

Artificial Neural Networks Application for Modelling of Wastewater Treatment Plant Performance

Ece Ceren Yilmaz¹ and Emrah Dogan²

¹ Sakarya University, Department of Environmental Engineering, Esentepe Campus, 54187 Sakarya, Turkey

² Sakarya University, Department of Civil Engineering, Esentepe Campus, 54187 Sakarya, Turkey

Abstract: Biological oxygen demand (BOD) has been shown to be an important variable in water quality management and planning. However, BOD is difficult to measure and needs longer time periods (5 day) to get results. Artificial Neural Networks (ANNs) are being used increasingly to predict and forecast water resources variables. The objective of this research was to develop artificial neural networks (ANNs) model to estimate daily biological oxygen demand (BOD) at the influent of wastewater biological treatment plant. The plantscale data set (364 daily records of the year 2005) were obtained from a local wastewater treatment plant. Various combinations of daily water quality data, namely chemical oxygen demand (COD), water discharge (Q_w), suspended solid (SS), total nitrogen (N) and total phosphorus (P) are used as inputs into the ANN so as to evaluate the degree of effect of each of these variables on daily influent BOD. The results of the ANN model is compared with multiple linear regression model (MLR). Mean square error, average absolute relative error and coefficient of determination statistics are used as comparison criteria for the evaluation of the model performances. Based on the comparisons, it was found that the ANN model could be employed successfully in estimating the daily influent BOD of wastewater biological treatment plant and also ANN model is superior to MLR technique.

Keywords: Water quality, waste water treatment plant, biological oxygen demand, artificial neural network, multiple linear regression analysis.

Bir Atıksu Arıtma Tesisi Performansının Modellenmesi İçin Yapay Sinir Ağlarının Uygulanması

Özet: Biyolojik oksijen ihtiyacı (BOİ), su kalitesinin yönetimi ve planlamasında en önemli parametrelerden biri olarak gösterilmektedir. Fakat son derece önemli olan bu parametrenin ölçümü zordur ve ölçüm sonuçlarının elde edilmesi 5 gün gibi uzun bir süre almaktadır. Ölçümlerin zorluğu ve zaman alması ölçümlerin maliyetini de arttırmaktadır. Bu tür zorlukların üstesinden gelmek için bilim adamları bazı metotlar geliştirme yoluna gitmişlerdir. Bu metotlardan birisi de son zamanlarda literatürde sıkça kullanılan yapay zeka modelleridir. Bu çalışmada bir atıksu arıtma tesisinin girişindeki BOİ yapay zeka metotlarından olan yapay sinir ağları (YSA) ile tahmin edilmeye çalışılmıştır. YSA modelinin kurulmasında atıksu arıtma tesisinin girişindeki günlük kimyasal oksijen ihtiyacı (KOİ), günlük su debisi (Q_w), günlük askıda katı madde (AKM), günlük toplam azot (N) ve günlük toplam fosfor (P) parametreleri girdi olarak kullanılırken günlük BOİ parametresi ise çıktı olarak kullanılmıştır. Ayrıca, en iyi sonucu veren modelin araştırılmasında çeşitli girdi kombinasyonları kullanılarak BOİ tahminleri yapılmıştır. YSA modellerinin çıktılarının performansları, determinasyon katsayısı (R²), ortalama mutlak hata (OMH) ve ortalama karesel hata (OKH) gibi hata performans fonksiyonları kullanılarak değerlendirilmiştir. Ayrıca YSA sonuçları çoklu regresyon analizi (ÇRA) sonuçları ile de karşılaştırılmıştır. Performans sonuçlarına bakıldığında YSA modelinin BOİ tahmininde ÇRA' dan çok daha etkili bir model olduğu ve gerçeğe çok yakın sonuçlar verdiği görülmüştür.

Anahtar Kelimeler: Su kalitesi, atıksu arıtma tesisi, biyolojik oksijen ihtiyacı, yapay sinir ağları, çoklu regresyon analizi.

Reference to this paper should be made as follows (bu makaleye aşağıdaki şekilde atıfta bulunulmalı):

E.C.Yilmaz and E.Dogan, 'Modelling of Wastewater Treatment Plant Performance Using Adaptive Neuro Fuzzy Inference Systems', Elec Lett Sci Eng, vol. 4(1), (2008), p1-9

* Corresponding author; Tel.: +(90) 264 295 5749, E-mail:emrahdogan77@gmail.com

1 Giriş

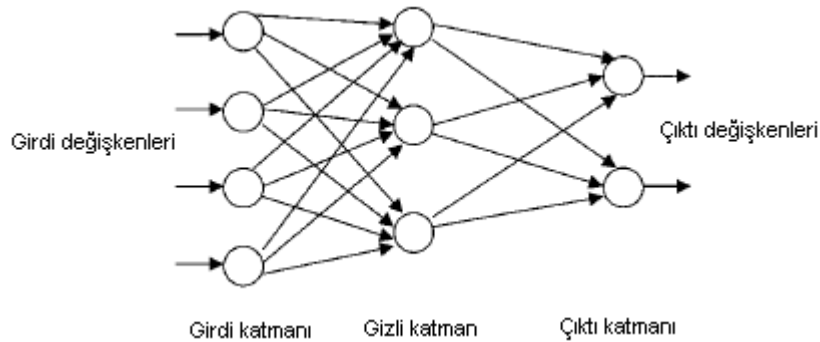
Su kalitesinin yönetimi ve planlanmasında son derece önemli olan BOİ miktarının doğru bir şekilde tahmin edilmesi çevre mühendisliğinde önem arz etmektedir. Çünkü BOİ miktarının tahmini, su kalitesinin çevreye olan etkilerinin belirlenmesi açısından önemli rol oynamaktadır [1]. Ancak bu kadar önemli olan bu parametrenin ölçümü zordur ve ölçüm sonuçlarının elde alınması 5 gün gibi uzun bir zaman almaktadır. Ölçümlerin zorluğu ve uzun süre alması ölçümlerin maliyetini de arttırmaktadır. Araştırmacılar su kalitesi parametrelerini kullanarak regresyon analizi yapmaktadırlar. Ancak bu klasik regresyon analizi problemin doğasından kaynaklanan lineer olmayan karmaşık ilişkiler sebebiyle iyi sonuçlar vermemektedir. Lineer olmayan problemler, klasik yöntemlere her zaman rahat modelleme imkanı vermez. Belirsizlik ve kesinsizlik durumlarında daha uygun olarak kullanılacak başka yöntemlerde mevcuttur. Literatürde esnek yöntemler (soft computing) olarak ifade edilebilecek bu grup içerisinde yapay sinir ağları (YSA), bulanık mantık (BM), adaptif sinirsel bulanık sistemler (ASBS) gibi yöntemler bulunmaktadır.

Bu nedenle son yıllarda su kalitesi parametreleri arasındaki ilişkiyi açıklamak için yapay zeka teknikleri kullanılmaya başlanmıştır [2, 3, 4, 5, 6, 7, 8, 9, 10]. Çalışma kapsamında, yapay sinir ağları (YSA) kullanılarak bir atıksu tesisinin girişindeki günlük BOİ ihtiyacı miktarının tahminleri yapılmıştır. Ayrıca YSA tahminleri çoklu regresyon analizi (ÇRA) tahminleri ile kıyaslanmıştır.

2 Yapay Sinir Ağları

Yapay sinir ağları (YSA) kavramı insan beyninin çalışma ilkelerinin sayısal bilgisayarlar üzerinde taklit edilmesi fikri ile ortaya çıkmış olup ilk çalışmalar nöronların matematiksel modellenmesi üzerinde yoğunlaşmıştır. Yapılan çalışmalar nöronların komşu nöronlarla bilgi alışverişinde bulunduğunu ortaya çıkarmıştır. Yapay sinir ağları diye isimlendirilen alan bu nöronların belli biçimlerde bir araya gelmesinden oluşmuştur. YSA modelleri, algoritmik olmayan paralel ve yayılı bilgi işleme yetenekleri ile klasik modellerden farklıdır. Farklı olan bu özellikleri sayesinde YSA karmaşık ve doğrusal olmayan hesapları kolaylıkla ve hızlı bir biçimde yapabilir. Algoritmik olmayan ve çok yoğun paralel işlem yapabilen YSA, ayrıca öğrenme kabiliyeti ve paralel dağıtılmış hafıza ile de hesaplamada yeni bakış açılarına sebep olmuştur. Girdi katman nöronları girdi bilgilerini alır bağlantılar vasıtasıyla bir sonraki bilgi işleme tabakası elemanlarına değerleri iletir. Bu işlem çıktı tabakasına ulaşıncaya kadar devam eder. Bu tür bilgi akışının bir yönde ilerlemesiyle oluşan ağ ileri beslemeli ağ olarak bilinir.

3 katmanlı ileri beslemeli tipik bir YSA modeli Şekil 1’de verilmiştir.



Şekil 1. Tipik 3 katmanlı ileri beslemeli YSA mimarisi

Günümüzde YSA bir çok bilim alanına uygulanmaktadır. Bu yaklaşım diğer bilim dallarında olduğu gibi hidrolik ve hidroloji bilim dallarında da iyi sonuçlar elde etmek için kullanılmaktadır. Su kaynakları sistemleri lineer olmayan ve pek çok parametreye sahip kompleks ilişkilerden oluşur. Bu tür problemler YSA kullanılarak etkili bir şekilde çözülebilir. Ayrıca YSA probleme kolayca uyum gösterebilmektedir. Hidroloji alanındaki çalışmalarda en yaygın olarak kullanılan YSA mimarisi çok katmanlı geri yayılım algoritmaları ile beslemeli ağ modelidir [11].

3 İleri Beslemeli Geri Yayınımı Yapay Sinir Ağları (İBGYYSA) Modeli

Bu çalışmada, YSA(i,j,k) mimarisi, sırasıyla i, j ve k simgelerinin girdi, gizli ve çıktı katmanlarını gösterecek şekilde oluşturulmuştur. Her katman birçok nörondan oluşmakta olup katmanlar arasında ağırlık kümeleri ile bağlanmaktadır. Bağlanma şekli ve her kısımdaki nöron sayısı değişebilmektedir. Aynı kısımdaki nöronlar arasında iletişim olmasına izin verilmemektedir. Bu eğitim sürecinin başında bağlantı kuvvetleri rasgele değerler olarak atanmaktadır. Öğrenme algoritması her iterasyonda eğitim başarı ile tamamlana kadar kuvveti değiştirmektedir. İterasyon süreci bir sonuca vardığında bağlantı kuvvetleri eğitim sürecinde kullanılan örneklerdeki mevcut bilgiyi elde eder ve saklar. Yeni bir girdi grubu sunulduğunda ileri doğru besleme ile yapay sinir ağının bağlantı kuvvetlerindeki öğrenilmiş ve saklanan bilgi sayesinde bir çıktı grubu elde edilir. Burada i ve k değerleri 1 olup j değerleri 2, 3, 5, 7 ve 10 değerleri alınarak YSA modelleri oluşturulmuştur. Bu çalışmada transfer fonksiyonu olarak sigmoid fonksiyonu, YSA eğitimi için de genelleştirilmiş delta kuralına dayalı olan geri yayılım algoritması kullanılmıştır [12].

4 Çoklu Regresyon Analizi (ÇRA)

Şayet bağımlı değişken (çıkıtı) olan Y' e etki eden m sayıda bağımsız değişkenler (girdiler) var ise (X_1, X_2, \dots, X_m) bağımlı değişken ile bağımsız değişkenler arasındaki lineer fonksiyon aşağıdaki gibi yazılabilir:

$$y = a + b_1x_1 + b_2x_2 + \dots + b_mx_m \quad (1)$$

Regresyon katsayıları (a, b_1, b_2, \dots, b_m) en küçük kareler yöntemiyle aşağıdaki gibi bulunabilir.

$$\sum_{i=1}^N e_{yi}^2 = \sum_{i=1}^N (y_i - a - b_1x_{1i} - b_2x_{2i} - b_mx_{mi})^2 \quad (2)$$

5 Dataların Toplanması

Bu çalışmada kullanılan 2005 yılına ait günlük kimyasal oksijen ihtiyacı (KOİ), günlük su debisi (Q_w), günlük askıda katı madde (AKM), günlük toplam azot (N), günlük toplam fosfor (P) ve günlük biyolojik oksijen ihtiyacı (BOİ) parametreleri Adapazarı Şehrindeki bir atıksu arıtma tesisinin girişinden toplanmıştır.

Bu parametrelerin her birinin istatistiksel analizleri Tablo 1'de verilmiştir. Bu tabloda x_{ort} , S_x , C_v , C_{sx} , x_{min} , and x_{maks} her bir parametrenin ortalaması, standart sapması, varyansı, çarpıklığı, minimum ve maksimum değerlerini göstermektedir. En çok değişkenlik gösteren parametrenin

AKM ($C_v=0.50$ mm) olduğu, KOİ parametresinin BOİ ile en yüksek korelasyona ($R=0.954$) sahip olduğu, su debisi Q_w nin BOİ ile ters orantılı ($R=-0.357$) olduğu da Tablo 1'den anlaşılmaktadır.

Tablo 1. Çalışmadaki herbir parametrenin istatistiksel analizi

Data set	x_{ort}	S_x	$C_v(S_x/x_{ort})$	C_{sx}	x_{min}	x_{maks}	BOİ ile korelasyon (R)
BOİ(mg/l)	237.054	99.13	0.42	-0.16	33	610	1.000
KOİ(mg/l)	445.407	178.85	0.40	-0.31	73	865	0.954
N(mg/l)	48.42	17.89	0.37	-0.53	9.5	81	0.871
P(mg/l)	4.35	1.97	0.45	0.24	0.5	9.8	0.650
AKM(mg/l)	322.228	162.28	0.50	1.87	52	1395	0.452
$Q_w(m^3/gün)$	63562	17233	0.27	-0.73	15920	94410	-0.357

6 İBGYYSA Modelinin Uygulanması

Beş girdi vektörü (KOİ, Q_w , AKM, N ve P) ve bir çıktı vektöründen (BOİ) oluşan 364 verinin analizi göz önüne alınmıştır. Bu veriler Denklem 3 kullanılarak 0.1-0.9 arasında normalize edilmiş eğitim ve test setlerini oluşturmak üzere iki gruba ayrılmıştır. Eğitim seti 244, geriye kalan 120 adet veri seti ise programın gerçek değerlere yaklaşım performansının değerlendirmesinde test seti olarak kullanılmıştır.

Bu çalışmada hataların değerlendirilmesi için, determinasyon katsayısı (R^2), ortalama karesel hata (OKH) ve ortalama mutlak hata (OMH) fonksiyonları kullanılmıştır.

Öncelikle modelin uygulanması için tüm veriler 0.1 ile 0.9 arasında normalize edilmiştir.

$$x_i = 0.8 \cdot (X_i - X_{min}) / (X_{mak} - X_{min}) + 0.1 \quad (3)$$

Burada, x_i normalize edilmiş değerleri göstermekte olup, X_{mak} and X_{min} maksimum ve minimum ölçülen değerlerdir. Normalizasyon yapılarak veriler boyutsuz hale getirilmiş olur.

$$R^2 = \frac{BOİ_o - BOİ_s}{BOİ_s} \quad (4)$$

burada:

$$BOİ_o = \sum_{i=1}^n (BOİ_{i(ölçülen)} - BOİ_{i(ortalama)})^2 \quad (5)$$

$$BOİ_s = \sum_{i=1}^n (BOİ_{i(ölçülen)} - BOİ_{i(tahmin edilen)})^2 \quad (6)$$

burada, $BOİ_{i(ölçülen)}$ ve $BOİ_{i(tahmin edilen)}$ sırasıyla günlük ölçülen ve İBGYYA ile hesaplanan BOİ miktarlarıdır. Ortalama günlük BOİ miktarı da $BOİ_{(ortalama)}$ olarak verilmiştir.

Ortalama karesel hata (OKH) şu şekilde tanımlanabilir,

$$MSE = \frac{1}{n} \sum_{i=1}^n (Q_{si(ölçülen)} - Q_{si(tahmin edilen)})^2 \quad (7)$$

Modelin güvenilirliğinin test edilmesinde global performans fonksiyonlarının (R^2 , OKH) haricinde ortalama mutlak hata (OMH) gibi performans fonksiyonu da kullanılmıştır.

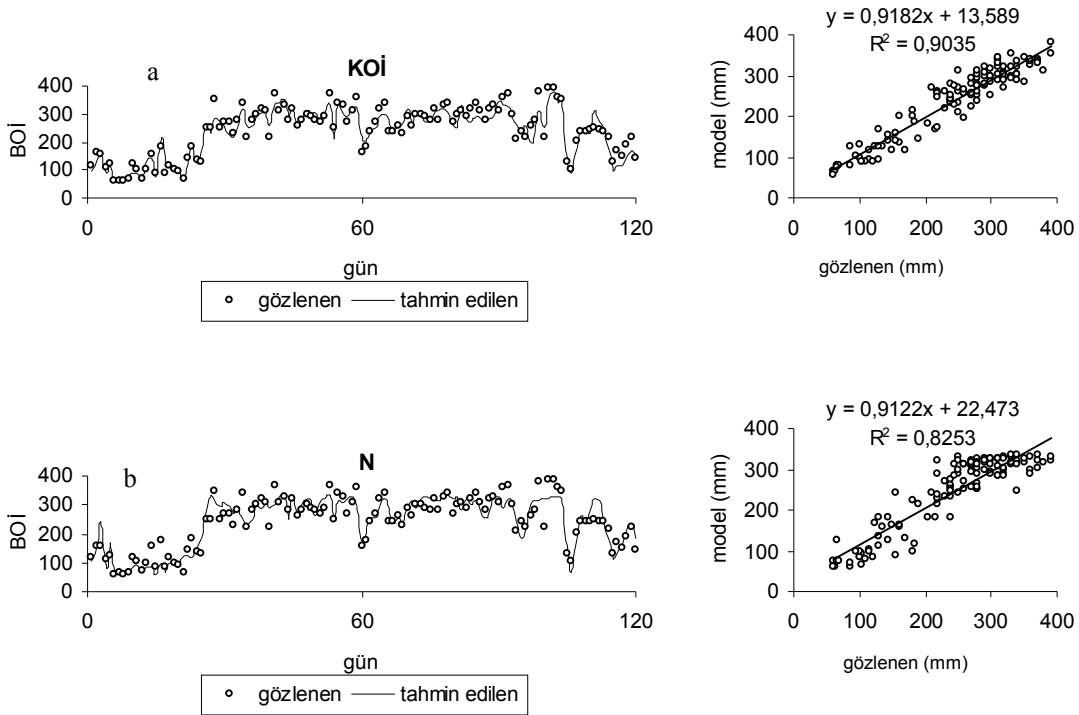
$$OMH = \frac{1}{n} \sum_{i=1}^n |RH| \quad \text{burada,} \quad (8)$$

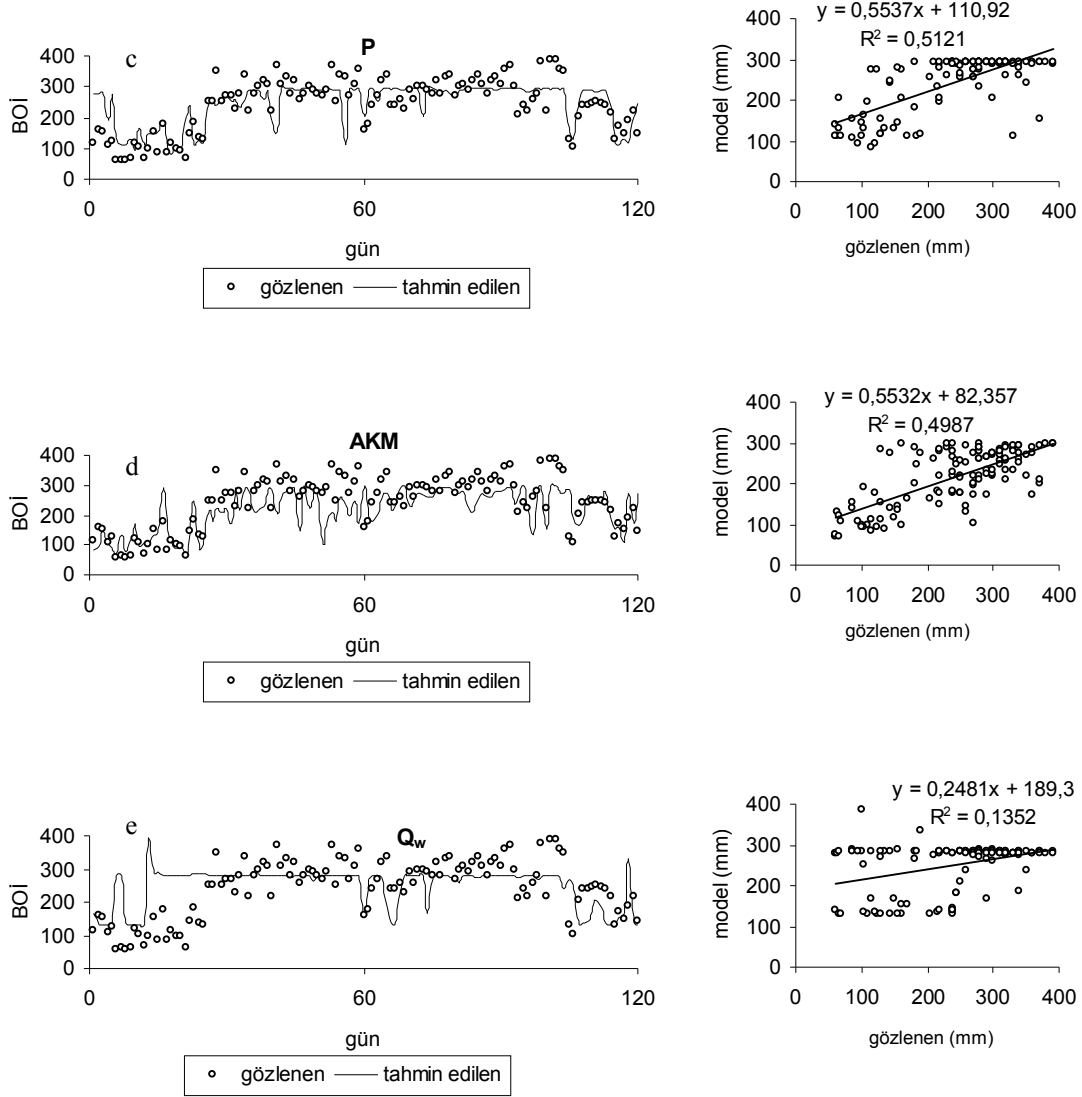
$$RH = \frac{BO\ddot{I}_{i(\ddot{o}l\ddot{c} \ddot{u}l\ddot{e}n)} - BO\ddot{I}_{i(\text{tahmin edilen})}}{BO\ddot{I}_{i(\ddot{o}l\ddot{c} \ddot{u}l\ddot{e}n)}} \cdot 100 \quad (9)$$

burada, RH rölatif hatadır ve yüzde olarak belirlenir.

7 Duyarlık Analizinin İBGYYSA Modeli Kullanılarak Yapılması

İBGYYSA modelinde kullanılacak olan girdi parametrelerinin seçimi modelin performansı açısından önem arz etmektedir. Modeldeki girdilerin etkinlik dereceleri duyarlılık analizi ile belirlenebilmektedir. BOİ tahmini için herbir girdi parametresi ayrı ayrı kullanılarak olaya en duyarlı parametre İBGYYSA ile bulunmuştur. Sonuçlar Şekil 2a-e de gösterilmiştir. Ayrıca hangi girdi kombinasyonlarının en etkili modeli vereceği duyarlılık analizi kullanılarak bulunmuştur (Tablo 2). Duyarlık analizi sonucunda tüm girdi parametrelerin BOİ tahmininde önem arz ettiği görülmüştür. KOİ girdi parametresinin BOİ tahmini için en etkili parametre olduğu, su debisi Q_w nin ise en az etkili parametre olduğu duyarlılık analizi ile tespit edilip sırasıyla Şekil 2a ve 2e' de gösterilmiştir.





Şekil 2a-e. Her bir girdi parametresinin BOİ tahminindeki performansının İBGYYSA ile Belirlenmesi

Tablo 2. En etkili İBGYYSA modelinin duyarlık analizi kullanılarak belirlenmesi

Performance	KOİ	KOİ+N	KOİ+N+P	KOİ+N+P+AKM	KOİ+N+P+AKM+Q _w
OMH (%)	10.52	10.25	10.18	10.10	10.03*
OKH	819.41	778.67	726.25	714.69	708.01*
R ²	0.903	0.903	0.911	0.915	0.919*

Not: En iyi sonuçlar ‘*’ ile gösterilmiştir.

8 En Uygun İBGYYSA Modelinin Belirlenmesi

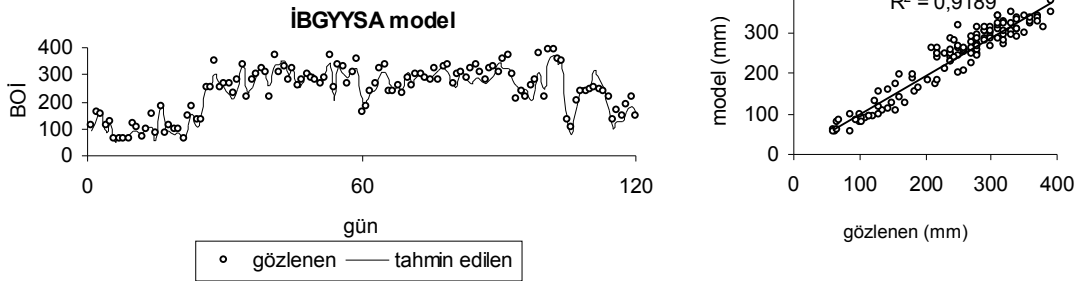
En iyi sonucu veren İBGYYSA mimarisi çeşitli deneme yapımlardan sonra elde edilmiştir. İBGYYSA mimarisinin belirlenmesi Tablo 3 de gösterilmiştir. En etkili İBGYYSA mimarisinin bulunmasında OKH ve R² performans fonksiyonları kullanılmıştır. Bu çalışmada, gizli katman nöron sayısı çeşitli denemelerden sonra test seti performans değerlerinden, Tablo 3 de

gösterildiği üzere 3 olarak belirlenmiştir. Performansı en yüksek olan YSA(5 3 1) modelidir. Ayrıca modelin eğitimi için en uygun iterasyon sayısının 1000 olduğu Tablo 3 den anlaşılmaktadır.

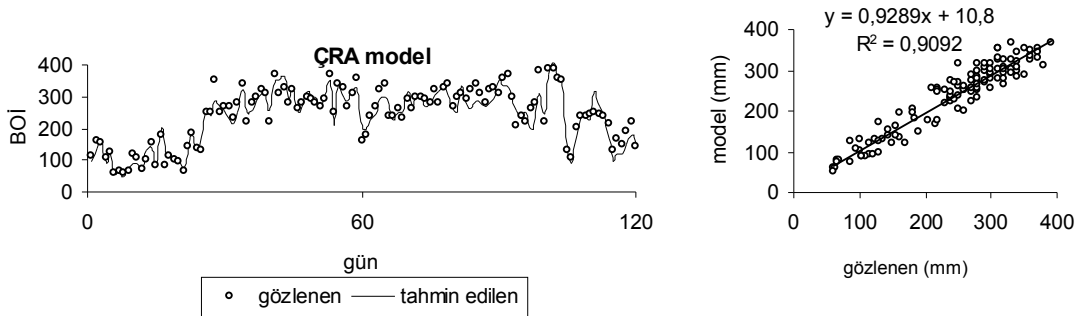
Tablo 3. En etkili İBGYYSA mimarisinin belirlenmesi

YSA mimarisi (tabakalardaki neron sayısı)	İterasyon Sayısı (Epoch)	Determinasyon Katsayısı (R^2)	Ortalama Karesel Hata (OKH)
YSA(5, 2, 1)	1000	0.917	716.21
YSA(5, 3, 1)	1000	0.919	708.01
YSA(5, 5, 1)	1000	0.915	737.30
YSA(5, 7, 1)	1000	0.913	757.42
YSA(5, 10, 1)	1000	0.895	869.38
YSA(5, 2, 1)	2000	0.908	743.79
YSA(5, 3, 1)	2000	0.909	755.84
YSA(5, 5, 1)	2000	0.894	864.95
YSA(5, 7, 1)	2000	0.898	828.60
YSA(5, 10, 1)	2000	0.884	966.52

İBGYYSA modelinin eğitilmesinden sonra model test edilmiştir. Test seti BOİ tahminleri ölçülen BOİ ler ile karşılaştırıldığında İBGYYSA tahminlerinin gözlenenlere çok yakın sonuç verdiği görülmüştür. Tahmin edilen değerlerin gözlenen değerlere çok yakın olduğu ve eğilimlerinin nerdeyse birebir yakın olduğu Şekil 3 de görülmektedir. Ayrıca İBGYYSA tahminleri ÇRA tahminleri ile kıyaslanmıştır. Bu kıyaslamalar da Şekil 4 de gösterilmiştir. Karşılaştırmalara bakıldığında İBGYYSA modelinin ÇRA modelinden daha iyi sonuçlar verdiği anlaşılmıştır.



Şekil 3. İBGYYSA modeli BOİ tahminlerinin gözlenen BOİ ler ile kıyaslanması



Şekil 4. ÇRA modeli BOİ tahminlerinin gözlenen BOİ ler ile kıyaslanması

9 Sonular

Bu alıřmada İBGYSA modeli kullanılarak bir atıksu arıtma tesisinin giriřindeki gnlk BOİ miktarları tahmin edilmeye alıřılmıřtır. Her bir girdi parametresi ayrı ayrı modelde girdi olarak kullanılmıř ve her bir parametrenin BOİ tahminindeki etkinlik dereceleri bulunmuřtur. Bunun sonucunda en etkili parametrenin KOİ olduėu, daha sonra sırasıyla N, P AKM etkili olduėu, en az etkili parametreninde Q_w olduėu tespit edilmiřtir. Ayrıca etkili bir İBGYSA modelinin girdi parametrelerine baėlı olmasından dolayısıyla en etkili modelin tayini iin duyarlık analizi yapılmıřtır. Duyarlık analizi sonucunda tm girdi parametrelerinin (KOİ, Q_w , AKM, N ve ,P) olaya etki ettiėi gzlemlenmiřtir. Ayrıca İBGYSA tahminleri RA tahminleri ile kıyaslanmıřtır. Kıyaslamalara bakıldıėında İBGYSA performansının RA performansından daha iyi olduėu anlařılmıřtır. Sonu olarak İBGYSA modeli BOİ tahmininde gereėe yakın ve gvenilir tahminler vermesi sebebiyle su kalitesi ynetiminde ok etkin bir model olarak kullanılabileceėi sylenebilir.

References (Referanslar)

1. Aguilera PA, Frenich JA, Torres, Castro H, Vidal JLM and Canton M (2001) Application of the Kohonen Neural Network in Coastal Water Management: Methodological Development for the Assessment and Prediction of Water Quality. *Water Research* 35 (17) 4053–4062.
2. J.P. Suen, J.W. Eheart, M. Asce, Evaluation of Neural Networks for Modelling Nitrate Concentration in Rivers, *J. of Wat. Res. Plan. and Manag.*, 129 (2003) 505-510.
3. P.A. Aguilera, A.G. Frenich, J.A. Torres, H. Castro, J.L.M. Vidal and M. Canton, Application of the Kohonen neural network in coastal water management: methodological development for the assessment and prediction of water quality, *Water Res.* 35 (2001) 4053–4062.
4. A.H. Lobbrecht, D.P. Solomatine Control of Water Levels in Polder Areas Using Neural Networks and Fuzzy Adaptive Systems. *Water Industry Systems: Modeling and Optimization Applications* 1 (1999) 509-518.
5. H.R. Maier, G.C. Dandy, The use of artificial neural networks for the prediction of water quality parameters, *Wat. Res. Research* 32 (1996) 1013–1022.
6. L.G. Sovan, A. Maritxu. J. Giraudel, Prediction of Stream Nitrogen Concentration From Watershed Features Using Neural Network, *Wat. Res.* 33 (1999) 3469-3478.
7. C.G. Wen, C.S. Lee, A Neural Network Approach to Multiobjective Optimization for Water Quality Management in a River Basin, *Wat. Res. Research* 34 (1998) 427–436.
8. I. Zaheer, C.G. Bai, Application of Artificial Neural Network for Water Quality Management, *Lowland Tech. Int.* 5 (2003) 10-15.
9. S. Fogelman, M. Blumenstein, H. Zhao, Estimation of chemical oxygen demand by ultraviolet spectroscopic profiling and artificial neural networks, *Neural Comput & Applic.*, 15 (2006) 197–203.

10. B. Sengorur, E. Dogan, R. Koklu, A. Samandar Dissolved Oxygen Estimation Using Artificial Neural Network For Water Quality Control, Fresenius Env. Bul. 15 (2006) 1064-1067.
11. Govindaraju, R.S., and Rao, R. A., (2000), Artificial Neural Networks in Hydrology, Kluwer Academic Publishers, 93-109.
12. ASCE Task Committee on Application of Artificial Neural Networks in Hydrology. (2000). "Artificial neural networks in hydrology. I: Preliminary concepts." J. Hydrologic Engrg., ASCE, 5(2),115–123.