chroma.2004.07.112>
- Quinlan, J. R. (1992). *Learning With Continuous Classes*. *World Scientific*, 92, 343–348. <https://doi.org/10.1.1.34.885>
- Sephton, P. (2001). Forecasting recessions: can we do better on MARS? *Review*, 83(2), 39–49.
- Sharda, V. N., Patel, R. M., Prasher, S. O., Ojasvi, P. R., & Prakash, C. (2006). Modeling runoff from middle Himalayan watersheds employing artificial intelligence techniques. *Agricultural Water Management*, 83(3), 233–242. <https://doi.org/10.1016/j.agwat.2006.01.003>
- Solomatine, D. P., & Xue, Y. (2004). M5 Model Trees and Neural Networks: Application to Flood Forecasting in the Upper Reach of the Huai River in China. *Journal of Hydrologic Engineering*, 9(6), 491–501. [https://doi.org/10.1061/\(ASCE\)1084-0699\(2004\)9:6\(491\)](https://doi.org/10.1061/(ASCE)1084-0699(2004)9:6(491))
- Zhang, W. G., & Goh, A. T. C. (2013). Multivariate adaptive regression splines for analysis of geotechnical engineering systems. *Computers and Geotechnics*, 48, 82–95. <https://doi.org/10.1016/j.compgeo.2012.09.016>
- Zhang, W., & Goh, A. T. C. (2014). Multivariate adaptive regression splines and neural network models for prediction of pile drivability. *Geoscience Frontiers*, 7(1), 45–52. <https://doi.org/10.1016/j.gsf.2014.10.003>