Bölüm 2

Rüzgar Türbini Aktif Güç Üretiminin Makine Öğrenmesi ve Derin Öğrenme Yöntemleriyle Tahmini

İzzet Yavuz[1]

Elif Öztürk[2]

Özet

Rüzgâr enerjisinin elektrik arzındaki payının büyümesiyle birlikte, üretimin kısa ve orta vadeli tahminlerinin doğru yapılması ihtiyacı ortaya çıkmaktadır. Bu bölümde, Türkiye'nin Marmara Bölgesi'nde faaliyet gösteren bir rüzgâr türbinine ait, bir yıllık (2018), 10 dakikalık gerçek SCADA ölçümlerinin (n=50.530) analizi, aktif gücü 1, 6 ve 24 saat sonrası için tahmin eden beş yöntemle (Naive Bayes, Random Forest, XGBoost, Uzun-Kısa Vadeli Bellek ağı ve Transformer) gerçekleştirilmektedir. Ölçekleme ve yüzdelik tabanlı rüzgâr eşikleri, sızıntıyı önlemek amacıyla yalnızca eğitim istatistiklerinden türetilmiştir. Bulgular, 1 saatlik ufukta hiçbir modelin Süreklilik (persistence) taban çizgisini istatistiksel olarak anlamlı biçimde geçemediğini (Diebold-Mariano p=0,44) göstermekte, buna karşılık 24 saatlik ufukta XGBoost'un persistence'a göre istatistiksel olarak anlamlı bir iyileşme sağladığını ($p<0,001$) ortaya koymaktadır; ancak açıklanan varyans pratikte sınırlı kalmaktadır ($R^2 \approx 0,00$). Nitekim 6 ve özellikle 24 saatlik ufuklarda hiçbir yöntem klasik anlamda gerçek bir tahmin becerisi sunmamaktadır. SHAP (SHapley Additive exPlanations) analizi, tahminlerin kısa ufuklarda ağırlıklı olarak en son gözlemlenen güç değerine, uzun ufuklarda ise esasen iklimsel bir temel çizgi işlevi gören mevsimsel (yılın günü) örüntülere dayandığını doğrulamaktadır. Bulgular, dışsal sayısal hava tahmini girdisi olmaksızın, yalnızca geçmiş SCADA kayıtlarına dayanan istatistiksel ve öğrenme temelli yöntemlerin doğal bir öngörülebilirlik ufku sınırıyla karşı karşıya olduğunu göstermektedir. Makine öğrenmesi bu sınırı ortadan kaldırmaktan çok, sınırın kendisini nicel olarak görünür kılmaktadır.

1 Öğretim Görevlisi, Ankara Üniversitesi Nallıhan Meslek Yüksekokulu, Elektrik Enerjisi Üretim, İletim ve Dağıtımı Programı, izzetyavuz@ankara.edu.tr, ORCID ID: 0000-0002-2309-3815

2 Araştırma Görevlisi, İstanbul Gelişim Üniversitesi Mühendislik ve Mimarlık Fakültesi, Elektrik Elektronik Mühendisliği (İngilizce), elozturk@gelisim.edu.tr, ORCID ID: 0009-0001-4006-3458

https://doi.org/10.58830/ozgur.pub1529.c6171

1. Giriş

Rüzgâr enerjisi, kurulu kapasitedeki hızlı büyümenin de etkisiyle birçok ülkenin elektrik üretim portföyünde giderek daha belirleyici bir rol oynamaktadır. Türkiye'de de rüzgâr enerjisi santrallerinin (RES) kurulu gücü son on yılda istikrarlı biçimde artmıştır. 2024 yılı sonu itibarıyla toplam kurulu rüzgâr gücü yaklaşık 13 GW'a ulaşmış, ülkenin 50 metre yükseklikte 6,8 m/s üzeri rüzgâr hızları için teknik rüzgâr potansiyeli 118,683 MWe olarak hesaplanmıştır (Sogukpinar vd., 2025). Rüzgâr enerjisindeki bu hızlı büyüme, üretim dalgalanmalarının önceden öngörülebilmesinin önemini daha da artırmaktadır. Doğası gereği değişken ve kesintili olan rüzgâr üretimi, saatlik ve günlük tahminlerin doğruluğunu hem sistem güvenilirliği hem de enerji sisteminin ekonomik işletimi açısından zorunlu kılmaktadır. Tahmin hataları ya da düşük doğruluk, yedek üretim planlamasında daha yüksek maliyetlere, dengeleme piyasasında cezalara yol açmaktadır. Elektrik-elektronik ve haberleşme mühendisliği açısından bakıldığında, rüzgâr güç tahmini yalnızca istatistiksel bir modelleme meselesi değil, aynı zamanda birçok mühendislik alanının kesişiminde yer alan ve SCADA (Denetleyici Kontrol ve Veri Toplama) sistemleriyle doğrudan bağlantılı bir uygulama alanıdır.

Son dönemde rüzgâr ve güneş enerjisi tahmininde ML ve DL tekniklerinin yaygın biçimde kullanılması, geleneksel istatistiksel zaman serisi modellerinin (ARIMA ve üstel düzeltme gibi) doğrusal olmayan ve karmaşık örüntüleri hesaba katamamasından kaynaklanmaktadır (Rajaperumal ve Christopher Columbus, 2025). Random Forest ve XGBoost gibi ağaç tabanlı topluluk yöntemleri, tablo biçimli SCADA verisi üzerinde sağlam ve anlaşılır çözümler olarak iyi performans gösterirken, LSTM sinir ağları ve son dönemde Transformer tabanlı derin sinir ağı modelleri , zaman serilerindeki uzun vadeli bağımlılıkları yakalama kapasiteleriyle öne çıkmaktadır. Bununla birlikte, mevcut literatür bu yöntemlerin göreli başarısının tahmin ufkuna (ufuk uzadıkça öngörülebilirliğin nasıl değiştiğine) ve girdi verisinin kapsamına (yalnızca geçmiş SCADA verisi mi, yoksa NWP verisini de içeriyor mu) bağlı olarak önemli ölçüde değiştiğini göstermektedir. Dolayısıyla, tek bir modelin diğerinden üstün kabul edilemeyeceği ve ufka ve veri kapsamına duyarlı karşılaştırmalı çalışmalara ihtiyaç duyulduğu açıktır.

Literatürde sıklıkla karşılaşılan bir diğer sorun, model karşılaştırmalarının çoğunlukla yalnızca nokta tahmin metriklerine (MAE ve RMSE gibi) dayanması ve gözlemlenen farkların istatistiksel olarak anlamlı olup olmadığının test edilmemesidir. Bir modelin sayısal olarak diğerinden daha düşük hata vermesi, bu farkın örneklem gürültüsünden ayırt edilebilir olduğu anlamına gelmez. Özellikle yüksek örneklem hacmine sahip ama güçlü otokorelasyon taşıyan

enerji tahmini gibi zaman serilerinde, görünürdeki "iyileşmeler" istatistiksel testten geçirilmediğinde yanıltıcı sonuçlara yol açabilmektedir. Bu boşluk göz önünde bulundurularak, bu bölüm model karşılaştırmalarını yalnızca tanımlayıcı metriklerle değil, biçimsel anlamlılık testleriyle de desteklemeyi amaçlamaktadır.

Bu bölümün amacı; gerçek, tek bir rüzgâr türbinine ait, kamuya açık SCADA verisi kullanılarak, istatistiksel olarak sağlam bir çerçevede (kronolojik ayırım, sızıntısız ölçekleme, Diebold-Mariano anlamlılık testi ve blok bootstrap güven aralıkları ile) (i) Naive Bayes bir persistence taban çizgisinin, (ii) iki ağaç tabanlı topluluk yönteminin (Random Forest, XGBoost) ve (iii) iki derin öğrenme mimarisinin (LSTM, Transformer) 1, 6 ve 24 saatlik ufuklardaki aktif güç tahmin performansını karşılaştırmaktır. Başlangıçta belirtmek gerekir ki bu karşılaştırmanın en dikkat çekici bulgusu, hangi yöntemin "kazandığı" değil, ufuk uzadıkça -özellikle 6 ve 24 saatte- hiçbir yöntemin dışsal meteorolojik girdi olmaksızın gerçek bir tahmin becerisi sergilememesi, bunun yerine giderek mevsimsel bir temel çizgiye yakınsamasıdır. Bu bulgu Bölüm 4 ve 5'te nicel olarak ortaya konmaktadır. Ayrıca, XGBoost modelleri üzerinde SHAP (SHapley Additive exPlanations) analizi uygulanarak, modellerin farklı ufuklarda hangi değişkenlere dayandığı yorumlanabilir biçimde ortaya konmaktadır. Bu çalışma, elektrik-elektronik ve haberleşme mühendisliği alanında yenilenebilir enerji entegrasyonu ile makine öğrenmesinin kesişiminde, tamamen gerçek veri ve hesaplamalara dayanan, tekrarlanabilir bir uygulama örneği sunmayı amaçlamaktadır. Çalışmanın üç temel katkısı şöyledir: (i) beş farklı yöntem sınıfının -bir istatistiksel taban çizgisi, ağaç tabanlı iki topluluk yöntemi ve iki derin öğrenme modeli- aynı veri seti, aynı eğitim-test ayırımı ve aynı değerlendirme prosedürüyle doğrudan karşılaştırılması; (ii) modeller arasındaki farkların istatistiksel testlerle (Diebold-Mariano testi ve blok bootstrap) kanıtlanması; ve (iii) SHAP yorumlama analizinin, yalnızca hangi modelin en iyi olduğuna değil, neden öyle çalıştığına da ışık tutacak biçimde kullanılması.

Çalışmanın geri kalanı şu şekilde düzenlenmiştir: İlgili literatürün gözden geçirilmesi Bölüm 2'de sunulmaktadır. Bölüm 3'te veri seti, veri ön işleme, öznitelik çıkarımı, model mimarisi ve değerlendirme metodolojisi ele alınmaktadır. Bulgular; Bölüm 4'te sunulmuş olup Bölüm 5'te ise literatür ışığında yorumlanmaktadır. Sonuçlar Bölüm 6'da verilmektedir. Bölüm 7'de çalışma sınırlılığı hakkında bilgiler verilmiştir.

2. Literatür Taraması

Son beş yılda rüzgâr ve güneş üretimi tahminine ilişkin çalışmaların taranması, makine öğrenmesi (ML) ile derin öğrenme (DL) arasındaki karşılaştırmanın giderek artan bir eğilim olduğunu ortaya koymuştur. Rajaperumal ve Christopher Columbus (2025), yalnızca derin öğrenme algoritmalarını değil, klasik makine öğrenmesi algoritmalarını (Random Forest, Support Vector Regression) da topluluk yöntemleri ve diğer tekniklerle birlikte, aralarında bu çalışmada da kullanılan Türkiye rüzgâr türbinleri Kaggle SCADA veri tabanının da bulunduğu birden çok veri seti türü üzerinde analiz etmiştir. Yazarlar, parametrelerin doğru ayarlanmasının ve model seçiminin, tek bir evrensel yapıya değil, veri setinin özelliklerine ve tahmin ufkuna bağlı olduğunu ileri sürmektedir.

Mazen (2025), aynı SCADA verisini farklı bir açıdan ele alarak, güç tahmininden önce iki aşamalı bir aykırı değer tespiti (One-Class SVM tabanlı) uygulamıştır. Çalışma, rüzgâr hızının küpü ve artık güç gibi fizik temelli özniteliklerin tahmin performansını artırdığını bulmuştur. Bu sonuçlar, SCADA verisinin hem aykırı değerler hem de sansürlenmiş/bekleme durumundaki kayıtlar içerdiğine işaret etmesi bakımından, bu bölümün veri kalitesi bulgularıyla (Bölüm 3.1) tutarlıdır. Aykırı değerleri modele dahil edip yalnızca gerçek zaman boşluklarını NaN olarak ele alan mevcut yöntemin aksine, Mazen (2025) bunları ön işleme yoluyla ayıklamaya çalışmaktadır.

Rüzgâr hızı ve gücü tahmininde geleneksel ML algoritmalarının performansı yoğun biçimde incelenmiştir. Mugware vd. (2024)'e göre, Güney Afrika'nın farklı iklimlere sahip üç bölgesinde rüzgâr hızı tahmini için CNN, DAN2, Random Forest ve XGBoost modelleri kullanıldığında, en iyi algoritma bölgeye göre değişebilmektedir. Böylece hiçbir mimari tüm koşullarda üstünlük sağlayamamıştır. Bu sonuç, mevcut çalışmanın (Bölüm 4.1) model sıralamasının tahmin ufkuna bağlı olduğu yönündeki bulgusuyla örtüşmektedir. Hussain vd. (2025)'in bir başka çalışması, aylık ortalama rüzgâr hızı verisine dayalı bir XGBoost tahmin modelinin geliştirilmesini içermektedir; ayrıca yazarlar, hiperparametre optimizasyonunun (Bayesian hiperparametre optimizasyonu dahil) önemli rolünü vurgulamaktadır. Xiong vd. (2022), Bayesian hiperparametre ayarlamasına sahip XGBoost modelinin, özellikle olumsuz hava koşullarının neden olduğu rüzgâr rampalarını ve düşük rüzgâr hızı dönemlerini tahmin ederken SVM, KELM ve LSTM yöntemlerinden daha az tahmin hatası ürettiğini ileri sürmektedir. Bu bulgu, mevcut çalışmada rüzgâr rejimine göre hata dağılımına ilişkin elde edilen sonuçlarla (Bölüm 4.3) örtüşmektedir. Son olarak, Mollasalehi ve Farhadi (2025) güncel literatür derlemelerinde LSTM, Random Forest ve XGBoost modellerini rüzgâr ve

güneş güç tahmini için karşılaştırmış ve her yöntemin kendine özgü güçlü ve zayıf yönleri olduğu, hiçbirinin her durumda daha iyi olmadığı sonucuna varmıştır. Bu bölüm, varılan sonuca ampirik bir katkı sunmaktadır.

Ağaç tabanlı topluluk öğrenmesi yöntemleri, kökenini Random Forest'tan (Breiman, 2001) almakta ve daha sonra gradyan artırmalı ağaçların ölçeklenebilir bir uygulaması olan XGBoost (Chen ve Guestrin, 2016) ile genişlemektedir. Bu yöntemler, görece küçük tablo biçimli veri setlerinde hesaplama açısından verimli ve yorumlanabilir taban çizgileri sunmaktadır. Derin öğrenme modelleri arasında LSTM'ler (Hochreiter ve Schmidhuber, 1997), kapı mekanizmalarına dayanarak uzun zaman serilerindeki bağımlılıkları geleneksel tekrarlayan ağlara kıyasla daha etkili biçimde yakalamaktadır. Daha yakın tarihli Transformer mimarisi (Vaswani vd., 2017) ise öz-dikkat mekanizmasıyla dizi adımları arasında doğrudan bağlantılar kurmakta ve böylece, özellikle uzun dizilerle çalışırken, tekrarlayan ağların sıralı işleme doğasından bağımsızlaşmaktadır. Zhou vd. (2021), Informer mimarisini geliştirerek Transformer'ların uzun dizili zaman serisi tahminindeki verimliliğini daha da artırmış ve bu, bu sınıftaki modellerin enerji dahil farklı uygulamalarda yaygınlaşmasına katkı sağlamıştır. Wu vd. (2022), zamansal füzyon Transformer'ının çok değişkenli rüzgâr hızı zaman serilerinde hem doğru hem de yorumlanabilir tahminler üretebildiğini kanıtlamıştır. Ancak Zeng vd. (2023), Transformer'ların zaman serisi tahminindeki etkinliğini açıkça sorgulamış ve birçok uzun ufuklu tahmin probleminde eklenen karmaşıklığın doğruluğu mutlaka artırmadığını, basit doğrusal modellerin bile Transformer tabanlı mimarileri geçebildiğini göstermiştir. Bu sonuç, mevcut çalışmanın bulgularıyla (Bölüm 4.1, 5) tutarlıdır; zira bu çalışma da Transformer'ın hiçbir ufukta ağaç tabanlı algoritmalardan mutlaka daha iyi performans göstermediğini ortaya koymaktadır. Enerji tahmini literatüründe Transformer tabanlı mimarilerin, özellikle daha uzun tahmin ufuklarında LSTM'lerle benzer düzeyde performans gösterdiği bildirilmekle birlikte, bu üstünlüğün evrensel olmadığı da belirtilmektedir. Nitekim bu durum mevcut çalışmanın bulgularıyla da (Bölüm 4) doğrulanmaktadır.

Yorumlanabilirlik açısından, kooperatif oyun teorisinden türetilen Shapley değerlerine dayanan SHAP çerçevesi (Lundberg ve Lee, 2017), karmaşık model tahminlerinin tutarlı biçimde değişken düzeyinde ayrıştırılmasına imkân tanımakta ve enerji tahmini uygulamalarında giderek yaygınlaşmaktadır. Türkiye'de Cakiroglu vd. (2024), 2011-2020 dönemine ait Çanakkale meteorolojik verileriyle altı regresyon modelini (XGBoost, Random Forest, LightGBM, CatBoost, AdaBoost, M5-Prime) karşılaştırmış ve SHAP analizi yardımıyla güç tahmini için en önemli değişkenin rüzgâr hızı olduğunu, ancak sıcaklık ve hava yoğunluğu gibi ikincil değişkenlerin de anlamlı katkı sağladığını bulmuştur. Dolayısıyla, bu bölümün SHAP bulgularının (Bölüm

4.4), Türkiye rüzgâr enerjisi alanındaki bu uygulama biçiminde doğrudan bir desteğe sahip olduğu söylenebilir. Rakip modellerin performans farklarının istatistiksel anlamlılığının değerlendirilmesinde ise klasik Diebold-Mariano testi (Diebold ve Mariano, 2002), iki rakip tahminin hata dağılımları arasındaki farkın tamamen rastlantısal olup olmadığını belirlemenin iyi bilinen bir yoludur; bununla birlikte enerji tahmini literatüründe modeller genellikle yalnızca MAE ve RMSE gibi nokta tahmin ölçütleri üzerinden, istatistiksel anlamlılık analizi yapılmaksızın karşılaştırılmaktadır. Bu çalışmada bu sorun, DM testi ve blok bootstrap güven aralıkları aracılığıyla kısmen ele alınmaktadır.

Literatür taraması, tahmin ufku uzadıkça model performansının sistematik biçimde bozulduğunu; bu bozulmanın, yalnızca geçmiş güç üretimi ve rüzgâr hızı bilgisine dayanan ve hava koşullarına ilişkin dışsal tahminlere dayanmayan modellerde daha belirgin olduğunu göstermiştir. Bu bölümde 24 saatlik tahmin ufkunda gözlemlenen sıfıra yakın açıklanan varyans ($R^2 \approx 0$), bu olguyla tutarlıdır.

Görece kısa sürelerde güç çıktısındaki belirgin değişimlerle karakterize edilen rüzgâr rampa olayları, literatürde ayrı bir alan oluşturmaktadır. Rampa olayları ilk olarak, optimize edilmiş bir "swinging door" algoritması uygulayan ve elde edilen öznitelikleri bir derin öğrenme modeline girdi olarak kullanan Han vd. (2020) tarafından ele alınmıştır. Yu ve Hur (2023) ise, rampa olayı tahminini değerlendirmek için kullanılan performans ölçütlerinin kendisinin de sorunlu olabileceğini ve Güney Kore'deki rüzgâr santralleri için geliştirilmiş bir değerlendirme ölçütüyle daha güvenilir bir rampa tahmini değerlendirmesi yapılabileceğini bildirmiştir. Bu bölümde rampa olayları ayrı bir kategori olarak modellenmemiş, bunun yerine rüzgâr hızının istatistiksel uç değerlerine (düşük ve yüksek rejimler) dayalı dolaylı bir yaklaşım benimsenmiştir (Bölüm 3.2, 4.3). Bu, Han vd. (2020) ile Yu ve Hur (2023)'un doğrudan rampa odaklı yaklaşımlarına kıyasla daha basit ama daha az özelleşmiş bir tercih olarak değerlendirilebilir.

Bu çalışmanın odağı rüzgâr enerjisi olmakla birlikte, güneş enerjisi tahmini literatürü de yöntemsel açıdan paralel dersler sunmaktadır. Sedai vd. (2023), uzun vadeli güneş enerjisi üretim tahmininde istatistiksel, ML ve DL modellerini karşılaştırmış ve ufuk uzadıkça tüm model ailelerinin performansının belirgin biçimde kötüleştiğini bildirmiştir. Bu bulgu, mevcut çalışmada rüzgâr gücü için gözlemlenen ufka bağlı performans bozulmasıyla (Bölüm 4.1) örtüşmekte ve enerji kaynağından (güneş ya da rüzgâr) bağımsız bir örüntüye işaret etmektedir. Türbin güç eğrisinin fiziksel temelleri açısından ise klasik rüzgâr enerjisi mühendisliği literatürü (Manwell vd., 2010), bu bölümde (Bölüm 4.3) gözlemlenen rejime bağlı hata örüntüsünün fiziksel

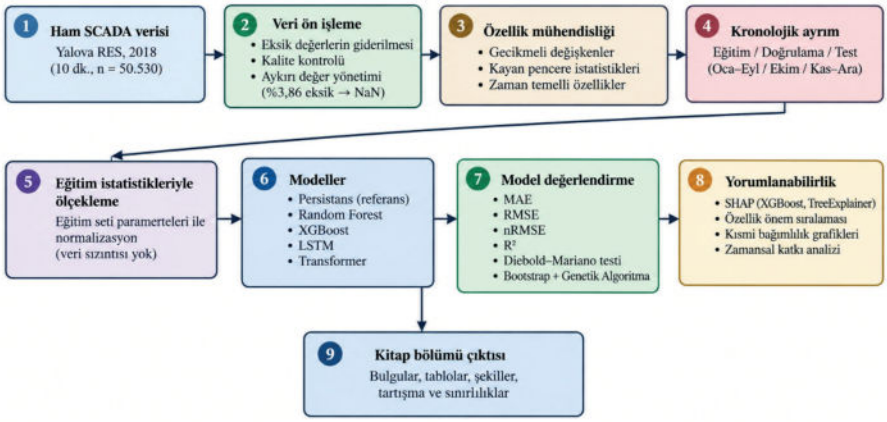
açıklamasını, aerodinamik güç katsayısının (Cp) teorik üst sınırı (Betz limiti, ~0,593) ile kesme ve nominal hız arasındaki kübik güç-hız ilişkisi üzerinden sağlamaktadır.

Literatür ayrıca, model seçiminin tek başına değil, veri ön işleme ve öznitelik mühendisliğinin kalitesiyle birlikte ele alınması gerektiğini vurgulamaktadır. Zaman damgasına dayalı gecikme değişkenleri, hareketli pencere istatistikleri ve periyodik (saatlik, günlük, mevsimsel) kodlamalar, ham SCADA ölçümlerine kıyasla modellerin örüntü yakalama kapasitesini belirgin biçimde artırmaktadır. Bu bulgu, mevcut çalışmanın öznitelik mühendisliği tasarımını (Bölüm 3.3) doğrudan motive etmektedir. Aynı şekilde, kronolojik (zaman sırasına dayalı) eğitim-test ayırımının, rastgele ayırıma kıyasla daha gerçekçi ve daha az iyimser performans tahminleri ürettiği yaygın biçimde kabul görmektedir; zira zaman serisinde rastgele ayırım, gelecekteki bilginin eğitim sürecine dolaylı biçimde sızmasına (veri sızıntısı) yol açabilmektedir (Lones, 2021). Albelali ve Ahmed (2025), LSTM tabanlı zaman serisi tahmininde dizi pencerelerinin veri bölünmesinden önce oluşturulmasının -tam da bu çalışmada özenle kaçınılan hatanın- performans metriklerini sistematik biçimde iyimser yönde çarpıttığını deneysel olarak göstermiştir. Bu bulgu, mevcut çalışmanın önce kronolojik ayırımı yapıp öznitelik ve dizi pencerelerini yalnızca ilgili alt küme içinde oluşturma yaklaşımını (Bölüm 3.2-3.3) doğrudan gerekçelendirmektedir. Belirsizlik tahmini tarafında ise, bağımlı (otokorelasyonlu) zaman serisi gözlemleri için geçerli güven aralıkları oluşturmanın standart bir yolu olan ve Politis ve Romano (1994) tarafından önerilen blok yeniden örnekleme ilkesi, bu çalışmada kullanılan blok bootstrap yaklaşımının (Bölüm 3.5) yöntemsel temelini oluşturmaktadır. Bu çalışmada da kronolojik ayırım ilkesi, hem model eğitiminde hem de ölçekleme ve eşik hesaplamalarında tutarlı biçimde uygulanmıştır.

Sonuç olarak, mevcut literatür bu konuda üç temel eğilim ortaya koymaktadır: (a) modellerin doğruluk açısından sıralaması tahmin dönemine göre değişken olduğundan, "kazanan" model hakkında çıkarım yaparken temkinli olunmalıdır; (b) SHAP gibi yorumlanabilirlik araçları, hem karşılaştırmalı doğruluğu hem de bir modelin davranışının ardındaki nedenleri anlamak için giderek daha önemli hale gelmektedir; ve (c) enerji tahmini literatüründe anlamlılık testi eksikliğine rağmen, bu testler karşılaştırmalı sonuçlar üretmek için kritik önemdedir.

3. Veri ve Yöntem

Şekil 1, araştırmada kullanılan genel veri işleme ve model değerlendirme sürecini göstermektedir. Ham veriden bulgulara uzanan her adım, aşağıdaki alt bölümlerde ele alınmaktadır.



Şekil 1. Veri İşleme ve Model Değerlendirme Akış Şeması

3.1. Veri Seti ve Veri Kalitesi

Araştırma, anket ya da birincil veri kullanmadan kamu istatistiklerine, hakemli yayınlara ve kurumsal ikincil kaynaklara dayanan nicel bir senaryo analizidir. Yöntem dört aşamadan oluşur. İlk aşamada 2024 dizel ağır ticari araç filosu için temel senaryo kurulmuştur. İkinci aşamada 2030 ve 2035 yılları için yüzde 25, yüzde 50, yüzde 75 ve yüzde 100 elektrifikasyonu kapsayan sekiz senaryo tanımlanmıştır. Üçüncü aşamada her senaryo için depodan tekerleğe karbon emisyonu, lastik aşınması ve fren aşınması kaynaklı partikül madde bakımından nokta tahminleri ile olasılıksal belirsizlik aralıkları hesaplanmıştır. Son aşamada karbon tasarrufu ile artan lastik aşınması arasındaki ödünleşim Pareto sınırıyla gösterilmiştir.

Bu araştırmada kullanılan veri seti, Türkiye'nin Marmara Bölgesi'nde, Yalova ilinde (koordinatlar: 40,655°K 29,277°D) bulunan, nominal kapasitesi 3.600 kW olan bir rüzgâr türbininin 2018 yılına ait yıllık SCADA verisinden oluşmaktadır (Kaggle "Wind Turbine Scada Dataset"). Veri seti açık kaynaklıdır ve GitHub üzerinden CSV dosyası biçiminde indirilebilir. Aynı veri seti daha önce Mazen (2025) ile Rajaperumal ve Christopher Columbus (2025) tarafından da kullanılmıştır. Ham veri, 10 dakikalık aralıklarla ölçülen dört değişken içermektedir: tarih/saat, aktif güç (kW), rüzgâr hızı (m/s), üreticiye ait teorik güç değeri (kWh) ve rüzgâr yönü (derece).

Veri kalitesi kontrolleri, ham verinin tam bir yılı kapsayan düzenli bir 10 dakikalık zaman izgarasına (n=52.560) oturtulmasıyla yapılmıştır. Bu eşleştirme, 50.530 gözlem ile beklenen adım sayısı arasında, 32 ayrı ve en uzunu 104,2 saate varan kesinti bloğu içinde dağılmış, eksik zaman damgalarını

temsil eden 2.030 adımlık (%3,86) bir fark ortaya koymuştur. Muhtemelen bakım duruşlarından veya kayıt kesintilerinden kaynaklanan bu veri boşlukları doldurulmamıştır. Bu dönemler gerçek boşluklara karşılık geldiğinden NaN olarak işaretlenmiş ve eksik bir dönem içeren herhangi bir veri penceresi, model girdisi hazırlanırken nihai veri setinden çıkarılmıştır. Enterpolasyon veya başka bir yöntemle doldurma işlemi yapılmamıştır. 57 gözlemde negatif aktif güç değeri bulunmaktadır. Bu, negatif aktif gücün türbinin şebekeden çektiği bekleme veya öz tüketim gücünü ifade ettiği, bilinen bir durumdur.

Ham veri setindeki dört değişkenin ayrıntıları ve birimleri Tablo 1'de gösterilmektedir: Tarih/Saat (10 dakikalık zaman damgası); Aktif Güç (türbinin şebekeye verdiği anlık güç, kW); Rüzgâr Hızı (göbek yüksekliğinde ölçülen anlık rüzgâr hızı, m/s); ve Rüzgâr Yönü (türbinin yöneldiği, göbek yüksekliğindeki rüzgâr yönü, derece). Teorik güç eğrisi, üretici tarafından sağlanan ve yalnızca rüzgâr hızının bir fonksiyonu olarak ideal güç üretimini gösteren referans bir değişkendir. Bu değişken hiçbir modele dahil edilmemiş, yalnızca Şekil 3'te görsel karşılaştırma amacıyla kullanılmıştır.

Tablo 1. Ham SCADA Veri Setindeki Değişkenler

Değişken	Tanım	Birim
Tarih/Saat	10 dakikalık zaman damgası	-
LV ActivePower	Türbinin şebekeye verdiği aktif güç	kW
Wind Speed	Göbek yüksekliğinde ölçülen rüzgâr hızı	m/s
Theoretical_Power_Curve	Üretici güç eğrisine göre beklenen güç (yalnızca görsel karşılaştırma; model girdisi değil)	kWh
Wind Direction	Göbek yüksekliğinde rüzgâr yönü	derece

3.2. Kronolojik Ayırım ve Sızıntı Önleme

Veri kronolojik olarak üç döneme ayrılmıştır: eğitim (1 Ocak - 30 Eylül 2018), doğrulama (1-31 Ekim 2018) ve test (1 Kasım - 31 Aralık 2018). Bu ayırım, gelecekteki bilginin eğitim sürecine sızmasını önlemek amacıyla rastgele değil, kronolojik sıraya dayalı olarak yapılmıştır. Tüm ölçekleme (StandardScaler), rüzgâr rejimi eşikleri (aşağıda açıklanmaktadır) ve hedef normalizasyonu (derin öğrenme modelleri için) yalnızca eğitim dönemi istatistiklerinden hesaplanmıştır. Bu istatistikler daha sonra doğrulama ve test setlerine uygulanmış, ancak ne doğrulama ne de test verisi herhangi bir istatistiğin hesaplanmasına dahil edilmemiştir.

"Düşük rüzgâr" ve "yüksek rüzgâr" rejimleri, literatürdeki sabit fiziksel eşikler yerine, yalnızca eğitim döneminin rüzgâr hızı dağılımının 10. ve

90. yüzdelik dilimlerine dayanan veri temelli bir yaklaşımla tanımlanmıştır (p10=2,29 m/s, p90=13,12 m/s). Bu tercih, çalışmanın veri temelli eşiklerde yalnızca eğitim dönemi yüzdeliklerinin kullanılması ilkesiyle uyumludur ve farklı iklim ve bölge koşullarına genellenebilir bir yöntem sunmaktadır.

3.3. Öznitelik Mühendisliği

Ağaç tabanlı modeller (persistence, Random Forest, XGBoost) için düz bir öznitelik vektörü oluşturulmuştur: t anındaki rüzgâr hızı, rüzgâr yönünün sinüs ve kosinüs dönüşümleri, t anında ve t-10dk/t-30dk/t-1sa gecikmeli güç ve rüzgâr hızı değerleri, 1 saatlik hareketli ortalama ve standart sapma (güç ve rüzgâr hızı için ayrı ayrı hesaplanmıştır) ile günün saatinin ve yılın gününün periyodik (sinüs/kosinüs) kodlamaları. Toplam 18 öznitelik kullanılmıştır. LSTM ve Transformer modelleri için ise aynı bilgiye dayanan ancak dizi biçiminde yapılandırılmış bir girdi oluşturulmuştur. Son 18 adımlık (3 saatlik) rüzgâr hızı, güç ve rüzgâr yönü sinüs/kosinüs değerlerinin dizisi ham biçimde sunulmuş, böylece ağın zamansal örüntüyü kendi iç durumu aracılığıyla öğrenmesi sağlanmıştır.

Üç tahmin ufku tanımlanmıştır: Bunlar; aktif güç değerinden 1 saat (6 adım), 6 saat (36 adım) ve 24 saat (144 adım) sonrası. Her ufuk için hedef değişken, ilgili adım sayısı kadar ileriye kaydırılan güç değeridir. Gerekli gecikme, pencere veya hedef değerlerinden herhangi biri eksik (NaN) olan gözlemler, o ufuk için model girdisinden çıkarılmıştır. Bu nedenle kullanılabilir gözlem sayısı ufuklar arasında hafifçe farklılık göstermektedir (test setinde 1 saat için n=8.203, 6 saat için n=8.138, 24 saat için n=7.922; LSTM ve Transformer için ise dizi kesintisi nedeniyle bu rakamlar marjinal olarak daha da düşüktür).

3.4. Model Mimarileri

Süreklilik (Persistence) taban çizgisi, t anındaki gözlemlenen güç değerinin t+h anı için tahmin olarak kullanılmasına dayanır ve literatürde kısa vadeli rüzgâr tahmini için zorlu bir karşılaştırma ölçütü olarak kabul edilir. Random Forest (150 ağaç, maksimum derinlik 10, minimum yaprak örneklem büyüklüğü 8) ve XGBoost (120 ağaç, maksimum derinlik 4, öğrenme oranı 0,1, alt örnekleme oranı 0,8) modelleri, sırasıyla scikit-learn ve xgboost kütüphaneleri kullanılarak, sabit bir rastgele tohumla (42) eğitilmiştir. LSTM (tek katman, 32 gizli birim, ardından doğrusal bir çıktı katmanı) ve Transformer (2 kodlayıcı katmanı, 32 model boyutu, 4 dikkat başlığı, 64 ileri besleme boyutu, sinüzoidal konumsal kodlama) mimarileri PyTorch ile geliştirilmiştir. Her ikisi de Adam optimizasyon algoritması (öğrenme oranı 1e-3), ortalama kare hata (MSE) kayıp fonksiyonu ve doğrulama kaybı 4 ardışık epoch boyunca

iyileşmediğinde devreye giren erken durdurma ile eğitilmiştir. Yakınsama kararlılığı açısından, derin öğrenme modellerinin hedef değişkeni eğitim seti istatistikleriyle standartlaştırılmış (z-skoru), tahminler ise değerlendirme aşamasında gerçek kW ölçeğine geri dönüştürülmüştür.

3.5. Değerlendirme Çerçevesi

Model, dört farklı ölçütle değerlendirilmiştir. Ortalama mutlak hata (MAE), mutlak tahmin hatalarının ortalama değerini ifade eder ve kilovat (kW) cinsinden yorumlanabilir bir ölçüt sunar. Kök ortalama kare hata (RMSE), büyük hatalara daha fazla ağırlık verir; bu nedenle rampa olayları sırasındaki hızlı değişimlerde ortaya çıkan büyük farklara oldukça duyarlıdır. Normalize edilmiş RMSE (nRMSE), RMSE'nin türbinin kapasitesine (3.600 kW; veri setindeki teorik güç eğrisinin ulaştığı en yüksek değer) oranı olarak hesaplanmıştır. Böylece, farklı kapasitelere sahip türbinlerin veya çalışmaların kolayca karşılaştırılmasını sağlayan, birimsiz bir hata ölçütü elde edilir. Belirleme katsayısı (R^2), test veri setindeki bağımlı değişken varyansının model tarafından açıklanan oranını ifade eder. $R^2=1$ modelin veriyi kusursuz tahmin ettiği, $R^2=0$ tahmin doğruluğunun ortalamaya eşit olduğu, negatif R^2 ise ortalamanın altında bir performans anlamına gelir. İlerleyen bölümde görüleceği üzere, bu oldukça sık karşılaşılan bir durumdur.

Kullanılan dört ölçüt, n test gözlem sayısı, y_t gerçek değer, $\hat{y}_t$ tahmin değeri ve $\bar{y}$ test setinin ortalaması olmak üzere şöyle tanımlanır: $MAE = (1/n)\cdot\Sigma|y_t - \hat{y}_t|$; $RMSE = \sqrt{[(1/n)\cdot\Sigma(y_t - \hat{y}_t)^2]}$; $nRMSE = RMSE / Kapasite$ (burada Kapasite = 3.600 kW); ve $R^2 = 1 - [\Sigma(y_t - \hat{y}_t)^2 / \Sigma(y_t - \bar{y})^2]$. Bu dört ölçüt, modelin performansını farklı açılardan değerlendirir: MAE ve RMSE hatanın mutlak büyüklüğünü kilovat cinsinden ölçerken, nRMSE hatayı türbinin kapasitesiyle ilişkilendirir ve R^2 modelin sağladığı ek açıklama gücünü ölçer.

Tanımlayıcı istatistiklerin ötesinde, XGBoost tahmin modelinin persistence karşılaştırma ölçütüne göre üstünlüğü Diebold-Mariano testi (Diebold ve Mariano, 2002) uygulanarak değerlendirilmiştir. Diebold-Mariano testi, iki rakip tahminin kare hataları arasındaki farkın ($d_t = e_{persistence,t}^2 - e_{XGBoost,t}^2$) ortalamasının sıfırdan anlamlı biçimde farklı olup olmadığını, farkın otokorelasyonunu da (ufuk uzunluğuna eşit bir gecikme derecesiyle, Newey-West tipi düzeltme kullanılarak) hesaba katarak sınar. Pozitif ve istatistiksel olarak anlamlı bir DM istatistiği, XGBoost modelinin persistence karşılaştırma ölçütüne göre daha küçük kare hatalar ürettiği anlamına gelir. XGBoost modelinin MAE değeri için %95 güven aralığı açısından ise, zaman serisi bağımlılığını kısmen hesaba katan blok bootstrap yöntemi (144

adımlık/24 saatlik blok uzunluğu, 2.000 tekrar) kullanılmıştır. Klasik (gözlem bazlı) bootstrap'ın aksine, blok bootstrap örneklemdeki ardışık gözlemler arasındaki bağımlılığı dikkate alarak daha güvenilir güven aralığı tahminleri sunar. Blok uzunluğu (144 adım) sezgisel olarak, ufuk uzunluğuna eşit biçimde seçilmiştir. Sonuçların bu seçime duyarlılığını sınamak amacıyla aynı analiz, 72 (12 saat) ve 288 (48 saat) adımlık blok uzunluklarıyla da tekrarlanmıştır. Sonuçlar (Bölüm 4.2), güven aralığı genişliğinin üç blok uzunluğu arasında beklenen yönde (blok uzunluğu arttıkça hafifçe genişleyerek) değiştiğini, ancak bu değişimin niteliksel sonucu değiştirecek düzeyde olmadığını göstermiştir. Bu nedenle 144 adımlık seçimin, aşırı iyimser (dar) bir güven aralığı üretmediği için savunulabilir olduğu değerlendirilmiştir.

Son olarak, XGBoost modellerinin karar mekanizmasını yorumlamak amacıyla, test setinden sabit bir rastgele tohumla (42) seçilen 1.500 gözlemlik bir alt örneklem üzerinde SHAP TreeExplainer (Lundberg ve Lee, 2017) uygulanmış ve her ufuk için ortalama mutlak SHAP değerine dayalı bir değişken önem sıralaması elde edilmiştir. TreeExplainer, ağaç tabanlı modeller için kesin (yaklaşık değil) Shapley değerlerini hesaplayabilen, hesaplama açısından verimli bir algoritmadır; bunun tüm test seti yerine temsili bir alt örneklem üzerinde çalıştırılması, yalnızca hesaplama süresini kısaltmaya yönelik pratik bir tercih olup TreeExplainer çıktısının kesinliğini etkilememekte, örneklem seçiminden bağımsız kalmaktadır.

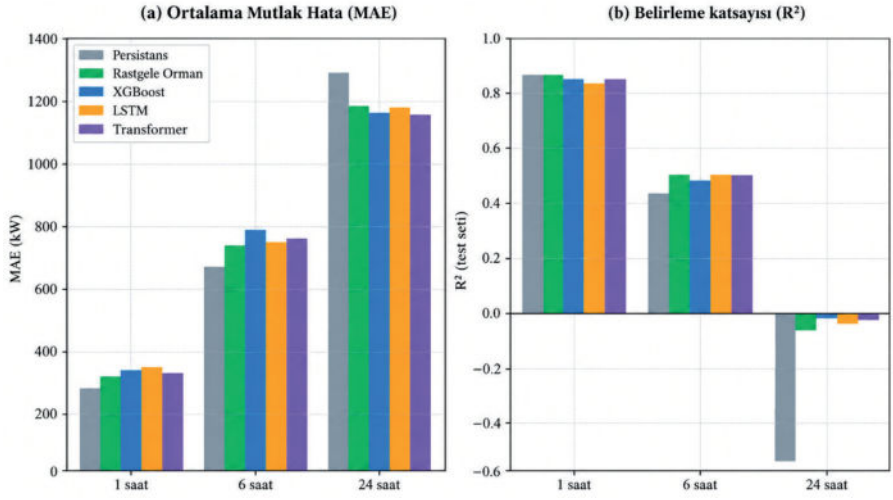
4. Bulgular

4.1. Genel Model Performansı

Şekil 2 ve Tablo 2, beş yöntemin üç tahmin ufkundaki test seti sonuçlarını göstermektedir. 1 saatlik ufukta, taban çizgisi modeli en düşük ortalama mutlak hatayla (MAE=283,7 kW; n=8.203) öne çıkmıştır. Random Forest tekniği en yüksek belirleme katsayısını (R^2=0,875) elde etmiş, ancak MAE açısından (321,4 kW) taban çizgisini geçememiştir. Transformer (MAE=336,5 kW, RMSE=507,3 kW, nRMSE=0,141, R^2=0,864) ve LSTM (MAE=354,3 kW, RMSE=518,3 kW, nRMSE=0,144, R^2=0,858), persistence'a kıyasla benzer ama hafifçe daha zayıf sonuçlar üretmiştir. XGBoost en yüksek MAE'ye (344,4 kW) sahip olmuş, ancak RMSE'si persistence ile karşılaştırılabilir düzeyde (496,4 kW) kalmış ve R^2'si yalnızca hafifçe daha düşük (0,869) olmuştur. Bu sonuçlar, rüzgâr türbinlerinin ürettiği elektriğin kısa ufuklarda yüksek düzeyde otokorelasyonlu olduğu ve bu nedenle, daha gelişmiş modellemeye rağmen Naive Bayesbir "değişim yok" karşılaştırma ölçütünü geçmenin zor olduğu yönündeki yaygın bilinen olguyu doğrulamaktadır.

Tablo 2. Beş Yöntemin Üç Tahmin Ufkundaki Test Seti Performansı

Ufuk	Model	n (test)	MAE (kW)	RMSE (kW)	nRMSE	R^2
1 saat	Persistence	8203	283,7	492,4	0,137	0,871
1 saat	RandomForest	8203	321,4	486,2	0,135	0,875
1 saat	XGBoost	8203	344,4	496,4	0,138	0,869
1 saat	LSTM	8167	354,3	518,3	0,144	0,858
1 saat	Transformer	8167	336,5	507,3	0,141	0,864
6 saat	Persistence	8138	671,3	1034,0	0,287	0,434
6 saat	RandomForest	8138	739,5	976,7	0,271	0,495
6 saat	XGBoost	8138	787,8	995,9	0,277	0,475
6 saat	LSTM	8102	741,2	968,5	0,269	0,503
6 saat	Transformer	8102	762,1	968,7	0,269	0,503
24 saat	Persistence	7922	1278,4	1723,7	0,479	-0,562
24 saat	RandomForest	7922	1187,5	1429,0	0,397	-0,074
24 saat	XGBoost	7922	1166,8	1378,3	0,383	0,001
24 saat	LSTM	7886	1184,6	1409,2	0,391	-0,044
24 saat	Transformer	7886	1158,5	1406,8	0,391	-0,040



Şekil 2. Tahmin Ufkuna Göre Model Performansı (test: Kasım-Aralık 2018).

6 saatlik ufukta sıralama tersine dönmeye başlamaktadır: LSTM (MAE=741,2 kW, RMSE=968,5 kW, nRMSE=0,269, R^2=0,5035) ve Transformer (MAE=762,1 kW, RMSE=968,7 kW, nRMSE=0,269, R^2=0,5033) en yüksek açıklanan varyansı elde ederken, Random Forest

en düşük MAE'yi (739,5 kW, R^2=0,495) vermiştir. Persistence yönteminin performansı bu ufukta belirgin biçimde kötüleşmiş (R^2 değeri 0,434'e düşmüş, MAE ise 671,3 kW olarak kalmıştır). Bu ufukta XGBoost, persistence hariç beş yöntem arasında en yüksek MAE'yi (787,8 kW) ve en düşük R^2'yi (0,475) göstermiştir. Görüldüğü üzere, tüm ufuklarda daha iyi performans gösteren baskın bir yöntem yoktur; model sıralaması tahmin ufkunun bir fonksiyonu olarak değişmektedir.

24 saatlik ufukta ise tüm modellerin performansı önemli ölçüde bozulmuştur: persistence karşılaştırma ölçütü negatif bir R^2 değeri (-0,562) göstermiş, bu da performansının ortalama tahminden bile düşük olduğu anlamına gelmektedir. Random Forest (MAE=1187,5 kW, R^2=-0,074) ile LSTM (MAE=1184,6 kW, R^2=-0,044) ve Transformer (MAE=1158,5 kW, R^2=-0,040) gibi derin öğrenme algoritmaları da negatif değerler göstermiştir. Yalnızca XGBoost modelinin pozitif bir R^2 değeri (0,0013) gösterdiği belirtilmelidir; ancak bu değer sıfıra oldukça yakındır. XGBoost ayrıca en düşük MAE değerini de (1166,8 kW, RMSE=1378,3 kW, nRMSE=0,383) göstermiştir. Transformer'ın MAE'sinin (1158,5 kW) XGBoost'unkinden hafifçe düşük olmasına rağmen R^2'sinin negatif kalması da belirtilmeye değerdir.

Sonuç olarak, hata büyüklüğünün (MAE ve RMSE) sistematik biçimde arttığını, R^2'nin ise sıfıra düştüğünü ya da negatife döndüğünü görmekteyiz. Bu, tahmin ufku uzadıkça mutlak hatanın yükseldiği, aynı zamanda modelin -basit ortalamaya kıyasla ek açıklama gücü olarak tanımlanan- bilgi değerinin düştüğü anlamına gelmektedir. Bu etkinin nedenleri, SHAP analizinden elde edilen sonuçlarla birlikte Bölüm 5'te açıklanmaktadır.

4.2. İstatistiksel Anlamlılık ve Belirsizlik

Tablo 3, Diebold-Mariano testine göre XGBoost ile persistence karşılaştırma ölçütü arasındaki farkın istatistiksel anlamlılığını, XGBoost'un ortalama mutlak hatası (MAE) için blok bootstrap %95 güven aralıklarıyla birlikte sunmaktadır. 1 saatlik tahmin ufkunda (DM=-0,78, p=0,44) ve 6 saatlik tahmin ufkunda (DM=1,07, p=0,28), iki yaklaşım arasındaki fark istatistiksel olarak anlamlı değildir. Bu, Bölüm 4.1'de bulunan küçük MAE farkının rastlantısallıktan kaynaklanabileceği anlamına gelmektedir. Dolayısıyla bu iki ufukta, bir yöntemi "daha iyi" diğerini "daha kötü" saymak için istatistiksel bir gerekçe yoktur; iki yaklaşım pratikte eşit kabul edilebilir. Buna karşılık, 24 saatlik tahmin ufkunda fark yüksek derecede istatistiksel olarak anlamlıdır (DM=3,46, p=0,0005). Örneklem büyüklüğü (n=7.922) göz önüne alındığında, XGBoost'un persistence'a göre kare hatada sağladığı iyileşme, büyüklüğü pratikte önemsiz olsa da ($R^2 \approx 0$), rastlantısallıkla açıklanamayacak kadar güvenilirdir. İstatistiksel

anlamlılık ile pratik önemsizliğin bu şekilde bir arada bulunması, nicel analizde sık karşılaşılan ama enerji tahmini problemlerinde nadiren vurgulanan bir durumdur.

Tablo 3. Diebold-Mariano Testi (persistence ile XGBoost karşılaştırması) ve XGBoost MAE İçin %95 Blok Bootstrap Güven Aralıkları.

Ufuk	XGBoost MAE (kW)	%95 Bootstrap GA (kW)	DM istatistiği	p-değeri	n (test)
1 saat	344,4	[305,2 – 380,0]	-0,780	0,4356	8203
6 saat	787,8	[715,9 – 866,3]	1,071	0,2840	8138
24 saat	1166,8	[1042,7 – 1297,8]	3,457	0,0005	7922

XGBoost modelinin bootstrap yöntemiyle hesaplanan MAE değeri için %95 güven aralıkları şöyledir: 1 saatlik tahmin ufkunda [305,2-380,0] kW (aralık genişliği ≈75 kW); 6 saatlik tahmin ufkunda [715,9-866,3] kW (aralık genişliği ≈150 kW); ve 24 saatlik tahmin ufkunda [1042,7-1297,8] kW (aralık genişliği ≈255 kW). Ufuk uzadıkça yalnızca ortalama hata değil, bu hatayı çevreleyen belirsizlik (güven aralığının genişliği) de sistematik biçimde artmıştır. Bu güven aralıklarının blok uzunluğu seçimine (144 adım/24 saat) duyarlılığı da ayrıca incelenmiştir: aynı analiz 72 (12 saat) ve 288 (48 saat) adımlık blok uzunluklarıyla tekrarlandığında, 24 saatlik ufuktaki aralık genişliğinin sırasıyla yaklaşık 217 kW, 257 kW ve 279 kW olduğu bulunmuştur. Bir başka deyişle, blok uzunluğu arttıkça aralık beklenen yönde hafifçe genişlemiş, ancak hiçbir blok uzunluğunda sıfırı dışarıda bırakacak kadar daralmamıştır. Bu sonuç, 144 adımlık seçimin aşırı iyimser (yapay biçimde dar) bir güven aralığı üretmediğini ve dolayısıyla Bölüm 4.2'deki anlamlılık yorumlarının blok uzunluğu seçimine karşı makul ölçüde dayanıklı olduğunu doğrulamaktadır. Genel bir örüntü olarak, uzun ufuklu tahminlerin yalnızca ortalama olarak daha kötü değil, aynı zamanda dönemden döneme daha değişken (daha az güvenilir) olduğu görülmektedir. Bu da şebeke işletmecilerinin 24 saatlik tahminleri tek bir kesin rakam yerine olası bir aralık olarak ele almasının önemini pekiştirmektedir.

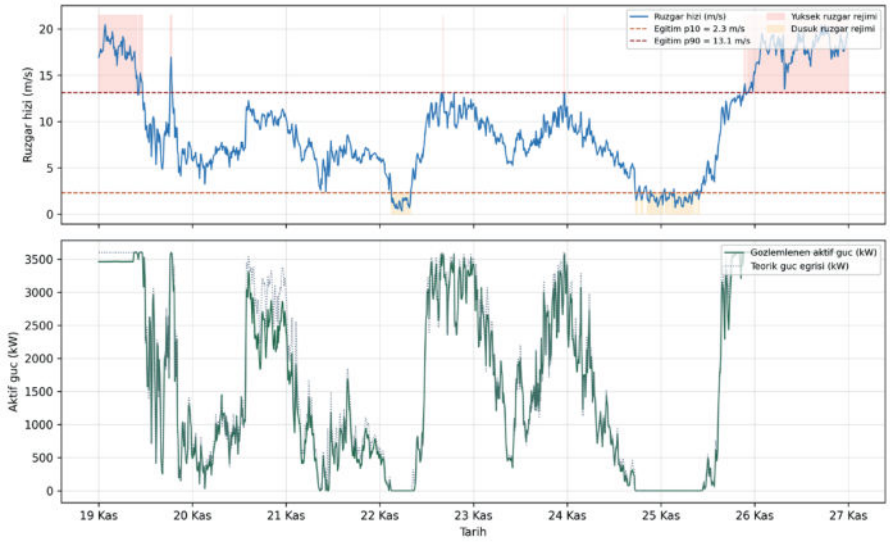
4.3. Rüzgâr Rejimine Göre Hata Dağılımı

Şekil 3'te görsel olarak gösterilen düşük, normal ve yüksek rüzgâr rejimlerine göre XGBoost hatası incelendiğinde (Tablo 4), her üç ufukta da en düşük MAE değerinin düşük rüzgâr rejiminde (rüzgâr hızı ≤2,29 m/s) ortaya çıktığı görülmektedir (1 saatte 136,6 kW, 6 saatte 573,1 kW ve 24 saatte 1298,7 kW; bu sonuncusu ilginç biçimde normal rejimden daha yüksektir). Bu durum; düşük rüzgâr rejiminde güç üretimi kesme hızının

altında kalır ve bu nedenle çoğunlukla sıfıra yakın, düşük varyanslıdır; bu da mutlak hatayı küçük tutar. Buna karşılık, gücün rüzgâr hızıyla birlikte en dik (ve en değişken) biçimde arttığı normal rejimde (1 saatte 374,7 kW, 6 saatte 819,6 kW), hata belirgin biçimde artmaktadır. Bu durum, tahmin probleminin rüzgâr türbini güç eğrisinin S-şekilli orta bölgesinde en yüksek olduğu yönündeki mevcut mühendislik görüşüyle tutarlıdır. Bu S-şekilli davranış, rüzgâr türbinlerinin temel aerodinamiğinden kaynaklanmaktadır: kesme hızı ile nominal hız arasındaki aralıkta türbinin ürettiği güç, rüzgâr hızının küpü olarak değişmekte ve Cp katsayısı teorik olarak her zaman Betz limitinin (0,593) altında kalmaktadır (Manwell vd., 2010). Bu aralıkta, rüzgâr hızı ölçümündeki veya tahminindeki küçük sapmalar, tahmin edilen güçte büyük sapmalara yol açmaktadır. Yüksek rüzgâr hızı rejiminde (≥13,12 m/s) ise tahmin hataları, normal rüzgâr hızı rejimine kıyasla daha küçüktür (1 saatte 286,2 kW, 6 saatte 725,0 kW, 24 saatte 1248,8 kW). Bunun nedeni, yüksek rüzgâr hızı rejiminde üretilen gücün pratikte nominal kapasiteye (3600 kW) yakın bir düzeyde sabitlenmesidir.

Tablo 4. Rüzgâr Rejimine Göre XGBoost Ortalama Mutlak Hatası (Test Seti).

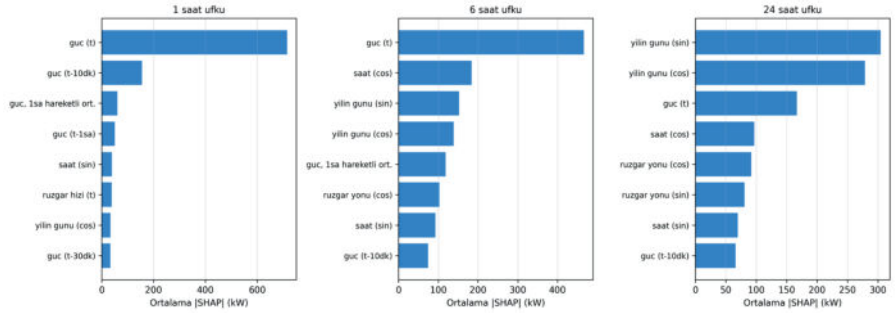
Ufuk	**Rüzgâr rejimi**	**n (test)**	**XGBoost MAE (kW)**
1 saat	Düşük rüzgâr (≤p10)	589	136.6
1 saat	Normal	6396	374.7
1 saat	Yüksek rüzgâr (≥p90)	1218	286.2
6 saat	Düşük rüzgâr (≤p10)	582	573.1
6 saat	Normal	6338	819.6
6 saat	Yüksek rüzgâr (≥p90)	1218	725.0
24 saat	Düşük rüzgâr (≤p10)	546	1298.7
24 saat	Normal	6158	1138.9
24 saat	Yüksek rüzgâr (≥p90)	1218	1248.8



Şekil 3. Örnek Bir Haftada (19-26 Kasım 2018) Rüzgâr Hızı ve Aktif Güç Üretimi (Eğitim Setine Dayalı Düşük ve Yüksek Rüzgâr Eşikleri İşaretlenmiştir).

4.4. SHAP ile Yorumlanabilirlik

Şekil 4, üç ufkun her birinde XGBoost modellerinin ortalama mutlak SHAP değerine göre sıralanmış en önemli sekiz değişkenini sunmaktadır. 1 saatlik ufukta, modelin kararı büyük ölçüde t anında gözlemlenen güç değerine dayanmaktadır (ortalama |SHAP|=714,9 kW); bunu 10 dakika gecikmeli güç (155,1 kW) ve 1 saatlik hareketli ortalama (60,3 kW) izlemektedir. Başka bir deyişle, kısa ufuklarda model esasen "yakın geçmişin ağırlıklı bir uzantısı" olarak işlev görmektedir. 6 saatlik ufukta, t anındaki güç hâlâ en önemli değişken olmakla birlikte (466,2 kW), önemi belirgin biçimde azalmış; günün saati (183,8 kW) ve yılın günü (152,3 ve 138,6 kW) gibi periyodik ve mevsimsel değişkenler öne çıkmaya başlamıştır. 24 saatlik ufukta ise tablo tamamen değişmekte, en önemli iki değişken yılın gününün sinüs ve kosinüs kodlamaları (304,5 ve 279,0 kW) olurken, t anındaki güç üçüncü sıraya (166,7 kW) gerilemektedir.



Şekil 4. XGBoost Modelleri İçin SHAP Değişken Önem Sıralaması (Test Alt Örneklemi, n=1500).

Bu bulgu, Bölüm 4.1'de bildirilen performans bozulmasıyla doğrudan tutarlıdır ve onu açıklayıcı niteliktedir. 24 saatlik ufukta model artık gerçek bir tahmin yapmamakta, ölçülebilir ölçüde mevsimsel bir iklimsel temel çizgiye (yılın o dönemi için tipik rüzgâr ve güç düzeyi) geri dönmektedir. Bu geçiş, dışsal sayısal hava tahmini girdisi bulunmayan bir kurulumda beklenen ve fiziksel olarak makul bir davranıştır.

Bu üç ufuk arasındaki fark, değişken önem sıralamasının ardındaki baskın bilgi kaynağının niteliğindeki değişim olarak görülebilir. Somut olarak, bilgi kaynağı 1 saatte anlık ve yakın geçmiş güç ölçümleri; 6 saatte yakın geçmiş ölçümler ile günün saati (kısmi periyodiklik); son olarak 24 saatte ise yılın günü (mevsimsel iklimsel temel çizgi) biçimindedir. Rüzgâr yönünün kosinüs ve sinüs bileşenleri de, özellikle 6 ve 24 saatlik ufuklarda, değişken önem sıralamasını etkilemektedir; bu da rüzgâr yönünde mevsimsel değişimler olduğuna (örneğin kış aylarında belirli bir rüzgâr yönünün daha baskın olması) ve modelin rüzgâr yönünü mevsimler için bir vekil gösterge olarak kullandığına işaret etmektedir. Buna karşılık, dönüştürülmemiş rüzgâr hızı değişkeni (wind_speed) SHAP önem sıralamasında yalnızca 1 saatlik ufukta anlamlı düzeydedir (38,97 kW) ve 6 ile 24 saatlik ufuklarda ilk sekize girmemektedir. Bunun nedeni, rüzgâr hızı değişkeninin yüksek kısa vadeli otokorelasyonu nedeniyle, mevcut ölçümün gelecekteki değeri tahmin etmek için önemini yitirmesidir.

4.5. Hesaplama Süresi ve Yorumlanabilirlik Karşılaştırması

Tahmin performansına ek olarak, hesaplama verimliliği ve model yorumlanabilirliği de pratik uygulama açısından tartışılmaya değerdir. Bu çalışmada, GPU optimizasyonu olmaksızın yalnızca CPU tabanlı bir ortamda eğitim süreleri tahmin edilmiş ve önemli yöntemsel farklılıklar gözlemlenmiştir. Örneğin, yaklaşık 37.000 ila 39.000 veri örneği üzerinde eğitilen Random

Forest ve XGBoost algoritmaları, ufuk başına birkaç saniyeden birkaç on saniyeye kadar süre gerektirmiştir. LSTM modeli, erken durdurma sayesinde ufuk başına yaklaşık 9-18 saniyede yakınsamıştır. Dikkat mekanizmasının getirdiği ek hesaplama yükü nedeniyle Transformer mimarisi, ufuk başına yaklaşık 27-58 saniye gerektirmiştir. Bir başka deyişle, en yavaş yöntem en hızlı yöntemden kabaca 5-10 kat daha fazla hesaplama süresi gerektirmiştir. Bu fark, gerçek zamanlı ya da sık güncellenen (örneğin her 10 dakikada bir yeniden eğitilen) operasyonel sistemlerde, Transformer'ın sunduğu doğruluk kazancının -özellikle 24 saatlik ufuktaki marjinal iyileşmenin- hesaplama maliyetiyle birlikte değerlendirilmesi gerektiğini göstermektedir.

Yorumlanabilirlik açısından, ağaç tabanlı yöntemler (Random Forest, XGBoost) hem yerleşik değişken önem ölçütleri hem de SHAP gibi son teknoloji açıklanabilirlik araçlarıyla doğrudan ve hesaplama açısından verimli biçimde analiz edilebilmektedir (Bölüm 4.4). Buna karşılık, LSTM ve Transformer gibi derin öğrenme mimarileri için SHAP benzeri açıklamalar teorik olarak mümkün olsa da (örneğin DeepExplainer veya GradientExplainer aracılığıyla), bunların hesaplama maliyeti TreeExplainer'a kıyasla çok daha yüksektir ve ortaya çıkan açıklamaların kararlılığı örnekleme stratejisine daha duyarlıdır. Bu nedenle, bu çalışmadaki yorumlanabilirlik analizi yalnızca XGBoost modelleri üzerinde yürütülmüştür. Burada izlenen yöntem, karar ağaçlarının yalnızca rekabetçi doğruluk düzeyleri sunmakla kalmayıp aynı zamanda hesaplama açısından verimli olduğu ve sonuçların yorumlanabilirliğini sağladığı yönündeki genel bulguyla tutarlıdır. Bu özellik, şebeke işletmecilerinin model sonuçlarını açıklaması gereken düzenleyici bağlamlarda özellikle önem taşımaktadır.

5. Tartışma

Bu çalışmanın temel bulgusu, model karmaşıklığının tahmin performansı üzerindeki etkisinin, tahmin ufkuyla nasıl etkileşime girdiğine bağlı olarak farklılaştığıdır. Daha gelişmiş modellerin (XGBoost, LSTM, Transformer) daha kısa ufukta (1 saat) MAE açısından persistence modelini geçememesi, ilk bakışta "makine öğrenmesinin işe yaramadığı" biçiminde yorumlanabilir. Ancak bu bulgu, rüzgâr tahmini literatüründe yerleşik olan, çok kısa ufuklarda otokorelasyonun görece yüksek olduğu ve bu nedenle persistence'ın bu tür durumlarda neredeyse kusursuz bir tahminci olduğu görüşüne açık bir destek sunmaktadır. Bu ufukta Diebold-Mariano testinin anlamlı çıkmaması ($p=0{,}44$), küçük MAE farklarının örneklem değişkenliği sınırları içinde kaldığını doğrulamaktadır.

24 saatlik ufuktan elde edilen sonuçlar ikili bir yorumu işaret etmektedir. Birincisi, XGBoost modelinin persistence modeline kıyasla anlamlı üstünlüğü

($p<0,001$), makine öğrenmesi yönteminin kare hata ölçütü açısından Naive Bayes modele göre sistematik bir avantaj sağladığı anlamına gelmektedir. İkincisi, iki model arasındaki farkın gerçek büyüklüğü küçüktür (sıfıra yakın R^2 ile gösterildiği üzere). SHAP analizinden elde edilen bulgularla birlikte değerlendirildiğinde, bu durum, farkın büyük ölçüde gerçek bir tahmin becerisinden değil, mevsimsel ortalamayı daha iyi hesaplayabilme yeteneğinden kaynaklandığı anlamına gelmektedir. Başka bir deyişle, bu karşılaştırma bir tahmin gücü ölçütü değil, çok kötü bir taban çizgisine (persistence modelinin 24 saatlik ufukta tamamen çöktüğü) göre üstünlüğün bir değerlendirmesidir.

Derin öğrenme mimarilerinin (LSTM, Transformer) hiçbir ufukta ağaç tabanlı yöntemleri açık ve tutarlı biçimde geçememesi, literatürdeki bir başka tekrarlayan gözlemle uyumludur: orta büyüklükteki (bu çalışmada yaklaşık 37.000-39.000 eğitim örneği), tablo biçimli veya kısa dizi girdili problemlerde, derin öğrenme mimarilerinin teorik esnekliği her zaman pratik bir avantaja dönüşmemektedir. Bununla birlikte, LSTM ve Transformer'ın 6 saatlik ufukta en yüksek R^2 değerlerine ulaşması, bu mimarilerin hem yakın geçmişin hem de periyodik örüntülerin birlikte önemli olduğu orta vadeli ufuklarda göreli bir avantaj sunabileceğini düşündürmektedir; ancak bu gözlem, tek bir türbin ve tek bir yılla sınırlı olduğundan temkinle yorumlanmalıdır. Bu çalışmada Diebold-Mariano testinin yalnızca XGBoost-persistence çifti için uygulandığı vurgulanmalıdır (Bölüm 3.5, 4.2). LSTM, Transformer ve Random Forest arasındaki (veya bu üçü ile XGBoost arasındaki) farkların istatistiksel anlamlılığı ayrıca test edilmemiştir. Bu nedenle, "LSTM/Transformer 6 saatte daha iyi performans gösteriyor" veya "XGBoost 24 saatte en düşük MAE'ye sahip" gibi ifadeler, yalnızca gerçekleştirilen nokta tahminlerine dayalı betimleyici karşılaştırmalardır. Bu farkların örneklem gürültüsünden ayırt edilebilir olup olmadığının belirlenmesi, ilgili her model çifti için ayrı Diebold-Mariano testleri gerektirir; bu ise bu çalışmanın kapsamı dışındadır (Bölüm 7'ye bakınız).

Rüzgâr rejimine dayalı hata analizi (Bölüm 4.3), güç eğrisinin S-şekilli yapısının tahmin görevinin karmaşıklık düzeyini doğrudan etkilediğini; tahminin en zor olduğu bölgenin düşük ya da yüksek rüzgâr değil, gücün rüzgâr hızına en bağımlı olduğu orta bölge olduğunu ortaya koymaktadır. Bu bulgu, rüzgârın uç değerlerinin mutlaka en yüksek tahmin hatasına yol açmadığını ve tahmin modellerinin güvenilirliğine ilişkin her analizin yalnızca uç değerleri değil, güç eğrisinin dinamik bölgelerini de göz önünde bulundurması gerektiğini göstermektedir. Ayrıca bu bulgu, şebeke işletmecileri açısından rezerv kapasite tahmini bakımından da pratik bir değer taşımaktadır; zira yalnızca "fırtına" ya da "sakin hava" durumları değil, bu durumda tahmin hatasındaki sistematik artış nedeniyle orta düzeydeki rüzgâr değişimleri de dikkate alınmalıdır.

Bu çalışmanın bulguları, aynı ya da benzer SCADA veri setini kullanan diğer çalışmaların bulgularıyla bir ölçüde örtüşmektedir. Özellikle Rajaperumal ve Christopher Columbus (2025), bu çalışmada da kullanılan, Kaggle'dan elde edilmiş Türkiye rüzgâr türbini veri setini üç veri durumundan biri olarak kullanmış ve topluluk entegrasyonu yaklaşımlarının tekil modellere kıyasla daha iyi performans gösterdiğini bulmuştur. Bu çalışmada topluluk entegrasyonu (modellerin çıktı sonuçlarının birleştirilmesi) uygulanmamış olsa da, farklı modellerin farklı tahmin ufuklarında en iyi performansı göstermesi bulgusu (Bölüm 4.1), gelecekteki araştırmalarda böyle bir yaklaşımın uygulanmasının -özellikle 6 saatlik ufukta LSTM/Transformer avantajının ve 24 saatlik ufukta XGBoost avantajının bir araya getirilmesinin- sağlayabileceği potansiyel faydalara işaret etmektedir. Mazen (2025), aynı veri setinde aykırı değer temizliğinin tahmin performansını artırdığını bulmuştur. Bu çalışmada aykırı değerlerin (örneğin 57 negatif güç kaydının) olduğu gibi bırakılması, mevcut sonuçları hafifçe kötüleştirmiş olabilir; ancak bu, gerçek dünya işletmesinin "temizlik öncesi" veri durumunu yansıtmak amacıyla bilinçli bir araştırma yöntemi tercihi olarak yapılmıştır. Bu araştırmanın bulguları ayrıca, hiçbir model sınıfının evrensel bir üstünlüğe sahip olmadığını ve uygunluğun bölgesel ve ufka özgü etkenlere bağlı olduğunu savunan Mugware vd. (2024) ile Mollasalehi ve Farhadi (2025) ile de örtüşmektedir; bu görüş, Bölüm 4.1-4.3'te gözlemlenen sıralama tersine dönmeleriyle tamamen doğrulanmaktadır. Aynı şekilde, uzun ufuklarda açıklanan varyansın azalması ve bunun meteorolojiye ilişkin dışsal girdi eksikliğinden kaynaklanması, Sedai vd. (2023)'nin güneş enerjisi tahmini çalışmalarında gözlemlediği eğilimle örtüşmektedir. Bu durum, söz konusu sınırlamanın yalnızca rüzgâr enerjisine özgü olmadığını, yalnızca SCADA verisine dayalı tahmin sistemlerinin karşılaştığı daha genel bir sorun olduğunu göstermektedir.

6. Sonuç ve Öneriler

Bu çalışma, Türkiye'deki bir rüzgâr türbinine ait gerçek SCADA verisi kullanılarak, beş tahmin yönteminin performansının üç farklı tahmin dönemi boyunca istatistiksel olarak anlamlı bir karşılaştırmasını sunmaktadır. Temel bulgular şunlardır: (i) çok kısa dönemde (1 saat) persistence, kullanılan tüm gelişmiş yöntemlerle istatistiksel olarak karşılaştırılabilir, güvenilir bir karşılaştırma ölçütü işlevi görmektedir; (ii) orta dönemde (6 saat) LSTM ve Transformer ağları R^2 açısından hafif bir üstünlük sağlamaktadır; (iii) uzun dönemde (24 saat) yalnızca XGBoost istatistiksel olarak anlamlı ama önemsiz düzeyde bir kazanç üretmekte, SHAP analizi ise bu dönemdeki tahminlerin büyük ölçüde mevsimselliğe bağlı olduğunu göstermektedir; ve (iv) tahmin hatası rüzgâr rejimine önemli ölçüde bağlıdır; en yüksek hatalar

düşük ya da yüksek rüzgârlardan değil, güç eğrisinin dik olduğu orta aralıktan kaynaklanmaktadır.

Pratik öneri, kazanan modele bağlı kalmak yerine ufka duyarlı bir model seçimi yaklaşımı kullanmaktır. Çok kısa vadeli operasyonel karar alma (örneğin dakikalar ile bir saat arasındaki yük dengeleme) için persistence gibi basit ve hesaplama açısından ucuz yöntemler yeterli olabilir. Buna karşılık, orta vadeli planlama (gün içi piyasa katılımı, 6 saatlik ufuk) için LSTM veya Transformer gibi derin öğrenmeye dayalı modeller değerlendirilmelidir. 24 saatlik ufuktaki bulgunun operasyonel karşılığı, XGBoost'un yaklaşık 1167 kW'lık bir MAE üretmesi biçiminde ifade edilebilir; bu da yaklaşık %32'lik bir nRMSE'ye ya da 3.600 kW'lık nominal kapasiteden %32'lik bir sapmaya karşılık gelmektedir (nRMSE≈0,38). Türkiye'de dengesizlik maliyetleri, gün öncesi piyasa rejimi çerçevesinde gerçekleşen ile planlanan üretim arasındaki sapmaya göre hesaplandığından, tek bir türbin için bile %32'lik bir sapma önemli mali riskler taşımaktadır. Santral düzeyinde (onlarca türbin) bu sapma, MWh başına büyük parasal tutarlara karşılık gelebilir; ancak kesin tutar bu araştırmanın kapsamı dışındadır ve güncel dengesizlik oranlarına dayalı ek bir ekonomik analiz gerektirir. Daha da önemlisi, bu çalışmanın gösterdiği gibi, 24 saat ve ötesindeki ufuklarda (gün öncesi piyasa teklifleri, yük tevzi planlaması) doğruluk artışı arayışında yalnızca geçmiş SCADA verisine güvenilemez. Rüzgâr hızı ve yönü tahminleri gibi NWP verilerinin girdi verisine dahil edilmesi, literatürde tespit edilen ve bu çalışmada da doğrulanan "uzun ufuk performans çöküşü" sorununu çözmenin tek yolu gibi görünmektedir.

Türkiye özelinde, EPİAŞ Şeffaflık Platformu'nda ücretsiz olarak sunulan sistem düzeyindeki rüzgâr ve güneş üretim gerçekleşme verilerinin, bu çalışmada benimsenen türbin düzeyindeki yöntemle birleştirilmesi, açık bir gelecekteki çalışma alanına işaret edebilir. Şöyle ki, santral veya sistem düzeyindeki toplulaştırma, türbin düzeyinde gözlemlenen coğrafi değişkenliğin bir kısmını istatistiksel olarak azaltabilmeli (coğrafi çeşitlendirme etkisi) ve özellikle daha uzun ufuklarda, bu çalışmada bulunandan daha yüksek bir açıklanan varyansa (R^2) yol açmalıdır. Bu hipotezi test etmek için ayrı bir ampirik analiz yürütülmelidir. Gelecekteki araştırmalar için şu alanlar potansiyel konular olarak değerlendirilebilir: (a) birden fazla türbin/santralden elde edilen çok yıllı veri analiziyle sonuçların zaman ve coğrafya boyunca kararlılığının test edilmesi; (b) sayısal hava tahminlerinin (ERA5, Open-Meteo, ECMWF kaynaklı rüzgâr hızı/yönü tahminleri) kullanılmasıyla 24 saatlik ufukta ve ötesinde elde edilebilecek kazancın nicel olarak belirlenmesi; (c) şebeke işletmecilerinin yalnızca nokta tahminleri değil olası bir değer aralığını da bilmeyi tercih etmesi nedeniyle, olasılıksal yaklaşımların (kuantil regresyon/olasılıksal derin öğrenme yoluyla tahmin belirsizliği aralıklarının) araştırılması;

ve (d) tüm ilgili model çiftlerinin (LSTM ve Transformer, Random Forest ve XGBoost vb.) karşılaştırılması için Diebold-Mariano testinin uygulanması.

7. Sınırlılıklar

Bu bulgular, birtakım önemli sınırlılıklar ışığında değerlendirilmelidir. Birincisi, analiz tek bir türbine ve tek bir takvim yılına (2018) dayanmaktadır. İkincisi, bulguların farklı türbin tiplerine, coğrafi bölgelere veya yıllar arası iklim değişkenliğine genellenebilirliği sınırlıdır. Bu bağlamda, 24 saatlik ufukta gözlemlenen düşük performansın (Bölüm 4.1, 5) her ikisinin birleşik bir sonucu olabileceği göz önünde bulundurularak, iki ayrı sınırlılık kaynağı birbirinden ayrıştırılmalıdır: (a) veri yetersizliği: tek yıllık veri, mevsimsel (yılın günü) örüntünün yalnızca bir kez gözlemlendiği anlamına gelir; bu nedenle bu kurulum, modelin gerçekten mevsimsel olmayan, yıldan yıla genellenebilir bir örüntü öğrenip öğrenmediğini test edememektedir; ve (b) yöntemsel kısıt: model girdisi, dışsal sayısal hava tahmini (NWP) verisi olmaksızın yalnızca aynı türbinin geçmiş SCADA kayıtlarıyla sınırlı tutulmuştur; bu, girdi kümesinin kendisinden kaynaklanan ve çok yıllı veriyle bile devam edecek bir sınırlılıktır. Bu çalışmanın tasarımı, bu iki kaynağı birbirinden ayıracak bir karşılaştırma (örneğin NWP olmaksızın çok yıllı veri ile NWP ile tek yıllık veri karşılaştırması) içermediğinden, 24 saatlik ufuktaki performans çöküşünün ne ölçüde (a)'dan, ne ölçüde (b)'den kaynaklandığı bu çalışma kapsamında ayrıştırılamamaktadır. Bu, gelecekteki çalışmalar için açık bir araştırma sorusu olarak kalmaktadır. Üçüncüsü, bu çalışmada "aşırı hava" koşulları yalnızca rüzgâr hızının istatistiksel (yüzdelik dilim tabanlı) uç değerleri üzerinden temsil edilmiştir. Fırtına veya ani rüzgâr rampa olayları gibi meteorolojik açıdan daha zengin aşırı olay tanımları, veri setinin kapsamı (yalnızca rüzgâr hızı, yönü ve güç içermesi) nedeniyle ayrı biçimde modellenememiştir. Dördüncüsü, derin öğrenme modelleri (LSTM, Transformer), hesaplama kaynağı kısıtları nedeniyle görece küçük mimarilerle (tek veya iki katman, düşük gizli boyut) ve sınırlı sayıda epoch ile eğitilmiştir; daha büyük mimariler, daha kapsamlı bir hiperparametre araması veya daha uzun eğitim süreleri farklı (muhtemelen daha iyi) sonuçlar verebilir. Beşincisi, Diebold-Mariano testi yalnızca XGBoost-persistence karşılaştırmasına uygulanmış ve tüm model çiftleri için sistematik biçimde tekrarlanmamıştır (Bölüm 5'e bakınız); bu nedenle bu çalışma, örneğin LSTM ile Transformer arasındaki farkların anlamlılığına ilişkin istatistiksel bir sonuç sunmamaktadır. Altıncısı, bu çalışmada kullanılan hiperparametreler (örneğin Random Forest için ağaç sayısı ve derinlik, XGBoost için öğrenme oranı, LSTM ve Transformer için gizli boyut ve katman sayısı), sistematik bir arama (grid ya da random search) yerine, hesaplama süresi kısıtları gözetilerek makul bir yargıya dayalı olarak seçilmiştir. Dolayısıyla, bildirilen performans

farkları kısmen model ailesinin doğal kapasitesini değil, hiperparametre seçimini yansıtıyor olabilir. Yedincisi, yalnızca tek bir türbine ait veri kullanıldığından, sonuçların santral ya da sistem düzeyinde toplulaştırılmış (birden fazla türbin üzerinden ortalanmış) üretim tahminine genellenebilirliği doğrudan test edilmemiştir. Toplulaştırma genellikle tekil türbin düzeyinde gözlemlenen yerel değişkenliği yumuşatarak tahmin doğruluğunu artırma eğiliminde olduğundan, bu çalışmanın bulgularının santral düzeyi bir uygulamada muhtemelen daha olumlu sonuçlara dönüşmesi beklenebilir. Son olarak, negatif güç kayıtları ve olası sensör kaynaklı aykırı değerler (Mazen, 2025'te olduğu gibi) bu çalışmada da temizlenmeden bırakılmıştır. Bu tercih gerçekçi bir işletme senaryosunu yansıtsa da, aykırı değerlerin model eğitimi üzerindeki marjinal etkisi bu çalışma kapsamında ayrıca izole edilerek ölçülmemiştir.

Tüm hesaplamalar Python 3.12 ortamında; veri işleme için pandas ve pyarrow, klasik makine öğrenmesi modelleri için scikit-learn ve xgboost, derin öğrenme modelleri için PyTorch, yorumlanabilirlik analizi için shap, istatistiksel testler için scipy ve görselleştirme için matplotlib kütüphaneleri kullanılarak gerçekleştirilmiştir. Tüm betikler, ham veriden nihai tablo ve şekillere kadar uçtan uca çalıştırılabilir, numaralandırılmış (01'den 13'e) ayrı Python dosyaları olarak düzenlenmiş; her betik, bir önceki adımın çıktısını (ara veri dosyaları, eğitilmiş model nesneleri) kendi girdisi olarak kullanmıştır. Bu modüler yapı hem tekrarlanabilirliği kolaylaştırmakta hem de tek bir adımın (örneğin yalnızca SHAP analizinin) tüm ardışık düzenin baştan çalıştırılmasına gerek kalmadan yeniden yürütülmesine imkân tanımaktadır.

Veri ve Kod Erişimi Beyanı

Bu deneyde kullanılan orijinal SCADA veri seti, Kaggle web sitesinde "Wind Turbine Scada Dataset" başlığı altında kamuya açık olarak bulunmaktadır (https://www.kaggle.com/datasets/berkerisen/wind-turbine-scada-dataset) ve GitHub üzerinde de yer almaktadır. Veri ön işleme, öznitelik mühendisliği, model eğitimi, istatistiksel testler, SHAP analizi ve sonuçların görselleştirilmesi dahil olmak üzere deneyleri gerçekleştirmek için kullanılan Python kodu, işlenmiş ara veri tabloları ve yüksek çözünürlüklü grafiklerle birlikte, bu bölümün teslim dosyalarına dahil edilmiştir. Rastgelelik içeren tüm kod boyunca sabit bir rastgele tohum değeri (42) kullanılmış olup, bu sayede burada elde edilen sonuçların, ortak ham veri üzerinde kodun sıfırdan yeniden çalıştırılmasıyla tekrarlanabilir olması sağlanmaktadır.

Kaynakça

Albelali, S., & Ahmed, M. (2025). *Hidden leaks in time series forecasting: How data leakage affects LSTM evaluation across configurations and validation strategies*. arXiv preprint arXiv:2512.06932. https://doi.org/10.48550/arXiv.2512.06932

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5-32. https://doi.org/10.1023/A:1010933404324

Cakiroglu, C., Demir, S., Ozdemir, M. H., Aylak, B. L., Sariisik, G., & Abualigah, L. (2024). Data-driven interpretable ensemble learning methods for the prediction of wind turbine power incorporating SHAP analysis. *Expert Systems with Applications, 237*(A), Article 121464. https://doi.org/10.1016/j.eswa.2023.121464

Chen, T., & Guestrin, C. (2016). *XGBoost: A scalable tree boosting system*. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining içinde (ss. 785-794). Association for Computing Machinery.

Diebold, F. X., & Mariano, R. S. (2002). Comparing predictive accuracy. *Journal of Business & Economic Statistics, 20*(1), 134-144. https://doi.org/10.1080/07350015.1995.10524599

Han, L., Qiao, Y., Li, M., & Shi, L. (2020). Wind power ramp event forecasting based on feature extraction and deep learning. *Energies, 13*(23), Article 6449. https://doi.org/10.3390/en13236449

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation, 9*(8), 1735-1780. https://doi.org/10.1162/neco.1997.9.8.1735

Hussain, I., Ching, K. B., Uttraphan, C., Tay, K. G., Memon, I., & Memon, S. A. (2025). Predicting monthly wind speeds using XGBoost: A case study for renewable energy optimization. *Processes, 13*(6), Article 1763. https://doi.org/10.3390/pr13061763

Lones, M. A. (2021). H*ow to avoid machine learning pitfalls: A guide for academic researchers*. arXiv preprint arXiv:2108.02497. https://doi.org/10.48550/arXiv.2108.02497

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems, 30*, 4765-4774.

Manwell, J. F., McGowan, J. G., & Rogers, A. L. (2010). *Wind energy explained: Theory, design and application* (2nd ed.). John Wiley & Sons.

Mazen, R. (2025). A two-stage model for outlier detection and power prediction of wind turbine using SCADA dataset. *International Journal of Energy Research*, *2025*(1), Article 2527561. https://doi.org/10.1155/er/2527561

Mollasalehi, A., & Farhadi, A. (2025). *Solar and wind power forecasting: A comparative review of LSTM, Random Forest, and XGBoost models*. arXiv preprint arXiv:2509.24059. https://doi.org/10.48550/arXiv.2509.24059

Mugware, F. W., Sigauke, C., & Ravele, T. (2024). Evaluating wind speed forecasting models: A comparative study of CNN, DAN2, Random Forest and XGBoost in diverse South African weather conditions. *Forecasting, 6*(3), 672-699. https://doi.org/10.3390/forecast6030035

Politis, D. N., & Romano, J. P. (1994). The stationary bootstrap. *Journal of the American Statistical Association, 89*(428), 1303-1313. https://doi.org/10.1080/01621459.1994.10476870

Rajaperumal, T. A., & Christopher Columbus, C. (2025). Enhanced wind power forecasting using machine learning, deep learning models and ensemble integration. *Scientific Reports, 15*(1), Article 20572. https://doi.org/10.1038/s41598-025-05250-3

Sedai, A., Dhakal, R., Gautam, S., Dhamala, A., Bilbao, A., Wang, Q., Wigington, A., & Pol, S. (2023). Performance analysis of statistical, machine learning and deep learning models in long-term forecasting of solar power production. *Forecasting, 5*(1), 256-284. https://doi.org/10.3390/forecast5010014

Sogukpinar, H., Bozkurt, I., Oyucu, S., & Aksoz, A. (2025). Unveiling the wind energy future of Türkiye with policies, technologies and potential. *Heliyon, 11*(4), Article e42592. https://doi.org/10.1016/j.heliyon.2025.e42592

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems, 30*, 5998-6008.

Wu, B., Wang, L., & Zeng, Y.-R. (2022). Interpretable wind speed prediction with multivariate time series and temporal fusion transformers. *Energy, 252*, Article 123990. https://doi.org/10.1016/j.energy.2022.123990

Xiong, X., Guo, X., Zeng, P., Zou, R., & Wang, X. (2022). A short-term wind power forecast method via XGBoost hyper-parameters optimization. *Frontiers in Energy Research, 10*, Article 905155. https://doi.org/10.3389/fenrg.2022.905155

Yu, S., & Hur, J. (2023). An enhanced performance evaluation metrics for wind power ramp event forecasting. *IEEE Access, 11*, 100195-100206. https://doi.org/10.1109/ACCESS.2023.3313632

Zeng, A., Chen, M., Zhang, L., & Xu, Q. (2023). Are transformers effective for time series forecasting? *Proceedings of the AAAI Conference on Artificial Intelligence, 37*(9), 11121-11128. https://doi.org/10.1609/aaai.v37i9.26317

Zhou, H., Zhang, S., Peng, J., Zhang, S., Li, J., Xiong, H., & Zhang, W. (2021). Informer: Beyond efficient transformer for long sequence time-series forecasting. *Proceedings of the AAAI Conference on Artificial Intelligence, 35*(12), 11106-11115. https://doi.org/10.1609/aaai.v35i12.17325