

DİNAMİK SU BÜTÇESİ MODELİNE DESTEK VEKTÖR REGRESYONU ENTEGRASYONU

Zeynep Beril ERSOY *^{ID}

Umut OKKAN **^{ID}

Okan FİSTİKOĞLU ***^{ID}

Alınma: 14.10.2021; düzeltme: 09.11.2021; kabul: 21.01.2022

Öz: Yağış-akış modelleri kapsamında ele alınan modeller içerisinde kavramsal modeller havza dinamiğini atanan parametreler yardımıyla benzeştirmeye çalışırken, kapalı kutu modelleri ise fiziksel süreci dikkate almadan veri işleme esaslı uygulanmaktadır. Her iki yöntemin de birbirine göre avantajlı ve dezavantajlı yönleri bulunmaktadır. Örneğin kavramsal modellerin bazı parametreleri doğrusal tanımlandıklarında simülasyonlarda yanlılıklar gözlemlenmektedir. Diğer yandan, kapalı kutu modelleri tutarlı bir simülasyon için gecikmeli yağış değerlerine ihtiyaç duymaktadır. Bu nedenle çalışmada her iki yaklaşımın iyi yönlerini birleştiren hibrit bir model yapısının ortaya konması amaçlanmıştır. Bu kapsamda, dinamik su bütçesi adı verilen kavramsal bir yağış-akış modelinin doğrusal davranış gösteren yeraltısu depolama elemanı yerine destek vektör makinesi eklenerek beş parametrelili hibrit bir model oluşturulmuştur. Destek vektör makinesi ilavesi ile doğrusal olmayan haritalama yetisi kazanan model Balıkesir'in İkizcetepeler Baraj Havzası'nda uygulanmıştır. Hibrit modelin kavramsal modele kıyasla kalibrasyon ve validasyon dönemlerinde sırasıyla %21 ve %14 daha düşük hata performansı vermesi istatistiksel açıdan anlamlı bulunmuştur.

Anahtar Kelimeler: Yağış-Akış Modellemesi, Kavramsal Model, Makine Öğrenmesi, Hibrit Modelleme, İkizcetepeler Barajı

Integrating Support Vector Regression into Dynamic Water Budget Model

Abstract: Among the various rainfall-runoff models, conceptual ones can simulate the basin dynamics by means of assigned parameters, while black-box models are applied as data-driven techniques which take no account of the physical process. Both types involve some advantages and shortcomings relative to each other. For instance, as some parameters assigned in conceptual ones are linear, the runoff simulations can be biased. Besides, black-box models generally require antecedent precipitation data to get a robust simulation. Therefore, in the study, it is intended to propose a hybrid model structure integrating the prominent aspects of both approaches. In this concept, the linear groundwater storage of the dynamic water budget model, one of the conceptual types, was eliminated and a support vector regression was included instead, and thus, a hybrid model with five parameters was built. The model, which achieved nonlinear mapping capability with the inclusion of support vector regression, was implemented for İkizcetepeler Dam located at Balıkesir. It was found statistically significant that hybrid model provided relatively lower error performance as 21% and 14% in calibration and validation periods, respectively, when it was compared to that of the conceptual one.

Keywords: Rainfall-Runoff Modelling, Conceptual Model, Machine Learning, Hybrid Modeling, İkizcetepeler Dam

* İnşaat Mühendisliği Bölümü, Hidrolik Anabilim Dalı, Balıkesir Üniversitesi, 10145, Balıkesir, Türkiye

** İnşaat Mühendisliği Bölümü, Hidrolik Anabilim Dalı, Balıkesir Üniversitesi, 10145, Balıkesir, Türkiye

*** İnşaat Mühendisliği Bölümü, Hidrolik Anabilim Dalı, Dokuz Eylül Üniversitesi, 35160, İzmir, Türkiye

İletişim Yazarı: Zeynep Beril ERSOY (zeynepberil.ersoy@balikesir.edu.tr)

1. GİRİŞ

Hidroloji literatüründe daha az veri veya parametre seti ile yeterli akım simülasyonu yapmanın önemi sıklıkla atfedilmektedir. Bu aşamada kavramsal yağış-akış modelleri (KYA) ve kapalı kutu modelleri (KK) pratik araçlar olarak görülmektedir. Bunlar arasında, havzanın homojen kabul edildiği kavramsal yağış-akış modelleri zemin nemi ve yeraltı suyu depolaması elemanları ile hidrolojik süreçler arasındaki ilişkileri çeşitli matematiksel ifadelerle temsil etmeye çalışmaktadır (Noori ve Kalin, 2016). Modellerde genellikle tanımlı süreç parametrelerinin toplu halde olduğu varsayılmakta ve simülasyonlarının hassasiyetleri parametrelerinin kalibrasyonunun ne ölçüde doğru yapıldığına bağlı olmaktadır (Okkan ve Kırdemir, 2020). Diğer bir pratik tür olan kapalı kutu modelleri ise meteorolojik girdiler ile hedef değişken olan akış çıktıları benzeştiren ve sıklıkla makine öğrenimi algoritmalarının kullanımına dayanan tekniklerdir.

Makine öğrenmesi (machine learning: ML) modellerinin fiziksel süreçleri detaylı ele almadan, hidro-meteorolojik veri setlerine bağlı doğrusal olmayan akım karakterizasyonu sağlaması ve geleneksel KYA modellerinin de parametrik yapıları ile havza hidrolojisini temsil edebilme yetileri araştırmacıları bu iki tekniği birlikte uygulamaya teşvik etmiştir (Ersoy 2021; Ersoy vd. 2021; Okkan vd. 2021). Hibrit modelleme olarak anılan bu uygulamalarda, hibritlenecek model elemanlarının birbiriyle uyumlu olup olmadığını kontrol etmek ve modelleme sürecinin zayıf yönlerini (başlangıç koşullarına duyarlılık ve aşırı parametre kullanımı) ortadan kaldırmak önem arz etmektedir. KYA modellerinin ve ML tekniklerinin (çoğunlukla yapay sinir ağları (ANN)) güçlü yönlerini birleştirmeyi esas alan hibrit modelleme çalışmalarına ilişkin kapsamlı bir literatür Okkan vd. (2021)'de sunulmuştur.

Bu modellerin kurulumu esas itibari ile KYA modellerinden elde edilen çıktıların (hidrolojik süreç bileşenleri ile ilişkili çeşitli simüle edilmiş veriler) ML modellerine tekrar girdi olarak sunulmasına dayanmaktadır. Bu kapsamdaki ilk örnek çalışma Ancia vd. (2004) tarafından gerçekleştirilmiştir. Çalışmalarında, kısa vadeli akış tahminini iyileştirmek amacıyla önceden tanımlanmış parametrelerle GR4J modelinden elde edilen simüle edilmiş zemin nem indeksi değerlerini ANN'ye girdi olarak sunmuşlardır. Ancia vd. (2004) çalışmasından ilham alan diğer araştırmalar da benzer uygulamalar içermektedir (Humphrey vd. 2016; Noori ve Kalin 2016; Kumanlioglu ve Fıstıkoğlu, 2019; Ersoy 2021; Okkan vd. 2021).

Bu hibritleme çalışmalarında, anlaşıldığı üzere KYA modelleri önce kalibre edilmekte ve daha sonra bunların belirli çıktıları ML modellerine yeni işlenmiş girdi seti olarak sunulmaktadır. Yani, her iki kalibrasyon süreci de ardışık olarak çalıştırılmakta ve ML tabanlı kısım, KYA modelinin kalibrasyonu tamamlanmadan devreye girmemektedir. Bu paralel hibritleme tarzı ile tahmin performansında anlamlı iyileşmeler öngörülse bile kalibrasyon aşamasında KYA model parametreleri ile ML model parametreleri arasındaki dinamik etkileşim, tanımlanmış tüm parametrelerin aynı iterasyon adımlarında birlikte optimize edilememesinden dolayı kısıtlanmaktadır. Buna ilaveten, hesap yoğunluğu da iki ardışık modelleme süreci ile gereksiz artmaktadır (Ersoy 2021).

Yukarıda değinilen eksikliklerin üstesinden gelme amacıyla hazırlanan bu çalışmada Destek Vektör Regresyonu (DVR) esaslı bir ML tekniği Dinamik Su Bütçesi modeli (*dynwbm*) adı verilen kavramsal bir modele entegre edilmiş ve tüm tanımlı parametreler (*dynwbm* parametreleri ve ilgili ML parametreleri) eşzamanlı olarak bu seri bağlı çalışan yapı içerisinde kalibre edilmiştir. Entegrasyon öncesinde *dynwbm*'nin parametrelerinin hassasiyeti varyans ayrıştırma analizi (ANOVA) ile irdelenip hibrit modelin ML içeren katmana hangi kavramsal çıktıların girdi olarak sunulacağı belirlenmiştir. Tekil biçimde kullanılan KYA ya da ML modellerine güçlü bir alternatif olarak ilk defa Okkan vd. (2021), Ersoy vd. (2021) ve Ersoy (2021) çalışmalarında önerilen ve Gediz Havzası bütününde test edilen bu seri hibrit modelleme stratejisi bu çalışmada Balıkesir-İkizcetepeler Baraj Havzası'nda uygulanarak ulusal ölçekteki hidrolojik model kullanıcılarına yeni bir bakış açısı kazandırmayı hedeflemektedir. Uygulanan

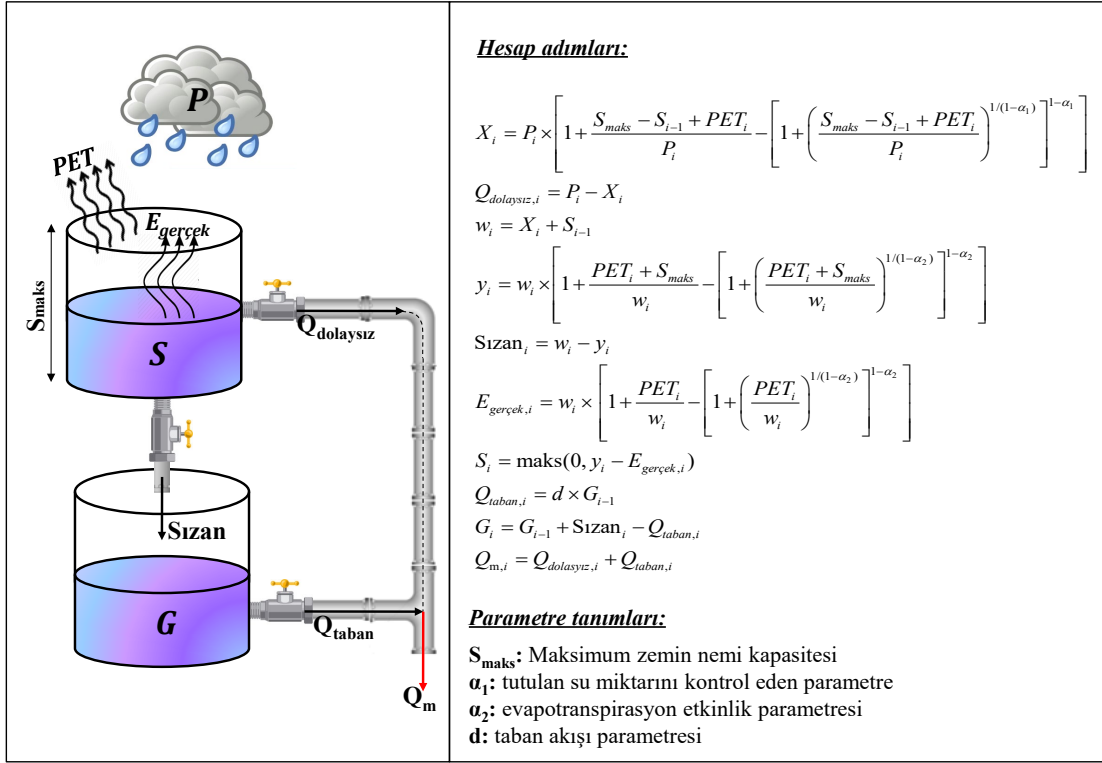
modelleme stratejisi ile ilgili ayrıntılar Bölüm 2'de sunulurken, çalışma alanı ve kullanılan veriler Bölüm 3'te, modellerden elde edilen bulgu ve analizler de Bölüm 4'te verilmiştir. Son bölüm ise çalışmanın tartışmalarını ve sonuçlarını kapsamaktadır.

2. YÖNTEM

Bu bölümde hibridizasyon aşamasında değerlendirilen *dynwbm* ve *DVR* tekniklerinin tanıtılmasının akabinde önerilen hibrit yaklaşım ile ilgili detaylar sunulmuştur. Bölüm sonunda ise performans değerlendirmesinde kullanılan kriterlerinden bahsedilmiştir.

2.1. Dinamik Su Bütçesi Modeli (*dynwbm*)

Çalışmada hibritlemenin Zhang vd. (2008) tarafından önerilen dinamik su bütçesi modeli (*dynwbm*) üzerinde denemesi kararlaştırılmıştır. Bu model Q_m aylık akış tahmininde girdi seti olarak aylık toplam yağış (P) ve potansiyel evapotranspirasyon (PET) verilerini kullanmaktadır. Modelin şematik bir diyagramı ve ilgili işlem adımları Şekil 1'de gösterilmekte olup, detaylar Zhang vd. (2008) çalışmasında verilmektedir.



Şekil 1:

Dinamik su bütçesi modelinin akış şeması (Ersoy vd., 2021; Okkan vd., 2021; Ersoy 2021)

Burada X havza tarafından tutulan su miktarını, S zemin nemi içeriğini, $Q_{dolaysiz}$ dolaysız akış miktarını, w mevcut su içeriğini, y evapotranspirasyon fırsatını, $Sızan$ yeraltı suyu depolamasına sızan su miktarını, E_{gercek} gerçek evapotranspirasyon miktarını, Q_{taban} taban akışını ve G ise aktif yeraltı suyu depolama miktarını göstermektedir.

2.2. Destek Vektör Regresyonu (*DVR*)

Destek vektör regresyonu, yağış-akış modellemesi ve akış simülasyonunda başarıyla kullanılan bir ML tekniğidir (Okkan ve Serbes 2012; Tongal ve Booij, 2018). ANN'de gizli

katmandaki hücre adedine bağlı olarak eğitilmesi gereken ağırlıkların fazlalaşabilmesi ve bundan dolayı aşırı öğrenme sorunu *DVR*'yi çoğu durumda daha pratik ve güvenilir kılmaktadır. *DVR*'nin ana önermesi, girdi verilerini doğrusal olmayan bir şekilde daha yüksek boyutlu özellik uzayına haritalayan bir kernel (çekirdek) fonksiyonunun kullanımına dayanmaktadır (Okkan vd. 2021; Ersoy 2021). *DVR*'den iyi sonuç elde etmek için, regülarizasyon parametresi (Cr) ve radyal tabanlı kernel fonksiyonunun genişlik parametresi (σ) olan iki kontrol parametresinin kalibre edilmesi gerekmektedir. ML tekniğinin diğer detayları Okkan ve Serbes (2012) çalışmasında bulunabilir.

2.3. Önerilen Hibrit Model: *d-DVR*

Şekil 2'de gösterildiği gibi önerilen hibrit model, *dynwbm* kavramsal yağış-akış modelinin destek vektör regresyonu ile entegrasyonundan oluşmaktadır. Burada *DVR* tabanlı ML kısmı, yeraltı suyu depolamasının doğrusal d parametresi ile yönetilen kavramsal taban akışı hesabının yerini almakta ve dört adet kavramsal çıktıyı akışa dönüştürmektedir. Kavramsal modelin taban akış bileşeninin seri şema içinde elimine edilmesinin nedeni bulgular bölümünde ayrıca açıklanmıştır. P , PET ve başlangıç zemin nemi içeriği (S_0) bilgisi istenen modelde 3 adet kavramsal parametre ve 2 adet de *DVR* parametresi olmak üzere 5 adet kalibre edilmesi gereken parametre mevcuttur. Hibrit modelde hangi kavramsal çıktıların kullanılacağına korelasyon analizleri neticesinde karar verilmiş olup, bunlar modelin dolaysız akış bileşeni ($Q_{dolaysız}$), sızan su miktarı ($Sızan$), zemin nemi (S) ve gerçek evapotranspirasyon ($E_{gerçek}$) çıktılarıdır. Bu seri şema içerisinde aynı zamanda temel bileşenler analizi (TBA) yaklaşımı da boyut küçültme aracı olarak değerlendirilmiştir.

TBA analizi ile kavramsal çıktıların boyutu ikiye indirgindikten sonra *d-DVR* olarak anılacak hibrit modelde, $[0 \ 1]$ arasında ölçeklenmiş $L \times 2$ boyutlu girdi matrisinin (Inp) derlenmesini takiben σ parametresine bağlı olarak kernel fonksiyonu matrisi oluşturulmaktadır. Radyal tabanlı kernel fonksiyonun öklid uzaklığı terimi (D) kalibrasyon ve validasyon veri setleri için Denklem 1 ve Denklem 2 eşitlikleri ile hesaplanabilmektedir.

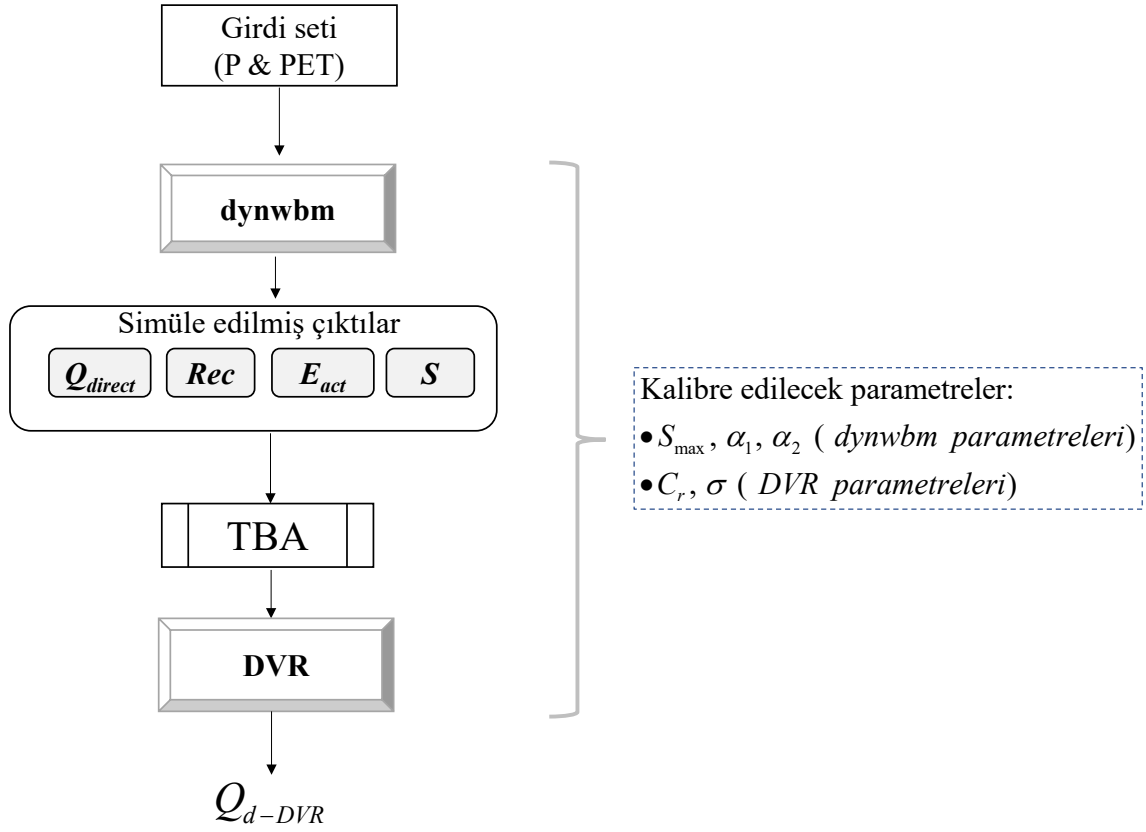
$$D = \text{diag}(X_{tr} \ X_{tr}^T) [1 \dots 1]_{1 \times L_{tr}} + \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}_{L_{tr} \times 1} \left[\text{diag}(X_{tr} \ X_{tr}^T) \right]^T - 2 X_{tr} \ X_{tr}^T \quad (1)$$

$$D = \text{diag}(X_{tr} \ X_{tr}^T) [1 \dots 1]_{1 \times L_{test}} + \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}_{L_{tr} \times 1} \left[\text{diag}(X_{test} \ X_{test}^T) \right]^T - 2 X_{tr} \ X_{test}^T \quad (2)$$

Burada L_{tr} ve L_{test} sırasıyla kalibrasyon ve validasyon verisinin uzunluğu, X_{tr} ve X_{test} ise Inp matrisi içerisinde sırasıyla kalibrasyon ve validasyon için ayrılan kısımları ifade etmektedir. D matrisi kalibrasyon ve validasyon dönemleri için elde edildikten sonra, kernel matrisine Denklem 3 ile erişilmektedir (D ve K kalibrasyon dönemi için $L_{tr} \times L_{tr}$ boyutlu iken validasyon döneminde $L_{tr} \times L_{test}$ boyutuna sahiptir).

$$K = \exp\left(-\frac{D}{2\sigma^2}\right) \quad (3)$$

Burada σ radyal tabanlı fonksiyonun genişlik parametresidir.



Şekil 2:
Önerilen hibrit modelleme stratejisi (Ersoy 2021; Okkan vd. 2021)

Modelin C_r regülarizasyon terimine (penaltı terimi) bağlı olarak Lagrange çarpanları (α) ve bias terimi (b) Denklem 4’te verilen ifadenin çözümü ile elde edilmektedir.

$$\begin{bmatrix} 0 & \bar{\mathbf{I}}^T \\ \bar{\mathbf{1}} & K + Cr^{-1}I \end{bmatrix}_{(L_{tr}+1) \times (L_{tr}+1)} \begin{bmatrix} b \\ \boldsymbol{\alpha} \end{bmatrix}_{(L_{tr}+1) \times 1} = \begin{bmatrix} 0 \\ \mathbf{y}_{tr} \end{bmatrix}_{(L_{tr}+1) \times 1} \quad (4)$$

Burada $\bar{\mathbf{1}} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}_{L_{tr} \times 1}$; $I = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}_{L_{tr} \times L_{tr}}$ şeklinde tanımlanmaktadır.

y_{tr} gözlenmiş akımları temsil etmekte olup α ile aynı boyuta ($L_{tr} \times 1$) sahiptir. Elde edilen α ve b değerlerine göre kalibrasyon ve validasyon dönemi çıktıları sırasıyla Denklem 5 ve Denklem 6 olarak elde edilmektedir.

$$[\mathbf{Q}_{m,s}]_{1 \times L_{tr}} = \boldsymbol{\sigma}^T x [K]_{L_{tr} \times L_{tr}} + b x [1 \dots 1]_{1 \times L_{tr}} \quad (5)$$

$$[\mathbf{Q}_{m,s}]_{1 \times L_{test}} = \vartheta^T x[K]_{L_{tr} \times L_{test}} + b x[1 \dots 1]_{1 \times L_{test}} \quad (6)$$

Hibrit modelin son aşamasında, ters ölçekleme (Denklem 7) ile normalleştirilmiş model çıktıları ($Q_{m,s}$) orijinal hedeflerle aynı birime sahip olan akış tahminlerine (Q_{model}) geri dönüştürülür.

$$Q_{model,t} = \mathbf{Q}_{m,s,t} \times (Q_{obs,max} - Q_{obs,min}) + Q_{obs,min}, \quad t = 1, 2, \dots, L \quad (7)$$

İşleyişi Denklem 1-7'de verilen ve beş parametre içeren bu hibrit model ile iki parametrelili DVR ve dört parametrelili *dynwbm* modellerinin kalibrasyonunda hibrit parçacık sürü optimizasyonu (HPSO) adı verilen bir optimizasyon algoritması kullanılmıştır. Algoritma hakkındaki detaylı bilgiye Ersoy (2021) ve Okkan ve Kirdemir (2020) çalışmalarından ulaşılabilir.

2.4. Performans Kriterleri

Çalışmada, kalibrasyon ve validasyon (doğrulama) dönemlerinde irdelenen toplam karesel hata veya hata kareler ortalamasının karekökü (HKOK) gibi ölçütlere ilaveten Nash-Sutcliffe katsayısı (NS) değerleri de sorgulanmıştır. Modeller, NS performansları açısından aşağıdaki gibi derecelendirilmektedir (Moriasi vd., 2007; Okkan vd., 2021; Ersoy, 2021):

Çok iyi: $0.75 < NS \leq 1$

İyi: $0.65 < NS \leq 0.75$

Yeterli: $0.50 < NS \leq 0.65$

Yetersiz: $NS \leq 0.50$

Çalışmada, Nash ve Sutcliffe (1970) tarafından önerilen NS tabanlı r^2 istatistikleri ile *dynwbm*'ye alternatif olarak uygulanan modellerin *dynwbm*'den ne ölçüde farklılık arz ettiği de sorgulanmıştır. Bu karşılaştırma ölçütü Denklem 8 ile ifade edilebilmektedir.

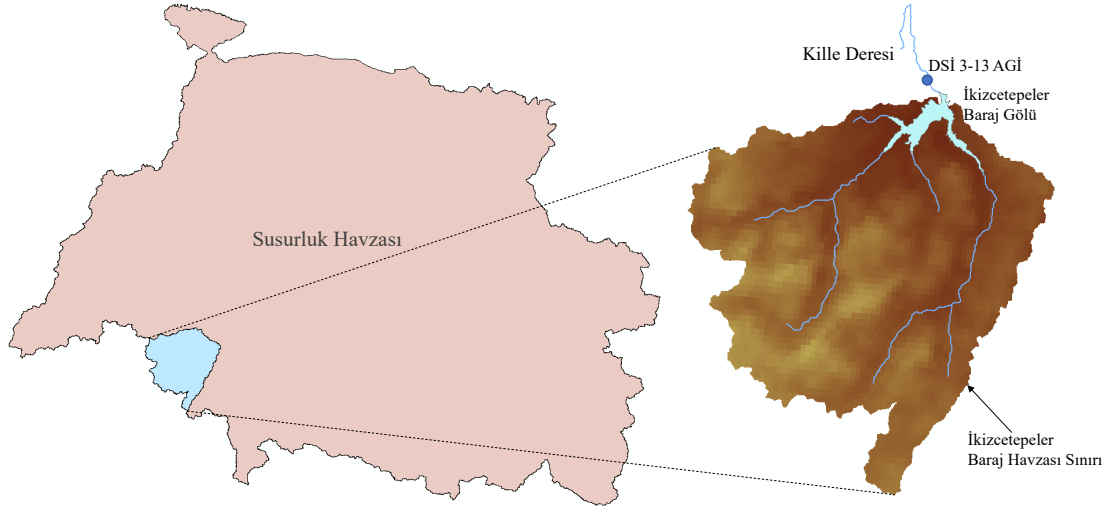
$$r^2 = \frac{NS_2 - NS_1}{1 - NS_1} \quad (8)$$

Burada NS_1 ve NS_2 , sırasıyla *dynwbm* ve uygulanan herhangi bir modelden elde edilen NS katsayılarıdır. Senbeta vd. (1999) çalışmasına göre, r^2 'nin %10'dan büyük olması halinde önerilen modelin karşılaştırılan referans modele kıyasla anlamlı iyileşme öngördüğü söylenebilmektedir.

3. ÇALIŞMA ALANI

Çalışmada uygulama alanı olarak İkizcetepeler Baraj Havzası seçilmiştir. Bu baraj Balıkesir kentinin içmesuyu ihtiyacını karşılamak amacıyla Susurluk Havzası sınırları içinde Kille Çayı üzerinde inşa edilmiş ve yaklaşık 500 km² drenaj alanına sahip bir barajdır. Baraj 1700 hektarlık bir alana yıllık 13,7 hm³'lük sulama hizmeti verirken yıllık 72 hm³'lük de içme suyu sağlamaktadır. İkizcetepeler baraj havzasına ait DSİ 3-13 İkizcetepeler akım gözlem istasyonu (AGİ) verileri Devlet Su İşleri (DSİ) Genel Müdürlüğü'nden, havzayı temsil eden

17150 kodlu meteoroloji istasyonuna ait meteorolojik gözlemler ise Meteoroloji Genel Müdürlüğü (MGM)'nden temin edilmiştir. Çalışmada baraj inşası öncesinde ölçülen 1965-1994 su yılları arasındaki akımlar ve aynı döneme ait aylık toplam yağış ve referans potansiyel evapotranspirasyon değerleri kullanılmıştır. MGM-17150 istasyonu baraj havzasına en yakın olan ve verileri eksiksiz olan tek istasyondur. 1965-1994 dönemi verilerine göre, yıllık toplam yağış, potansiyel evapotranspirasyon ve akış ortalamaları sırasıyla 556, 785 ve 210 mm civarında olup aridite indeksine göre havzada yarı-nemli iklim karakteristiği hakimdir. Değerlendirilen akım gözlem istasyonu ve baraj havzasının konumları ise Şekil 3'te sunulmaktadır (MGM istasyonu havzanın dışında yer aldığından ve havza içerisinde başka bir MGM istasyonu bulunmadığından Şekil 3'te sadece akım gözlem istasyonu belirtilmiştir).



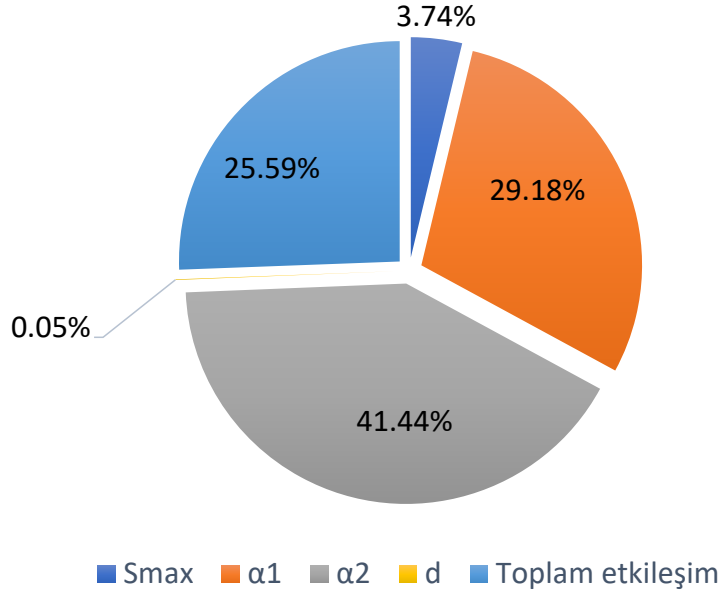
Şekil 3:
Uygulama havzasının, akım gözlem istasyonunun ve baraj gölünün konumu

4. BULGULAR

Çalışma kapsamında tüm modeller paket yazılım kullanılmadan MATLAB 2016a ortamında kodlanmıştır. Çalışmada önerilen seri hibridizasyonun simülasyon performansının etkinlik düzeyi değerlendirilirken tekil *DVR* ve *dynwbm* modeli çıktıları kıyaslama maksadıyla kullanılmıştır. Fakat öncesinde, hibrit tipteki modelin içindeki kavramsal kısımda taban akışı bileşeninin ortadan kaldırılmasına esas teşkil eden hassasiyet analizi Bölüm 4.1'de sunulmuştur. Akabinde, seri hibrit modelin nasıl kurulduğu Bölüm 4.2'de detaylandırılmış ve tüm modellerin simülasyon performanslarının karşılaştırılması Bölüm 4.3'te verilmiştir.

4.1. KYA Modelinde Parametre Hassasiyet Analizi

dynwbm'nin kalibrasyon ve validasyonunun akabinde, model parametrelerinin akış simülasyon performansı üzerindeki bireysel ve etkileşimli etkilerini ölçmek için ANOVA (varyans analizi) olarak adlandırılan hassasiyet analizine bağlı bir belirsizlik ayrıştırma yöntemi kullanılmıştır. Dört yönlü ANOVA ile kısmi varyanslar ayrıştırıldıktan sonra, ana kaynaklara ve etkileşim terimlerine ilişkin katkı sağlayan varyans oranları elde edilmiş olup, bunlar Şekil 4'te sunulmuştur. Şekil 4 incelendiğinde α_1 ve α_2 parametrelerinin *dynwbm* performansı üzerindeki hassasiyeti daha baskınken, taban akışını kontrol eden *d* parametresindeki değişimlerin akımlar üzerindeki etkisinin daha az olduğu görülebilmektedir.



Şekil 4:
Ana kaynaklar ve toplam etkileşim için varyans oranları

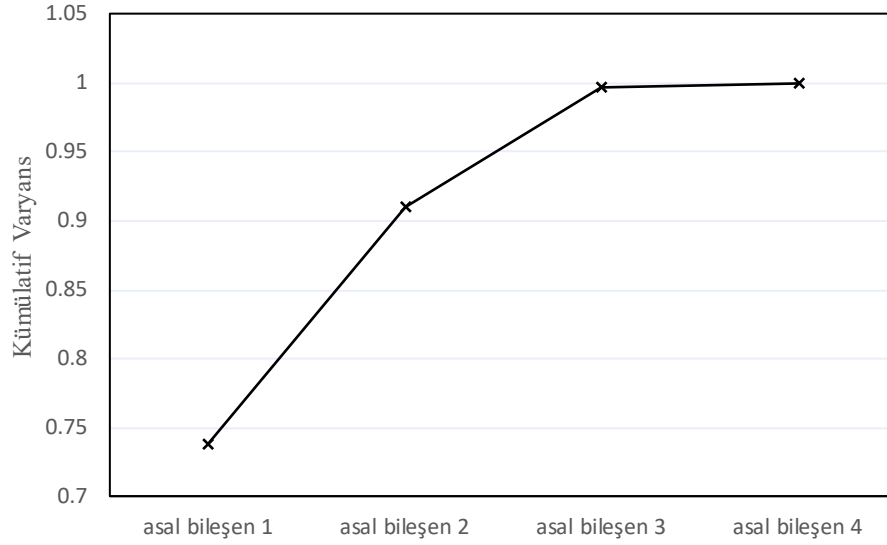
4.2. Seri Hibrit Modelin Kurulması

Bölüm 4.1’de belirtilen bulgular, taban akışı bileşeninin mevcut modelde açıklayıcı olmadığını ve ilgili sürecin geliştirilmesi gerektiğini göstermektedir. Bu nedenle Bölüm 2.3’te belirtildiği gibi, *dynwbm*’nin yeraltı suyu depolama elemanının yerini alabilecek bir makine öğrenmesi hibridizasyonunun daha başarılı simülasyonlar verip vermediği sorgulanmıştır. Önerilen hibritleme yönteminde, Bölüm 2.3’te de ifade edildiği üzere kavramsal modelin *Q_{dolaysız}*, *Sızan*, *S* ve *E_{gerçek}* çıktıları ML kısmı içinde işletilmiştir. Bununla birlikte gözlenmiş akımlar ile bazı kavramsal çıktılar arasında güçlü korelasyonların mevcut olduğu da bilinmektedir (Ersoy 2021; Okkan vd. 2021).

Fakat tüm olası değişkenlerin kullanımı sonrası doğabilecek çoklu doğrusallık hibrit şema içine eklenen ML bölümünün zayıf genelleme performansı göstermesine de neden olabilecektir. Öte yandan, ML operatöründeki giriş katmanına gereğinden fazla girdi yönlendirilmesi de kalibrasyon işlemini zorlaştırmaktadır. Bu aşamada ilişkili kavramsal çıktıları aşırı doğrusallıktan arındıran ve bunların boyutunu azaltabilen bir veri ön işleme tekniğine ihtiyaç duyulmuştur. Bu kapsamda Şekil 2’de belirtilen dört adet kavramsal çıktının ML kısmına yönlendirilmeden önce TBA’ya tabi tutulması uygun bulunmuştur. *d-DVR* modelinde, Şekil 5’te görüldüğü üzere, ilk iki temel bileşen varyansın büyük kısmını açıklamış (>0,9) ve dolayısıyla ML bölümünün giriş katmanına yönlendirilen değişken sayısı iki olarak alınmıştır.

4.3. Modellerin Kıyaslanması

Çalışmada modeller 1965-1979 su yılı verileri ile eğitilmiş 1980-1994 dönemi verileri ile test edilmiştir. Karşılaştırılan modellerin kalibrasyon ve validasyon dönemi performansları Tablo 1’de sunulmuştur. Tablo 1’de koyu hücreler elde edilen en iyi sonuçları temsil etmektedir.



Şekil 5:
TBA'da varyansın açıklanabilme mertebesi

Tablo 1. Model performanslarının karşılaştırması

(a) Kalibrasyon

Modeller	tüm veri		0,75 ve üstü yüzdellik	
	HKOK (mm)	NS	HKOK (mm)	NS
dynwbm	12,60	0,76	21,64	0,31
DVR	15,23	0,66	25,11	0,07
d-DVR	10,13	0,85	16,35	0,61

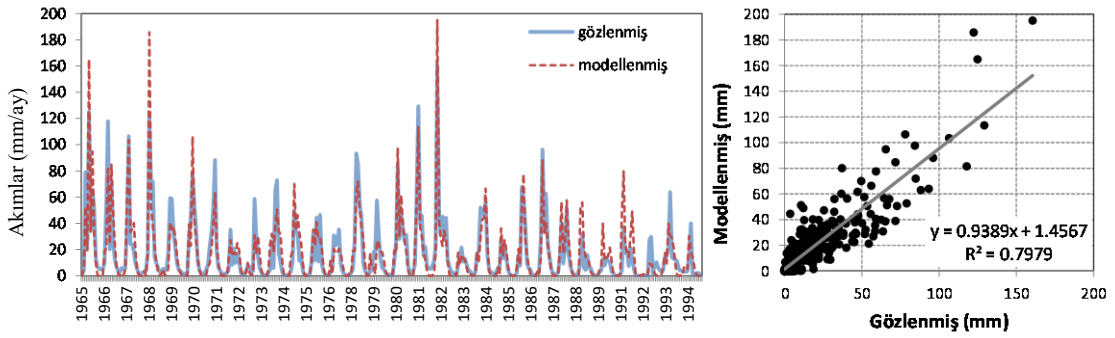
(b) Validasyon

Modeller	tüm veri		0,75 ve üstü yüzdellik	
	HKOK (mm)	NS	HKOK (mm)	NS
dynwbm	10,48	0,78	15,63	0,68
DVR	15,10	0,55	19,49	0,50
d-DVR	9,16	0,83	15,03	0,70

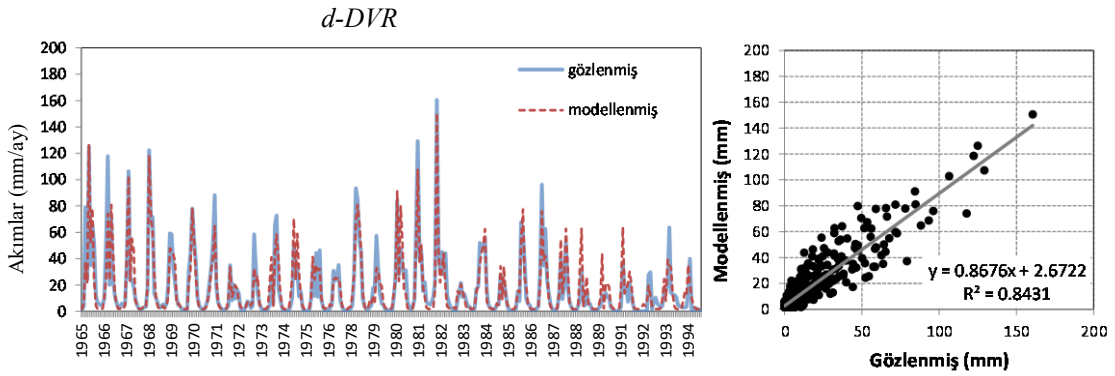
Gözlenen ve modellenen akış arasındaki hatanın bir ölçüsü olan HKOK istatistikleri de Tablo 1'de görüldüğü gibi her bir model için hesaplanmıştır. Öncelikle tüm veri noktaları için değerlendirme yapılmıştır. Buna göre, hibrit model, *dynwbm*'ye kıyasla kalibrasyon ve validasyon dönemlerinde sırasıyla %21 ve %14 daha düşük HKOK değerlerine sahiptir. Aynı değerlendirme *DVR* modeli ile yapıldığında bu dönemlerde HKOK değerlerinde %23 ve %47 artış görülmüştür. Dolayısı ile *DVR* tabanlı tekil ML modelinin kavramsal modelin performansını yansıtabilecek düzeyde olmadığı söylenebilmektedir.

Tablo 1'de belirtilen NS indisleri Moriasi vd. (2007) tarafından önerilen performans derecelendirmesi açısından irdelenmiştir. Tüm veri noktaları için yapılan değerlendirmeye göre, ML tekniği olan *DVR*'nin kalibrasyon ve validasyonda 'yeterli' kategorisinde olduğu görülmektedir. Öte yandan, önerilen seri hibrit model ve kavramsal model iki dönemde de 'çok

iyi' olarak derecelendirilse de *d-DVR* daha yüksek sonuçlar vermiştir. Kavramsal model ve hibrit modelin aynı kategoride olması sebebiyle, önerilen bu modelin referans model olan *dynwbm*'ye göre iyileşme gösterip göstermediğini tespit etmek için r^2 istatistikleri hesaplanmış ve *dynwbm*'ye alternatif olarak uygulanan modelin ne düzeyde farklılık sergilediği irdelenmiştir. %35 ve %24 olarak hesaplanan r^2 değerleri, Senbeta vd. (1999) tarafından önerilen %10'dan daha büyük olduğu için hibrit modelin kayda değer bir gelişme sağladığı söylenebilmektedir. Elde edilen bu bulgulara ek olarak *dynwbm* ve seri hibrit modelin toplam gözlem periyodu için (kalibrasyon ve validasyon birlikte olacak şekilde) gidişler ve saçılım grafikleri de sırasıyla Şekil 6 ve Şekil 7'de sunulmuştur. Buradan *dynwbm* modelinin yüksek akım koşullarında aşırı tahmin verme eğiliminde olduğu, *d-DVR*'nin ise bu bakımdan çok daha tutarlı davranış sergilediği ve ayrıca $y=x$ doğrusu etrafındaki sapmalarının *d-DVR* modelinde daha az olduğu görülebilmektedir.



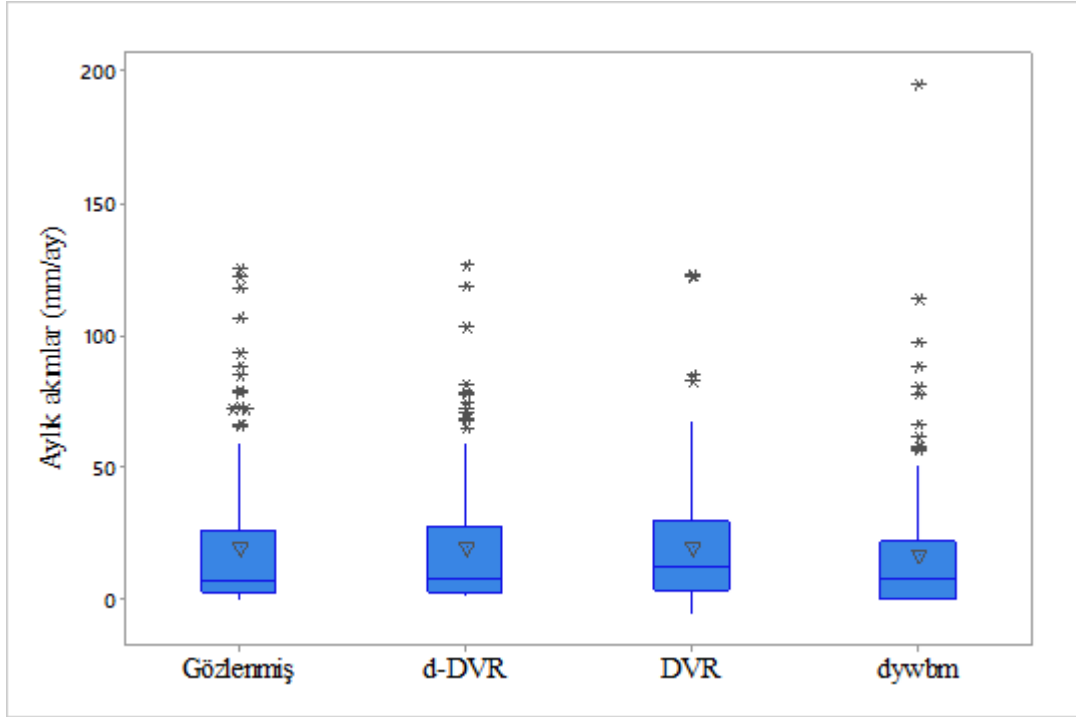
Şekil 6:
dynwbm modelinden üretilen akımlara ait gidiş ve saçılım grafikleri.



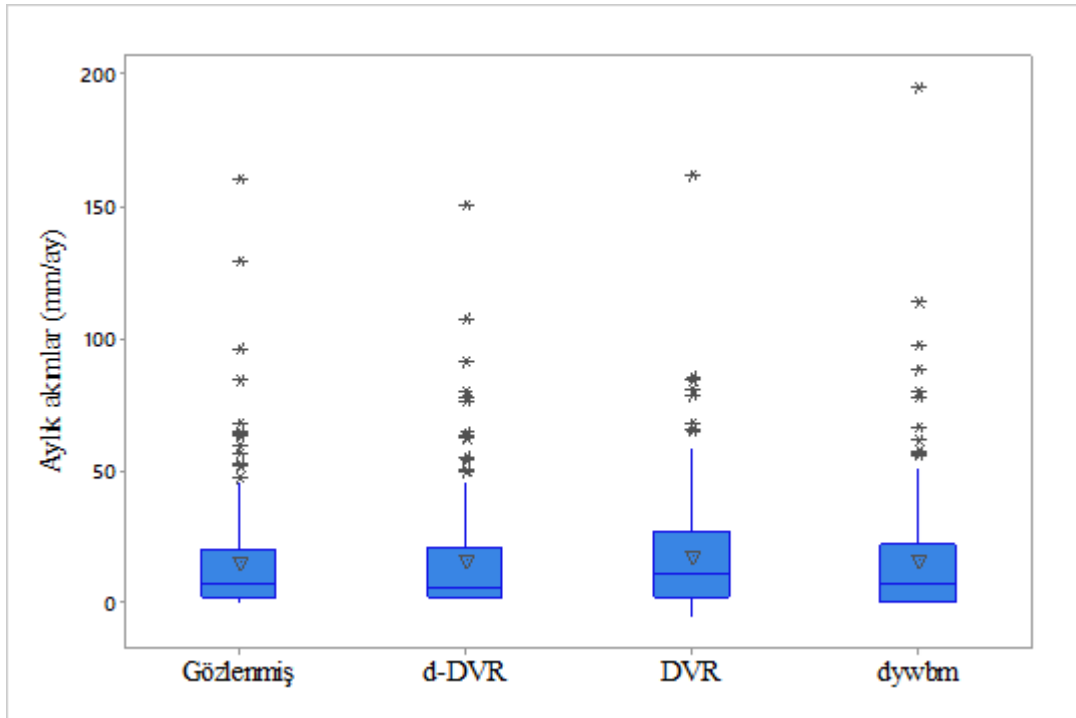
Şekil 7:
d-DVR hibrit modelinden üretilen akımlara ait gidiş ve saçılım grafikleri

Çalışmada ayrıca gözlenmiş akımların 0,75 ve üstü yüzdelerdeki değerleri çıkarılarak yüksek akım koşullarında model performanslarının sorgulanması da sağlanmıştır. Tablo 1'de bu koşullar için hesaplanan NS ve HKOK istatistikleri ayrıca belirtilmiştir. Sonuçlar *d-DVR* modelinin yüksek akım performanslarında da daha üstün olduğunu tekrar göstermiştir. Ayrıca modellerin kalibrasyon ve validasyon dönemindeki tahminleri kutu diyagramı ile de gösterilerek kartillerdeki uyumlar görsel olarak ortaya konmaya çalışılmıştır (Şekil 8). Buna göre, *d-DVR* modeli tüm kartillerde gözlenen verileri daha iyi benzeştirdiği anlaşılmaktadır.

(a) Kalibrasyon



(b) Validasyon



Şekil 8:

Modellerin (a) kalibrasyon ve (b) validasyon dönemi tahminlerinin kutu diyagramı ile gösterimi (kutulardaki üçgen sembolleri ortalama değeri göstermektedir).

5. SONUÇLAR

Çalışmada ortaya konulduğu üzere, *dynwbm* kavramsal yağış-akış modeli çeşitli parametreler vasıtasıyla benzeştirdiği havza yağış-akış ilişkisini *DVR* tabanlı makine öğrenmesi modeline kıyasla hem kalibrasyon hem de validasyon dönemlerinde daha güvenilir bir biçimde modellemiştir. Ancak kavramsal modelden elde edilen akım çıktılarında (özellikle yüksek akım koşullarında) bazı sistematik yanlışlıklar gözlenmiştir. Bu yanlışlıkların modelin yeraltı suyu depolamasının doğrusal parametreler ile kavramsallaştırılmasından kaynaklandığı düşünülmektedir. Bu tespit Ersoy (2021) ve Okkan vd. (2021) çalışmalarında özellikle vurgulanmıştır. Bu gerekçeden dolayı, *DVR* modelinin kavramsal modelin içine entegre edilerek yeraltı suyu depolaması bileşeninin elimine edilmesi ve simülasyon hassasiyetinin artırılması hedeflenmiştir. Bu hedef doğrultusunda iki farklı model yapısını kontrol eden parametrelerin birlikte optimizasyonunun nasıl bir etki sağladığı hibrit bir şema ile ortaya konulmuştur. Kavramsal bir yağış-akış modeli ile makine öğrenmesine dayanan veri güdümlü bir modelin ilk kez seri bir şema ile entegre edildiği bu çalışma gerçek bir hibrit model niteliği taşımaktadır. *d-DVR* olarak adlandırılan seri hibrit yapı HKOK ve NS gibi klasik performans ölçütleri ile yapılan sınamalara göre *dynwbm* ve *DVR* modellerinden daha üstün sonuç vermiştir. Bunun ana nedeni hibrit şema içerisine entegre edilen kavramsal modelin ve *DVR*'nin aynı iterasyon adımında iç içe çalıştırılması ve bütünleşik modelin değişken parametre varyasyonlarını dikkate alabilmesidir. Böylelikle kalibrasyon ve validasyon sonrasında modelin ürettiği zemin nemi içeriği ve gerçek evapotranspirasyon gibi çıktılar yeraltısuyu beslenimi, kuraklık analizi vb. uygulamalar için daha sağlıklı girdi niteliği taşımaktadır. Gelecekte önerilen hibritleme tarzı farklı kavramsal modellerin ve ML tekniklerinin entegrasyonunu kapsayacak şekilde farklı zaman ölçeklerinde farklı havzalara uygulanacak olup, bu konu ile ilgili çalışmalarımız devam etmektedir.

ÇIKAR ÇATIŞMASI

Yazarlar, bilinen herhangi bir çıkar çatışması veya herhangi bir kurum/kuruluş ya da kişi ile ortak çıkar bulunmadığını onaylamaktadırlar.

YAZAR KATKISI

Tüm yazarlar çalışmanın veri toplama, veri analizi ve yorumlama, makale taslağının oluşturulması, bilgisayar yazılımlarının hazırlanması, kavramsal ve tasarım süreçlerinin belirlenmesi, bu süreçlerin yönetilmesi ve fikrinsel içeriğin eleştirel incelemesi aşamalarında yer aldıklarını, makalenin son onayını ve tam sorumluluğunu kabul ederler.

KAYNAKLAR

1. Anctil, F., Michel, C., Perrin, C. ve Andréassian, V., (2004) A soil moisture index as an auxiliary ANN input for stream flow forecasting, *Journal of Hydrology*, 286 (1-4), 155–167. doi: 10.1016/j.jhydrol.2003.09.006
2. Ersoy, Z.B. (2021) Dinamik Su Bütçesi Modeline Makine Öğrenmesi Entegrasyonu ile Aylık Akış Tahminlerinin İyileştirilmesi. Balıkesir Üniversitesi Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, Balıkesir.
3. Ersoy, Z.B., Okkan, U., ve Fıstıkoğlu, O. (2021) Hybridizing a Conceptual Hydrological Model with Neural Networks to Enhance Runoff Prediction, *Manchester Journal of Artificial Intelligence and Applied Sciences*, 02, 176-178.

4. Humphrey, G.B., Gibbs, M.S., Dandy, G.C., ve Maier, H.R., (2016) A hybrid approach to monthly streamflow forecasting: Integrating hydrological model outputs into a Bayesian artificial neural network, *Journal of Hydrology*, 540, 623–640. doi: 10.1016/j.jhydrol.2016.06.026
5. Kumanlioglu, A.A., ve Fistikoglu, O., (2019) Performance enhancement of a conceptual hydrological model by integrating artificial intelligence, *Journal of Hydrologic Engineering*, 24 (11), 04019047. doi: 10.1061/(ASCE)HE.1943-5584.0001850
6. Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D., ve Veith, T.L., (2007) Model evaluation guidelines for systematic quantification of accuracy in watershed simulations, *Transactions of the ASABE*, 50(3), 885-900. doi: 10.13031/2013.23153
7. Nash, J.E., ve Sutcliffe, J.V., (1970) River flow forecasting through conceptual models, Part I- A discussion of principles, *Journal of Hydrology*, 10(3), 282-290. doi: 10.1016/0022-1694(70)90255-6
8. Noori, N., ve Kalin, L., (2016) Coupling SWAT and ANN models for enhanced daily streamflow prediction, *Journal of Hydrology*, 533, 141–151. doi: 10.1016/j.jhydrol.2015.11.050
9. Okkan, U., Ersoy, Z.B., Kumanlioglu A.A., ve Fistikoglu, O. (2021) Embedding machine learning techniques into a conceptual model to improve monthly runoff simulation: a nested hybrid rainfall-runoff modeling, *Journal of Hydrology*, 598, 126433. doi: 10.1016/j.jhydrol.2021.126433
10. Okkan, U., ve Kirdemir, U. (2020) Towards a hybrid algorithm for the robust calibration of rainfall–runoff models, *Journal of Hydroinformatics*, 22(4), 876-899. doi: 10.2166/hydro.2020.016
11. Okkan, U., ve Serbes, Z.A., (2012) Rainfall-runoff modeling using least squares support vector machines, *Environmetrics*, 23(6), 549–564. doi: 10.1002/env.2154
12. Senbeta, D.A., Shamseldin, A.Y., ve O’Connor, K.M., (1999) Modification of the probability-distributed interacting storage capacity model, *Journal of Hydrology*, 224(3-4), 149–168. doi: 10.1016/S0022-1694(99)00127-4
13. Tongal, H., ve Booi, M.J., (2018) Simulation and forecasting of streamflows using machine learning models coupled with base flow separation, *Journal of Hydrology*, 564(2018), 266-282. doi: 10.1016/j.jhydrol.2018.07.004
14. Zhang, L., Potter, N., Hickel, K., Zhang, Y., ve Shao, Q., (2008) Water balance modeling over variable time scales based on the Budyko framework - Model development and testing, *Journal of Hydrology*, 360(1-4), 117–131. doi: 10.1016/j.jhydrol.2008.07.021

