

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/373119058>

# Kompleks cephe tasarımlarında makine öğrenme eğrileri: Tahminleme modelleri doğru optimizasyonu beraberinde getiriyor mu?

Conference Paper · June 2023

CITATIONS

0

READS

106

1 author:



Berk Ekici

Izmir Institute of Technology

34 PUBLICATIONS 442 CITATIONS

SEE PROFILE

# Kompleks Cephe Tasarımlarında Makine Öğrenme Eğrileri: Tahminleme Modelleri Doğru Optimizasyonu Beraberinde Getiriyor mu?

BERK EKİCİ   
İzmir Yüksek Teknoloji Enstitüsü  
berkekici@iyte.edu.tr

## ÖZET

Birleşmiş Milletler Sürdürülebilirlik Kalkınma Hedefleri kapsamında sürdürülebilir bina performansına daha etkin bir şekilde ulaşılması hedeflenerek, araştırmacılar ve profesyoneller yapay zekâ yöntemlerini giderek artan bir şekilde tasarım süreçlerine dahil etmektedir. Makine öğrenmesi ile optimizasyon, bu bağlamda öne çıkan uygulamalar arasında yer almaktadır. Amaç, tahmin modelleri geliştirerek simülasyon gerektiren performans hedeflerini optimize ederken, gerekli olan hesaplama zamanının indirgenmesidir. Geliştirilen tahmin modellerinin doğruluğunu değerlendiren yöntemlerin genel yaklaşımı, toplanan sınırlı veri üzerinden bir çıkarım yapmaya yöneliktir. Halbuki, optimizasyon algoritmaları en iyi tasarım alternatifine ulaşabilmek için, arama alanlarındaki farklı bölgelere sıçrayabilme özelliğini taşımaktadır. Bu sebeple, sadece toplanan örnekler üzerinden tahmin modellerinin doğruluğunu test etmek, makine öğrenmesi ile optimizasyon sürecinde tahmin edilen sonucun gerçek arama alanında çok daha farklı bir sonuca karşılık gelmesi durumunu doğurabilmektedir. Bu çalışmanın amacı farklı makine öğrenme algoritmalarını kullanarak geliştirilen tahmin modellerinin optimizasyon süreçlerinde ne kadar doğru sonuçlar üretebildiğini araştırmaktır. Bu kapsamda makine öğrenmesi kullanarak optimizasyon sonuçlarının doğruluğunun test edildiği yeni bir yöntem sunulmuştur. 27 tasarım parametresine sahip 3 katlı bir binada yer alan kompleks bir parametrik cephe modeli, aydınlatma ve solar radyasyon simülasyon modellerine entegre edilerek geliştirilmiştir. Literatürde yaygın bir şekilde kullanılan yapay sinir ağları (ANN), çoklu doğrusal regresyon (MLR), destek vektör makineleri (SVM), rastgele orman (RF) ve karar ağacı (DT) algoritmaları 250 ile 2000 örneğe sahip toplanan her bir veri seti üzerinde aydınlatma ve solar radyasyon performans hedeflerine uygulanmıştır. Tahmin

sonularının optimizasyon sonuları ile kıyaslanması adına yüksek zaman gerektiren fonksiyon deęerlendirmeleri üzerine geliřtirilen radyal tabanlı iřlev yntemi ile optimizasyon (RBFOpt) algoritması tek hedefli ve iki kısıt fonksiyonu olan bir optimizasyon probleminde kullanılmıřtır. alıřma sonunda elde edilen sonular, sadece toplanan rneklere dayanan tahmin modellerinin optimizasyon srecinde hatalı sonulara neden olabileceęini gstermektedir. Makine ęrenmesi algoritmaları toplanan veriler hakkında doęru tahminlerde bulunabilirken, bu tahminler optimizasyon sreci boyunca toplanan verilerin dıřına ıktıęı zaman yanlış tahmin sonularını doęurabilmektedir. Bu kapsamda nerilmiř olan yntem karar vericiler iin hangi makine ęrenmesi algoritmasının daha doęru optimizasyon sonuları verebileceęi hakkında yardımcı olmaktadır.

**Anahtar Kelimeler:** Parametrik cephe tasarımı, Makine ęrenmesi, Optimizasyon, ęrenme eęrileri, Performansa dayalı tasarım, Srdrlebilir binalar

# Machine Learning Curves in Complex Façade Designs: Do predictive models bring along the correct optimization?

BERK EKİCİ   
İzmir Institute of Technology  
berkekici@iyte.edu.tr

---

## ABSTRACT

Within the scope of the United Nations Sustainability Development Goals, researchers and professionals are increasingly incorporating artificial intelligence methods into design processes to achieve sustainable building performance more effectively. Optimization with machine learning is one of the prominent applications in this context. The aim is to optimize performance targets that require simulation with predictive models while reducing the computational time required. The general approach of the methods that evaluate the accuracy of the developed prediction models is to make an inference based on the collection of limited data. However, optimization algorithms can jump to different regions in the search space in order to reach the best design alternative. For this reason, testing the accuracy of prediction models, only considering the collected samples, may lead to a situation where the predicted result in the optimization process may correspond to a different solution in the actual search space. This study aims to investigate the accuracy of the prediction models developed by different machine learning algorithms that can estimate accurate results in optimization processes. In this context, a new method is presented to test the accuracy of optimization results using machine learning. A complex parametric façade model of a 3-story building with 27 design parameters is developed by integrating lighting and solar radiation simulation engines. Artificial neural networks (ANN), multiple linear regression (MLR), support vector machines (SVM), random forest (RF) and decision tree (DT) algorithms, which are widely used in the literature, are applied to lighting and solar radiation performance targets on each collected data set with 250 to 2000 samples. In order to compare the prediction results



with the optimization results, the radial basis function optimization (RBFOpt) algorithm, which was developed for computationally expensive function evaluations, was used in an optimization problem having single objective and two constraint functions. The results obtained at the end of the study show that the prediction models, which only based on the collected samples, may cause incorrect results in the optimization process. While machine learning algorithms can make accurate estimations using the collected data, predictions can lead to incorrect results as they go beyond the sample set during the optimization process. In this context, the proposed method helps decision-makers to determine which machine learning algorithm can provide more accurate optimization results.

**Keywords:** Parametric façade design, Machine learning, Optimization, Learning curves, Performance-based design, Sustainable buildings

## 1. GİRİŞ VE PROBLEMİN TANIMI

Birleşmiş Milletler Sürdürülebilirlik Kalkınma Hedefleri (SDG) (UN, 2022) kapsamında binaların sürdürülebilir performans kriterlerinin daha etkin bir şekilde sağlanması hedeflenerek, araştırmacılar ve profesyoneller yapay zeka yöntemlerini giderek artan bir şekilde tasarım süreçlerine dahil etmektedir (Roman et al., 2020). Makine öğrenmesi ile optimizasyon, bu bağlamda öne çıkan uygulamalar arasında yer almaktadır (Westermann & Evins, 2019). Amaç, tahminleme modelleri geliştirerek simülasyon gerektiren performans hedeflerini optimize ederken, gerekli olan hesaplama zamanının kabul edilebilir sürelerle indirgenmesidir. Böylelikle, tasarım problemi için çoklu performans hedefleri göz önünde bulundurularak tasarım kararları alınabilir, büyük ölçekteki tasarımlarda kapsamlı tasarım alternatifleri irdelenir ve yüksek sayıda tasarım parametresini bünyesinde barındıran tasarım problemlerinde optimale en yakın sonuçlara erişim sağlanır.

Kompleks tasarım problemlerinde, tasarım parametrelerinin sayısı ve aldığı değerler göz önünde bulundurulduğu zaman, oldukça yüksek alternatifleri içerisinde barındıran arama alanları optimizasyon süreçleri içerisinde meydana gelmektedir (örneğin 10 değer alabilen 10 tasarım parametresinin olduğu bir problemde 1010 tasarım alternatifine sahip bir arama alanının meydana gelmesi). Oluşturulan tasarım modellerinden veri toplanarak geliştirilen tahminleme modelleri genellikle sınırlı sayıda tasarım örneğini barındırır (örneğin 1010 tasarım alternatifinin bulunduğu bir arama alanından 500 tasarım örneğinin toplanması). Kullanılan örnekleme yöntemleri tarafından arama alanlarının farklı bölgelerinden toplanan tasarım alternatifleri, girdi-çıkı ilişkisinin makine tarafından öğrenilebilmesini sağlamaktadır. Geliştirilen tahminleme modellerinin doğruluğunu değerlendiren yöntemlerin genel yaklaşımı, toplanan sınırlı veri üzerinden bir çıkarım yapmaya yöneliktir. Halbuki, optimizasyon algoritmaları sahip oldukları farklı yöntemler gereği, en iyi tasarım alternatifine ulaşabilmek için, arama alanlarındaki farklı bölgelere sıçrayabilme özelliğini taşımaktadır. Bu sebeple, sadece toplanan örnekler üzerinden tahminleme modellerinin doğruluğunu test etmek, makine öğrenmesi ile optimizasyon sürecinde tahmin edilen sonucun gerçek arama alanında çok daha farklı bir sonuca karşılık gelmesi durumunu doğurabilmektedir. Böylelikle, ulaşılan performans değerleri yanlış olabileceği gibi, tasarım kararları alınırken yanlış parametreler üzerinde değerlendirmeler yapılmasına da sebep olabilir.

Bu çalışmanın amacı farklı makine öğrenme algoritmalarını kullanarak geliştirilen tahminleme modellerinin optimizasyon süreçlerinde ne kadar

doğru sonuçlar üretebildiğini araştırmaktır. Literatürde yer alan makine öğrenmesi ile optimizasyon yöntemleri genel olarak (i) form üretimi, (ii) performans değerlendirmesi, (iii) veri üretimi, (iv) makine öğrenimi, (v) test kriterleri ve (vi) optimizasyon şeklinde kurgulanmaktadır (Roman et al., 2020). Bu araştırma kapsamında makine öğrenmesi sürecinde optimizasyon sonuçlarının doğruluğunun test edildiği yeni bir yöntem önerilmektedir. Sinüs fonksiyonu ile üretilen 27 tasarım parametresine sahip kompleks bir parametrik cephe modeli, 10m/30m/10.5m ölçülerine sahip 3 katlı bir binada aydınlatma ve solar radyasyon simülasyon modelleri entegre edilerek geliştirilmiştir. 5 farklı makine öğrenme algoritmasına ait öğrenme eğrilerinin irdelendiği bu araştırmada, simülasyona dayalı optimizasyon problemlerinde kullanılan bir optimizasyon algoritmasının sonuçları, geliştirilen tahminleme modelleri ile karşılaştırılmış ve önerilen yeni yöntemin katkıları bu çalışma kapsamında tartışılmıştır.

## 2. YÖNTEM

Makine öğrenmesi ile optimizasyon kapsamında önerilen yöntem, parametrik cephe modeli, aydınlatma ve solar radyasyon performans analizleri, çeşitli makine öğrenmesi ve optimizasyon algoritmalarını bünyesinde barındırmaktadır. Yöntem ve araştırma sürecinde uygulanan adımlar aşağıda açıklanmıştır.

### 2.1. Önerilen Makine Öğrenmesi ile Optimizasyon Yöntemi

Form üretimi ve performans değerlendirmesini kavramsal tasarım sürecinde bir otomasyona dönüştüren optimizasyon algoritmalarının mimarlıktaki temel amacı, nicel ve nitel performans hedeflerindeki avantajları maksimize etmek, dezavantajları minimize etmektir (Sarıyıldız, 2012). Sürdürülebilirlik kapsamında ele alınan performans hedefleri, örneğin aydınlatma, enerji tüketimi, solar radyasyon, görsel ve termal konfor vb., optimizasyon süreçlerinde hesaplanırken simülasyon motorlarının beraberinde getirdiği yüksek hesaplama süreleri nedeniyle oldukça yüksek optimizasyon süreleri ortaya çıkmaktadır. Hem tasarımların kavramsal aşamadan itibaren sürdürülebilirlik ile ilgili performanslarını optimize etmek hem de bu süreç için oldukça fazla hesaplama zamanına ihtiyaç duyulması bir ikilem yaratmaktadır. Bu ikileme bir çözüm olarak araştırmacılar ve profesyoneller içerisinde makine öğrenim süreçlerini de kapsayan, RBFOpt (Costa & Nannicini, 2018) gibi optimizasyon algoritmalarına yer vermektedir. Bu algoritmalar, genetik algoritma gibi meta sezgisel yöntemlerden daha az fonksiyon değerlendirmesi yaparak iyi bir tasarım sonucu bulmak üzere geliştirilmiştir. Ancak, büyük

ölçekli tasarım problemlerinde RBFOpt gibi algoritmalar, makine öğrenmesi süreçlerini barındırmalarına rağmen, optimale yakın sonuçları keşfedebilmek için oldukça yüksek sürelerle ihtiyaç duymaktadır (Ekici et al., 2022). Bu sebeple, bahsi geçen ikilime ikinci bir çözüm olarak, bu araştırmanın da odağında olan, makine öğrenmesi algoritmaları tarafından oluşturulan tahminleme modellerinin hedef fonksiyonlar olarak kullanıldığı optimizasyon süreçleri karşımıza çıkmaktadır.

Makine öğrenmesi ile optimizasyonun bir türü olan bu yaklaşım, genel olarak (i) form üretimi, (ii) performans değerlendirmesi, (iii) veri üretimi, (iv) makine öğrenmesi, (v) test kriterleri ve (vi) optimizasyon şeklinde uygulanmaktadır. Ancak bu yaklaşım, sınırlı sayıda toplanan örnekler üzerinden tahminleme modellerinin doğruluğunu test etmek üzerine kurgulandığı için, optimizasyon sürecinde edinilen sonucun gerçekten çok daha başka bir sonuca karşılık gelmesi durumunu meydana getirebilmektedir. Bu sebeple, literatürdeki uygulamanın farklı bir versiyonu bu çalışma kapsamında önerilmektedir (Şekil 1). Önceki çalışmalar ile karşılaştırıldığı zaman iki yaklaşım arasındaki temel fark, tahminleme modellerinin toplanan örnekler ve yapılan kısa süreli bir optimizasyon sonucu üzerinden değerlendiriliyor olmasıdır. Bu sayede, sadece tahminleme modellerinin kendi veri setleri içerisinden alınan doğrulama örnekleri değil, optimizasyon sürecindeki arama alanında bulunan oldukça başka bir bölgeye ait bir tasarım örneği de kullanılarak, hedefleme modellerinin doğruluğu hakkında daha objektif bir değerlendirme amaçlanmıştır.

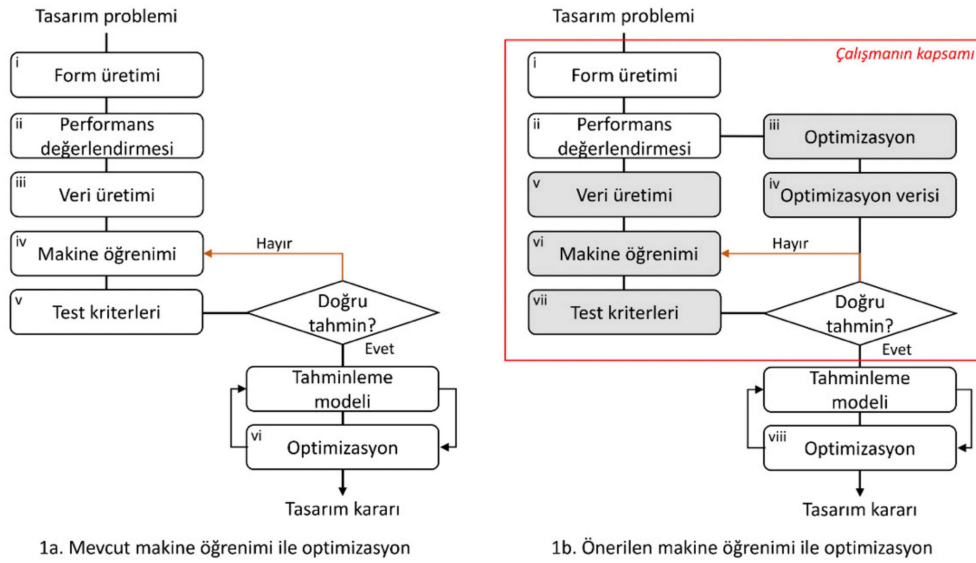
## 2.2. Parametrik Cephe Modeli

Önerilen yöntemde kullanılmak üzere cephe uzunluğu ve yüksekliği 30m/10.5m, mekân derinliği 10m olan İzmir'deki 3 katlı bir binanın belli bir kesimi üzerinde çalışılmıştır. Katlardaki her bir uzun kenar kullanılmak üzere, organik bir cephe tasarımı oluşturmak için Denklem 1'de verilen sinüs fonksiyonu 4 kenar için ayrı olarak uygulanmıştır.

$$a \times \sin(b \times x - c) + d \quad (1)$$

a parametresi oluşturulan organik cephedeki sapmayı, b parametresi sinüs periyodunu, c parametresi sinüs periyodunun yatay eksenindeki ötelenmesini ve d parametresi sinüs periyodundaki dağılım dengesini kontrol ederken, x parametresi pencere açıklığı sayısı kadar lineer olarak dizilmiş [0,3] arasındaki sayıları ifade etmektedir. Elde edilen organik formun sinüs etkisini arttırmak veya azaltmak için bütün denklem ile çarpılan ayrıca F parametresi

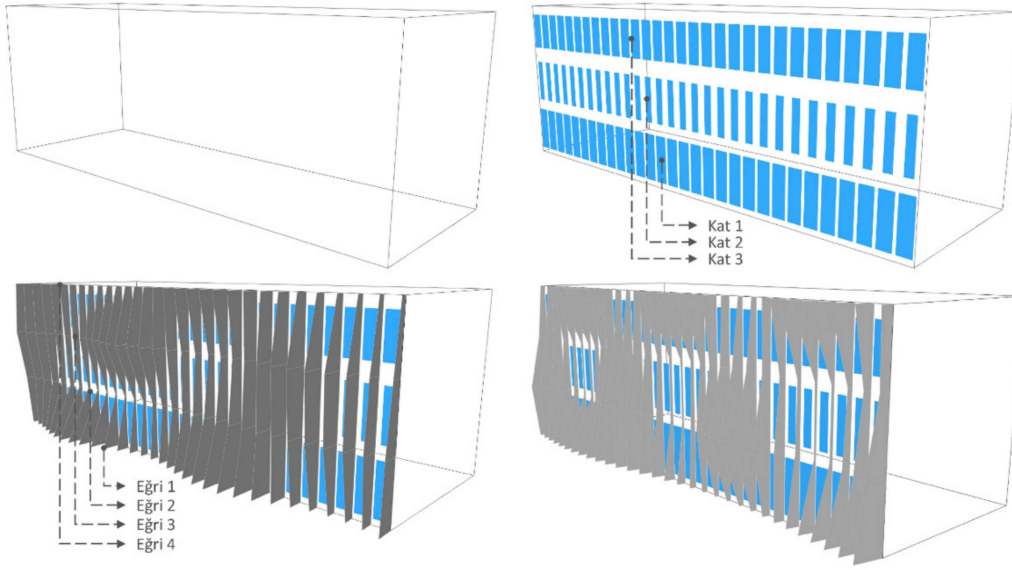
tanımlanmıştır. Oluşturulan yeni eğriler kullanılarak düşey gölgeleme elemanları üretilmiştir. Bütün binadaki düşey gölgeleme elemanlarının rotasyonunu sağlayan bir derece ( $r$ ) parametresine ek olarak, her bir katta bulunan pencere açıklıklarını  $x$  ve  $y$  yönlerinde ölçekleyen 6 adet parametre göz önünde bulundurulmuştur. Toplamda 27 parametrenin bulunduğu modele ait detaylar Tablo 1’de verilmiş olup, Grasshopper 3D (McNeel, 2014) ile geliştirilen parametrik modele ait farklı süreçlerden alınan görseller Şekil 2’deki gibidir.



**Şekil 1:** Önerilen yöntem

**Tablo 1:** Cephe tasarım modelinin parametreleri

Parametre	Açıklama	Eğri				Kat #			Birim	Aralık	Adet
		1	2	3	4	1	2	3			
$a$	Sapma	✓	✓	✓	✓				-	[-1.000, 1.000]	4
$b$	Periyod	✓	✓	✓	✓				$\pi$	[0.000, 6.000]	4
$c$	Yatay öteleme	✓	✓	✓	✓				-	[0.000, 5.000]	4
$d$	Dağılım dengesi	✓	✓	✓	✓				-	[1.000, 2.000]	4
$F$	Etki	✓	✓	✓	✓				$m$	[0.005, 1.000]	4
$x$	X yönü açıklık ölçeği					✓	✓	✓	-	[0.05, 0.95]	3
$y$	Y yönü açıklık ölçeği					✓	✓	✓	-	[0.05, 0.95]	3
$r$	Düşey eleman dönüş derecesi								$^\circ$	[-60.00, 60.00]	1



**Şekil 2:** Parametrik cephe tasarım örnekleri

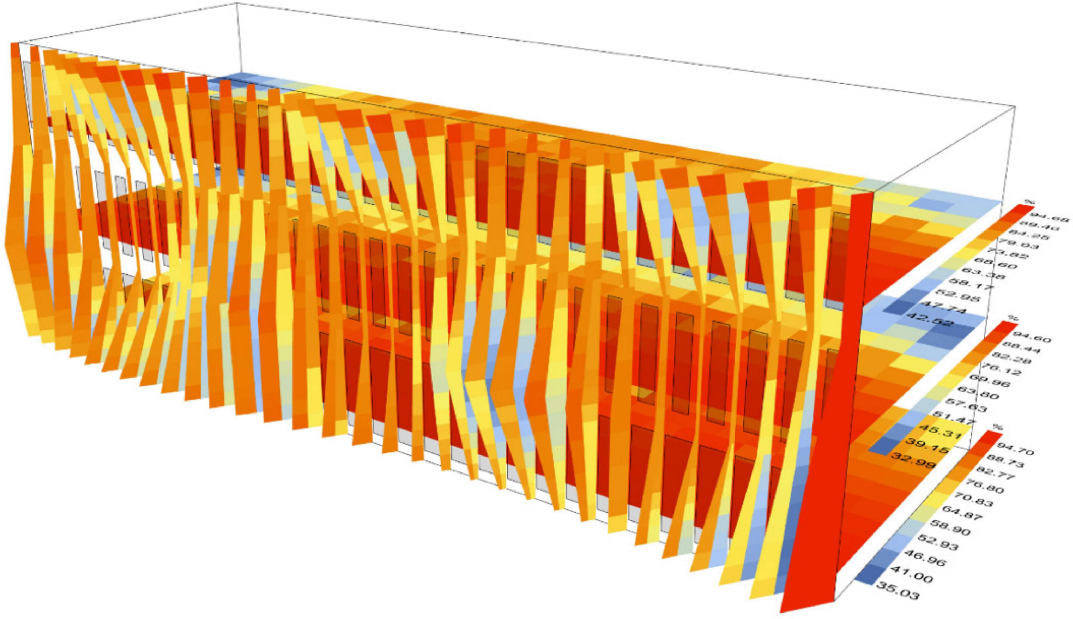
### 2.3. Performans Kriterleri ve Veri Üretimi

Geliştirilen parametrik modeldeki düşey gölgeleme elemanlarına fotovoltaik paneller yerleştirilerek bütün bir yıl boyunca güneş enerjisinden fayda sağlanması amaçlanmıştır. Bu hedef için binadaki bütün düşey elemanlara solar radyasyon simülasyonu entegre edilmiş olup bütün bir yıl elde edilen yenilenebilir enerji gücü ( $E_y$ ) hesaplanmıştır. Optimizasyon sürecinde bu hedef ( $F_x$ ) maksimize edilmek istendiği için yüksek yoğunlukta düşey elemanların oluşması muhtemeldir. Güneş ışığı yetersizliği sebebi ile fazladan yapay aydınlatma kullanılarak enerji sarfiyatının önüne geçmek için her bir katta devamlı güneş ışığı otonomisi ( $cDA$ ) metriği kullanılarak üç adet güneş ışığı simülasyon motoru modele entegre edilmiştir (Şekil 3). Her bir kattaki ortalama güneş ışığı seviyesi ( $cDA_{(1,...,3)}$ ) ve dağılımı ( $dist_{(1,...,3)}$ ) olmak üzere toplam 6 kısıt fonksiyonu ( $V_x$ ) belirlenmiştir. Latin hiper küp örnekleme (LHS) yöntemi kullanılarak 250 ile 2000 arasında oluşturulan 8 farklı veri seti uzunlukları için, geliştirilen parametrik model ve Ladybug araçları (Roudsari et al., 2013) kullanılarak yapılan simülasyonlar üzerinden, tasarım örnekleri ve performans sonuçları toplanmıştır (Şekil 4).

### 2.4. Makine Öğrenmesi

Bu çalışma kapsamında literatürde yaygın bir şekilde kullanılan yapay sinir ağları (ANN), çoklu doğrusal regresyon (MLR), destek vektör makineleri (SVM), rastgele orman (RF) ve karar ağacı (DT) algoritmaları toplanan her bir veri seti üzerinde uygulanmıştır. Makine öğrenme algoritması test edilirken kullanılan

veri setleri 0.2 oranında öğrenme ve test setleri olarak ikiye ayrılmıştır. Yapay sinir ağı modeli 1-n-n-1 mimarisine sahip olmakla her bir katmanda 50 nöron (n) kullanılmıştır. Stokastik gradyan inişi algoritması kullanılarak geliştirilen modellerde 300 iterasyon ile her bir performans kriterinin öğrenme süreçleri tamamlanmıştır.



**Şekil 3:** Solar radyasyon ve güneşiği analizleri

Çoklu doğrusal regresyon öğrenme ve test setleri üzerinde direk uygulanarak sonuçlar raporlanmıştır. Destek vektör makinelerinde ise radyal tabanlı bir işlev yöntemi kullanılarak tahminleme modelleri geliştirilmiştir. Rastgele ormanda tahminleyicilerin sayısı 30 olarak belirlenirken karar ağacının parametreleri varsayılan değerleri ile kullanılmıştır. Elde edilen tahminleme modelleri test setleri kullanılarak tahminleme doğruluğu Denklem 2’de verilen  $R^2$ , Denklem 3’teki ortalama kare hatası (MSE) ve Denklem 4’deki ortalama mutlak hata (MAE) yöntemleri ile analiz edilmiştir. Farklı istatistik yöntemlere yer verilmesi, oluşturulan tahminleme modellerinin farklı açılardan doğruluğunun kontrol edilmesini sağlamaktadır. Bu kapsamda  $x_i$  toplanan performans verisini,  $\bar{x}$  toplanan performans verisinin ortalamasını,  $y_i$  tahmin edilen performans verisini ve  $n$  toplanan performans verisinin uzunluğunu ifade etmektedir. Bunun yanı sıra, her bir algoritmanın ihtiyaç duyduğu öğrenme süreleri saniye (s) bazında kaydedilmiştir. Bu süreçte bütün adımlar Python programlama dilindeki Scikit-learn v1.2.1 (Pedregosa et al., 2011), Keras v2.10.0 (Chollet, 2015), TensorFlow v2.10.1 (Abadi et al., 2016) kütüphaneleri kullanılarak tamamlanmıştır.

$$R^2 = 1 - \frac{\sum_i (x_i - y_i)^2}{\sum_i (x_i - \bar{x})^2} \quad (2)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2 \quad (3)$$

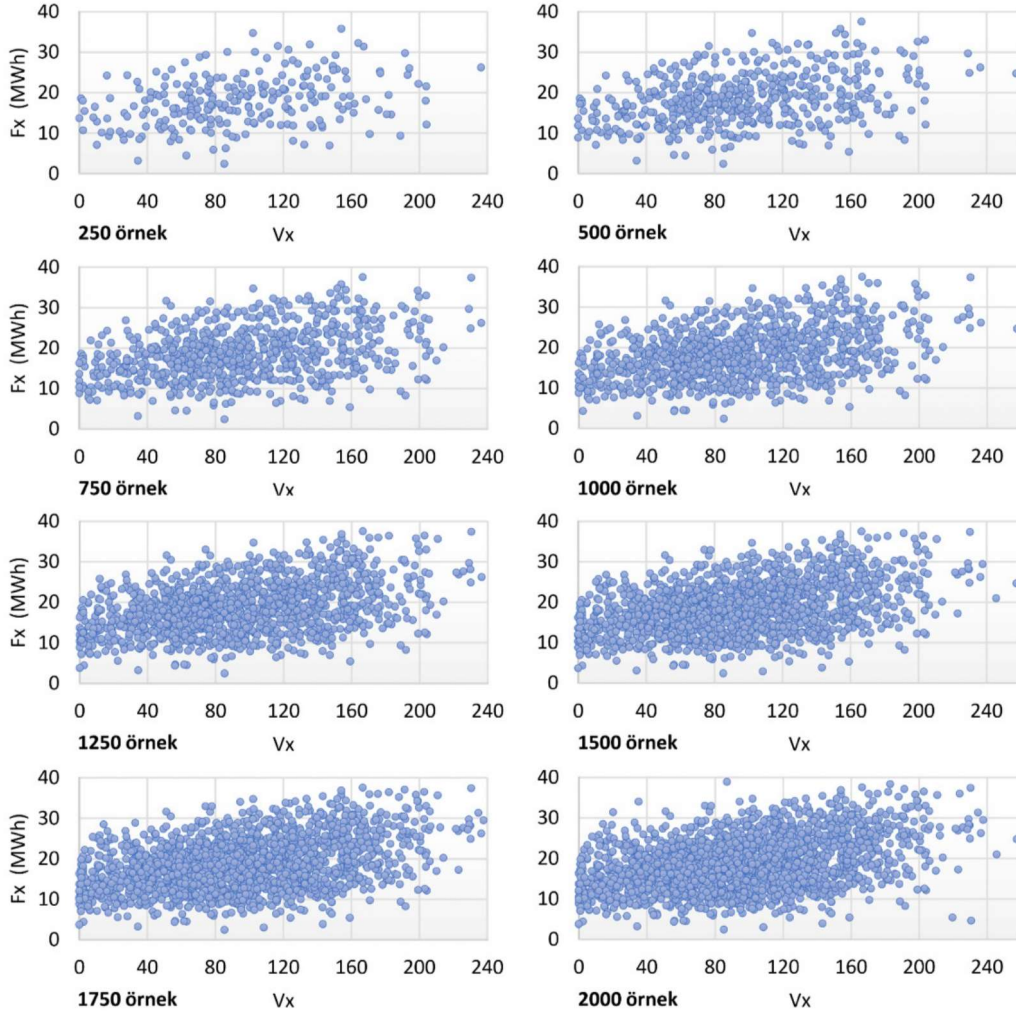
$$MAE = \frac{\sum_{i=1}^n |y_i - x_i|}{n} \quad (4)$$

## 2.5. Optimizasyon ve Önerilen Yöntemin Doğrulanması

Geliştirilen modelin 1 tasarım alternatifi için gereken hesaplama süresi yaklaşık 1 dk. olarak ölçülmüştür. Bu sebeple, yüksek zaman gerektiren fonksiyon değerlendirmeleri üzerine geliştirilen radyal tabanlı işlev yöntemi ile optimizasyon (RBFOpt) algoritması Opossum eklentisi (Wortmann, 2017) kullanılarak, Denklem 5'teki optimizasyon problemi için modele entegre edilmiştir. Toplamda 6 farklı kısıt fonksiyonu tanımlanması sebebiyle, standart bir penaltı fonksiyonu yerine, Optimus (Cubukcuoglu et al., 2019) eklentisinin içerisinde bulunan fizibilite eşliğine yakın (NFT) kısıt işleme yöntemi optimizasyon sürecine dahil edilmiştir.

$$\begin{array}{ll} \text{maks.} & E_y \\ \text{kısıt} & cDA_{1,\dots,3} \geq 75\% \\ & dist_{1,\dots,3} \geq 50\% \end{array} \quad (5)$$





**Şekil 4:** LHS örnekleme yöntemi ile toplanan farklı büyüklüklerdeki veri setlerinin Fx ve Vx dağılımları

Amaç,  $V_x=0$  iken ulaşılabilecek en yüksek  $E_y$  değerine ulaşmaktır. Optimizasyon sonrasında elde edilen sonuçlar, geliştirilen makine öğrenme modelleri kullanılarak tahmin edilen Fx ve Vx değerleri ile karşılaştırılmıştır. Böylelikle, önerilen yöntemin doğruluğu, literatürdeki pek çok örnek gibi sadece toplanan veri setleri üzerinden değil, veri seti dışarısında olabilecek optimizasyon sonucunda elde edilmiş olan parametre seti üzerinden de test edilmiştir.

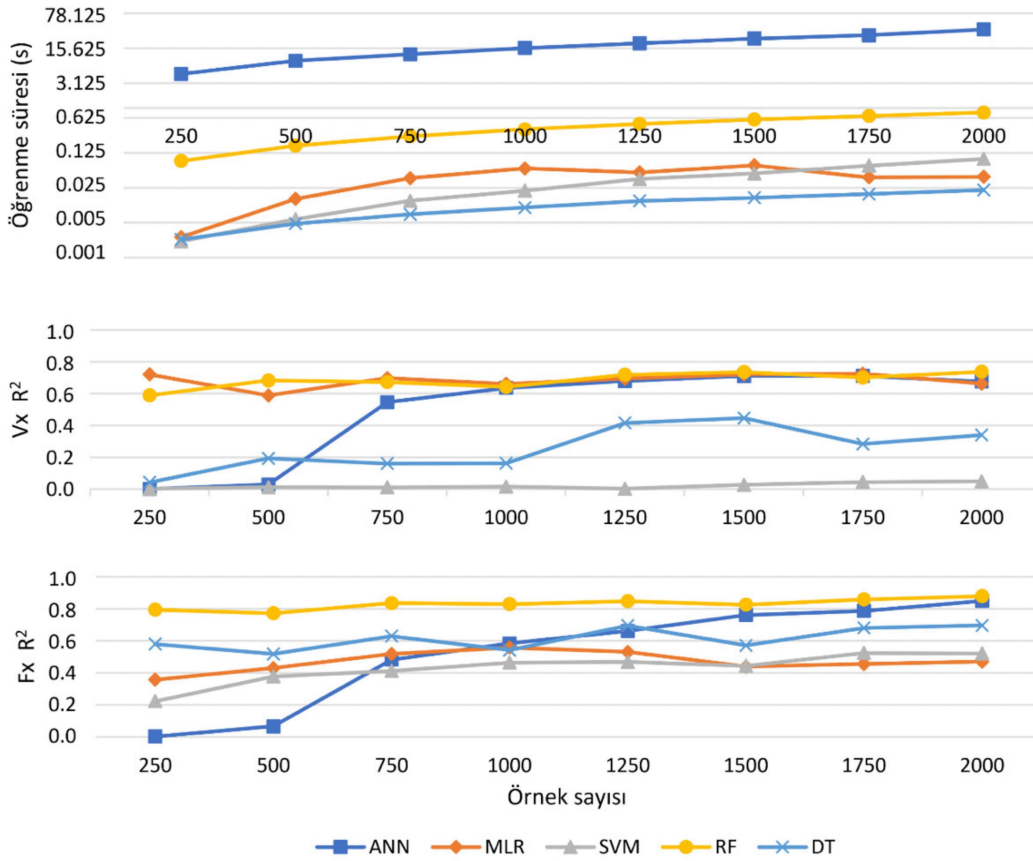
### 3. SONUÇLAR

Bu bölümde toplanan farklı boyutlardaki 8 veri seti üzerinde uygulanmış olan makine öğrenme sonuçları ile tamamlanmış olan optimizasyon sonuçları karşılaştırılmış ve tahminleme modellerinin doğruluğu tartışılmıştır. Yapılan makine öğrenmesi ve optimizasyon uygulamalarında Intel i7 – 12700H 2.50 GHz işlemcili, 32-GB DDR5 ram bellek, 1024-GB SSD sabit

disk ve Nvidia GeForce RTX 3060 6GB ekran kartı olan bir laptop bilgisayar kullanılmıştır. ANN modellerinin öğrenme süreçlerinde ekran kartı, diğer makine öğrenme süreçlerinde ise bilgisayarın işlemcisi gerekli hesaplamaların tamamlanmasında kullanılmıştır.

### 3.1. Makine Öğrenme Sonuçları

Makine öğrenme algoritmalarının tahminleme performanslarına dair ilk karşılaştırma Şekil 5'te verilen hesaplama süreleri ve tahminleme modellerinin  $R^2$  değerleri üzerinden yapılmıştır. Ekran kartı kullanımına rağmen, veri setlerinin uzunluklarıyla orantılı bir değişim gösteren ANN modelleri, 4.77 ile 37.17 saniye arasında öğrenim süreçlerini tamamlayabilmıştır. Diğer bütün makine öğrenme algoritmaları, işlemci kullanımına rağmen, 1 saniyeden daha kısa bir sürede aynı hesaplamaları tamamlamışlardır.

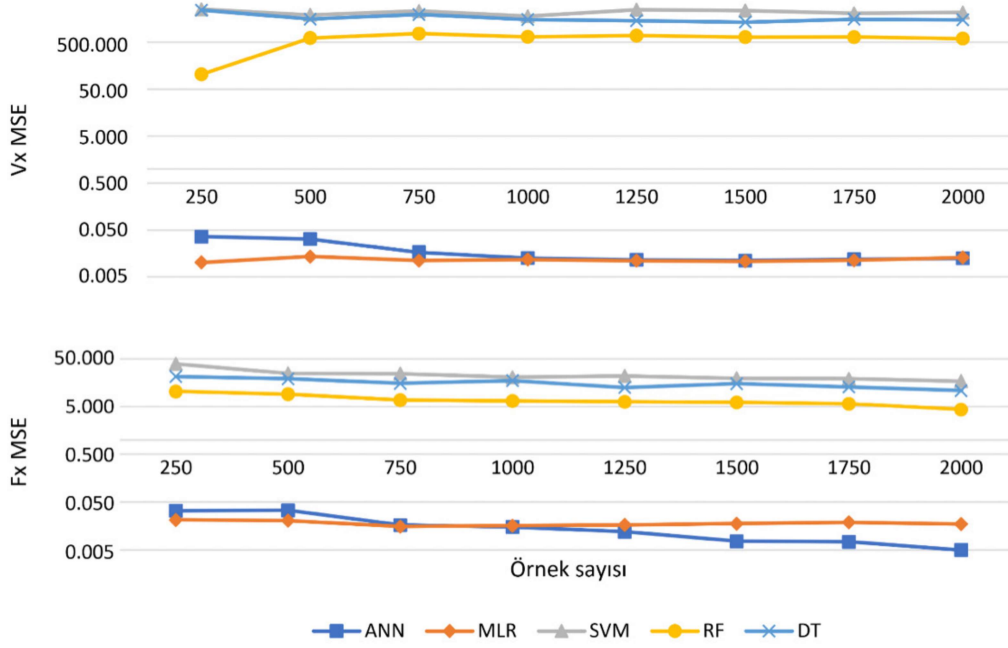


Şekil 5: Öğrenme süresi (saniye) ve  $R^2$  sonuçları

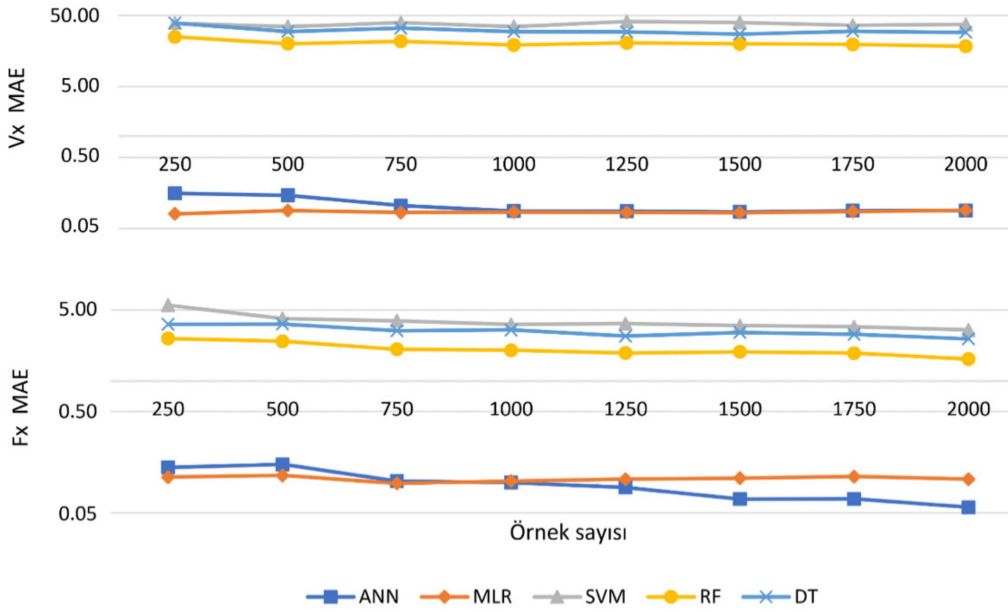
$R^2$  değerleri karşılaştırıldığı zaman, ANN modellerinin veri setlerindeki örnek artışına bağlı olarak kısıt ( $V_x$ ) fonksiyonu için öğrenme skorlarının 0.0006'dan 0.71'lere, hedef ( $F_x$ ) fonksiyonları için 0.002'den 0.84'lere yükseldiği

gözlemlenmiştir. MLR modelleri Vx tahminlerinde 0.58 ile 0.72 arasında R2 sonuçları verirken, bu değerler Fx için 0.35 ile 0.55 arasındadır. SVM modelleri Vx tahminlerinde 0.002 ile 0.04 arasında R2 sonuçları ile başarısız tahminleme modelleri üretirken, Fx için 0.25 ile 0.52 arasındadır. RF yöntemi veri setindeki artışlara rağmen oluşturulan bütün modellerde Vx için 0.58 ile 0.73 arasında ve Fx için 0.77 ile 0.87 arasında performans değerleri göstererek diğer yöntemlerden daha tutarlı bir öğrenme eğrisi sergilemiştir. Son olarak DT modellerinde veri seti uzunluğundaki artışa bağlı olarak Vx modellerindeki artışın 0.04'ten 0.44'e yükseldiği ve Fx tahminleme modelleri için 0.51 ile 0.69 arasında tahminleme skorlarına ulaştığı gözlemlenmiştir.

Sadece R2 sonuçları değerlendirilerek doğru tahmin yapabilen modellere erişim sağlanamayacağı için MSE (Şekil 6) ve MAE (Şekil 7) sonuçları da çalışma boyunca raporlanmıştır. Oluşturulan bütün Vx modellerine bakıldığı zaman, SVM, RF, ve DT makine öğrenme yöntemlerinin 103.74 ile 2516.94 arasında çok yüksek MSE sonuçlarına ulaştıkları gözlemlenmiştir. Fx modelleri açısından 4.35 ile 39.08 arasında yine yüksek hatanın oluştuğunu gösteren MSE değerlerine erişilmiştir. Yalnızca ANN ve MLR modellerinin MSE sonuçları 0.005 ile 0.03 arasında olması sebebiyle başarılı modeller olarak kategorize edilmiştir. Benzer bir durum MAE sonuçlarında da gözlemlenmiştir. Oluşturulan bütün Vx modellerine bakıldığı zaman, SVM, RF, ve DT makine öğrenme yöntemlerinin 18.45 ile 41.23 arasında çok yüksek MAE skorlarını raporladıkları gözlemlenmiştir. Fx modelleri ise 1.64 ile 5.60 arasında yine yüksek sayılabilecek MAE sonuçlarını vermektedir. MSE sonuçlarında olduğu gibi, yalnızca ANN ve MLR modellerinin MAE skorları 0.05 ile 0.15 arasında olması sebebiyle başarılı modeller olarak nitelendirilmiştir. Sonuç olarak, R2, MSE ve MAE sonuçlarından başarılı tahminleme yapabilen modellerin ANN ve MLR olarak saptanması sebebiyle, bir sonraki aşamada sadece bu yöntemlerin sonuçlarına yer verilmiştir.



Şekil 6: Ortalama kare hatası sonuçları

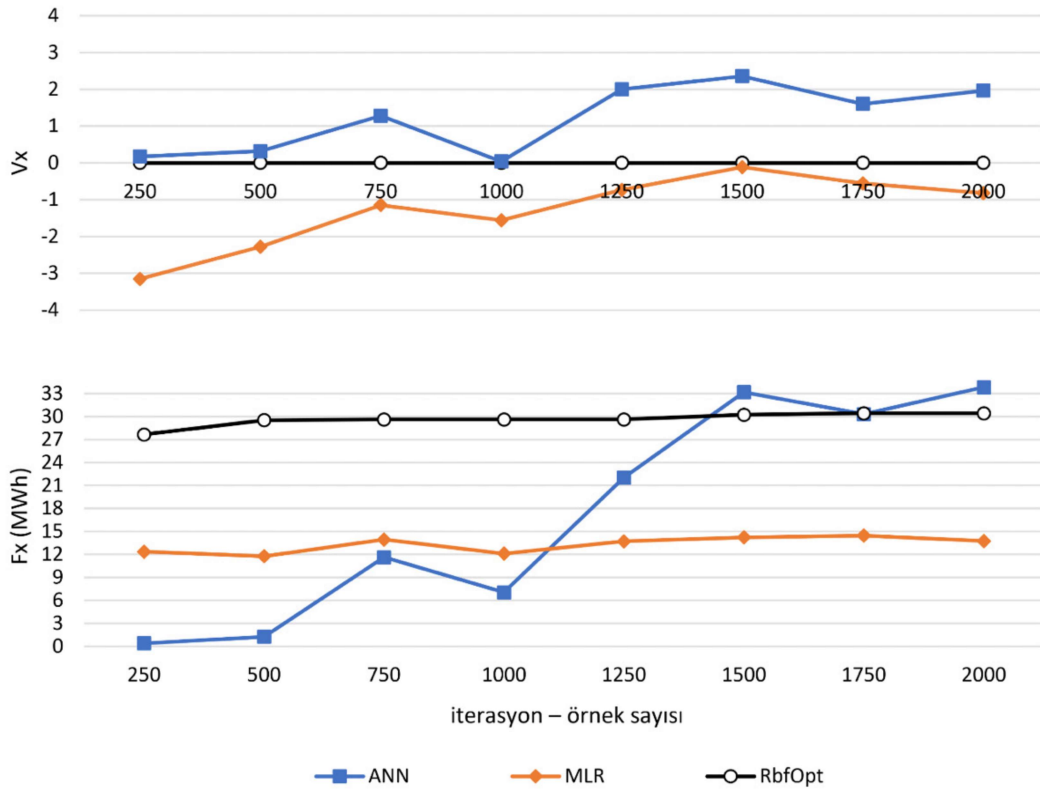


Şekil 7: Ortalama mutlak hata sonuçları

### 3.2. Tahminleme Modellerinin Doğruluğu

Yapılan çalışmalar sonucunda daha detaylı tahminleme analizlerinin yapılmasına karar verilen ANN ve MLR makine öğrenme yöntemlerine ait öğrenme eğrileri, oluşturulan her bir veri seti için RbfOpt optimizasyon

algoritmasının elde ettiği  $V_x$  ve  $F_x$  değerleri ile karşılaştırmalı olarak Şekil 8'de verilmiştir. Optimizasyon sürecinde, 32. iterasyonda bütün kısıtları ( $V_x=0$ ) sağlamayı başarabilen RbfOpt algoritması, 12.84 MWh enerji üretimini sağlayabilen bir cephe tasarım alternatifine erişmiştir. 2000'inci iterasyona kadar bu üretim miktarı 30.43 MWh'e kadar yükselmiş olsa da son 500 iterasyon boyunca hedef fonksiyon sonucunda çok büyük değişiklikler gözlemlenmemiştir. Optimize edilmiş cephe tasarımı alternatifinden elde edilen tasarım parametrelerine ait değerler bir önceki adımda oluşturulan makine öğrenimi modellerinde kullanıldığı zaman Şekil 8'deki tahminleme sonuçlarına ulaşılmıştır.



**Şekil 8:** Tahminleme modellerinin doğruluğu

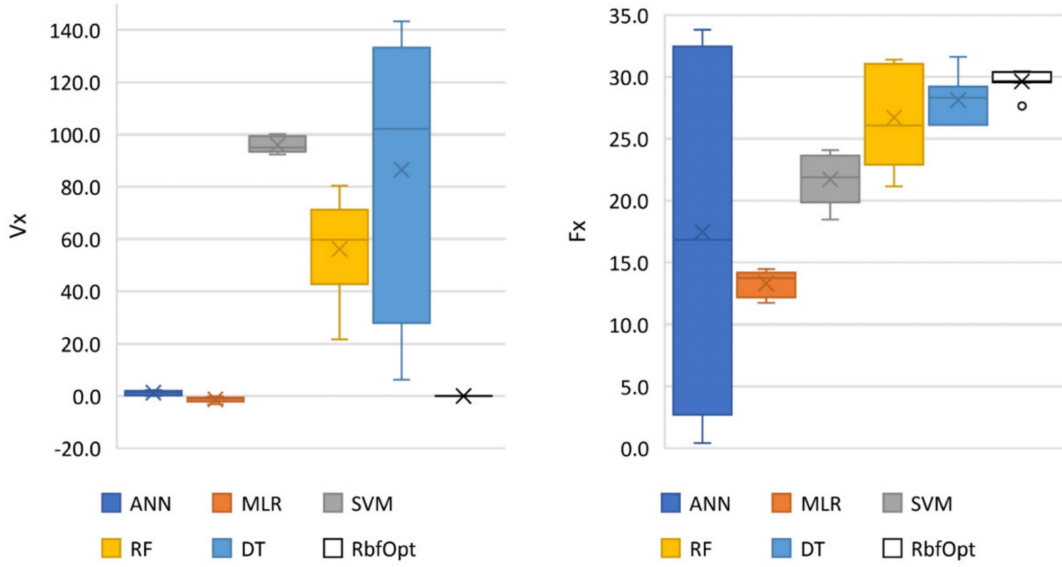
ANN modelleri düşük veri setlerinde, optimize edilmiş gerçek  $F_x$  değerlerinden oldukça uzak tahminleme performansı sergilemiştir. Ancak, veri setlerindeki örnek miktarının artması ile birlikte 1000 adet ve üzeri örneğe sahip veri setlerinden itibaren gerçek optimizasyon sonuçlarına doğru giderek artan bir öğrenme eğilimi gözlemlenmiştir. 1500 ile 2000 arasında örneğe sahip veri setleri kullanılarak oluşturulan tahminleme modelleri 30.30 MWh ile 33.81 MWh arasında tahminlemeler yaparak optimizasyon sonucunda elde edilen 30.43 MWh'lik gerçeğe en yakın sonuçları vermiştir. ANN modellerinin  $V_x$  tahmin sonuçlarına bakıldığında kısıt fonksiyonlarına ait tahminlemelerin

0.03 ile 2.35 arasında olduğu, RbfOpt algoritmasının 0.0 kısıt sonucuna kabul edilebilir uzaklıklarda tahminler yaptığı gözlemlenmiştir.

MLR modelleri ise düşük ve yüksek örneğe sahip veri setlerinde Fx tahminlerinin 11.76 MWh ile 14.47 MWh arasında olduğu, optimizasyon sonucunda elde edilen 30.43 MWh'lik gerçek değere oldukça uzak olduğu gözlemlenmiştir. Vx tahmin sonuçlarına bakıldığı zaman kısıt fonksiyonlarına ait tahminlemelerin -3.15 ile -0.12 arasında olduğu, RbfOpt algoritmasının 0.0 kısıt sonucuna kabul edilebilir uzaklıklarda tahminler yaptığı gözlemlenmiştir. ANN modellerinin aksine, veri setindeki artış MLR modellerinin Vx tahminlemelerinde 0 ile -1 arasında bir iyileşmeye doğru gitmesine sebep olduğu gözlemlenmiştir. Ancak, MLR modellerinin Fx tahminlemelerinde gerçek sonuçlardan oldukça uzak tahminler yapması sebebiyle çalışma kapsamında ANN modellerinin daha gerçekçi optimizasyon sonuçlarını tahmin ettiği sonucunu ortaya çıkmıştır. SVM, RF ve DT sonuçlarını da içeren bütün tahminleme ve optimizasyon sonuçlarına ait kutu grafikleri Şekil 9'da verilmiştir.

### 3.3. Çalışmanın Limitleri ve Uygulanabilirliği

Bu çalışma kapsamında oluşturulan veri setlerindeki örnek ve optimizasyon sürecindeki iterasyon sayıları en fazla 2000 olacak şekilde sınırlandırılmıştır. Bu kararın başlıca sebeplerinden biri 2000 iterasyon boyunca çalışan RbfOpt algoritmasının süreci tamamlamak için ihtiyaç duyduğu süre 33.3 saat olarak gözlemlenmiştir. Veri setlerindeki örnek sayısı ANN gibi bazı makine öğrenmesi algoritmalarında performans tahminleme skorlarının artışına sebep olabilmektedir. Bunun yanı sıra, performans tahminleme modelleri ile daha gelişmiş hibrit sezgisel optimizasyon algoritmaları kullanılarak 10000 ve üzeri iterasyonlardaki tasarım sonuçlarına kabul edilebilir hesaplama süreleri içerisinde ulaşılabilir. Bu sayede optimizasyona bağlı tasarım araştırmalarının derinliği artırılabilir.



**Şekil 9:** Çalışma boyunca elde edilen tahminleme ve optimizasyon sonuçlarına ait kutu grafikleri

Kullanılan makine öğrenmesi algoritmalarında bulunan parametrelerin ayarlamaları yine bu çalışmanın limitleri arasındadır. Mimari uygulamalara odaklanan bazı makine öğrenmesi çalışmaları gösteriyor ki, öğrenim süreçlerini kontrol eden farklı parametre ayarları analiz edilerek tahminleme modellerindeki doğruluk artırılabilir (Chatzikonstantinou & Sariyildiz, 2016). Ancak, bu çalışmanın sonuçları doğrultusunda parametre ayarı yapıldıktan sonra, optimizasyon sonuçlarına ait tahminleme skorlarının tekrar karşılaştırılıp nihai model seçimi yapılması önem arz etmektedir. Bunun yanı sıra, k katlı çapraz geçerlilik gibi daha detaylı istatistik analizlere öğrenim süreçlerinde yer vermek, oluşturulan veri setlerinin farklı bölgelerinden örnekler alınarak geliştirilen tahminleme modellerindeki doğruluğun artmasını sağlayabilir.

Önerilen makine öğrenimi ile optimizasyon yönteminde 27 tasarım parametresine sahip 3 katlı bir yapının cephe tasarımına odaklanılmıştır. Tasarım probleminin sahip olduğu toplamdaki tasarım alternatif sayısı, yani optimizasyon arama alanında mevcut olan alternatif sayısı,  $5.06e+80$  olarak hesaplanmıştır. Farklı parametre boyutlarına sahip tasarım problemlerinde arama alanları büyüyüp küçüleceğinden, tahmin edilen ve optimize edilen tasarım sonuçları değişkenlik gösterebilir. Bu sebeple, farklı parametre sayılarının olduğu benzer tasarım problemleri üzerinden bu çalışma kapsamında önerilen makine öğrenimi ile optimizasyon yönteminin uygulanması, elde edilen sonuçların genellenebilirliği açısından faydalı olabilir. Ayrıca, tek hedef yerine çok hedefli optimizasyon problemlerinde yapılacak tahminlerin



bütün hedef fonksiyonlarını kapsayacak şekilde çalışmanın kurgulanması gerekmektedir. Diğer bir deyişle, farklı iterasyonlardan alınacak olan Pareto optimal kümelerin doğru tahmin edilmesi üzerinden bir değerlendirme yapılmalıdır.

Çalışma kapsamında dijital olarak üretilen bir cephe modeli ve çeşitli simülasyon motorlarına yer verilmiştir. Gerçek hayattaki bir tasarım problemi için bu çalışmada önerilen yöntemin uygulanması durumunda, simülasyon modellerinin ampirik veriler ile karşılaştırılması ve dijital modellerin kalibrasyonlarının yapılması gerekmektedir. Kalibre edilen simülasyon modelleri üzerinden yine önerilen yöntem dahilinde dijital veriler toplanabilir ve tasarım problemi optimize edilerek binalarda performans iyileştirmeleri sağlanabilir. Ulaşılan nihai tasarımın gerçek hayatta uygulanması, performans iyileştirmelerinin yer aldığı yeni ampirik veriler ile optimizasyon sonuçlarının kıyaslanmasına ve böylece önerilen yöntemin gerçek hayatta sağladığı faydaların doğrulanması adına önemli bir rol oynayabilir.

### 3.4. Varılan Sonuç

Yapılan çalışmalar sonucunda elde edilen veriler, sadece toplanmış olan örnekler üzerinden optimizasyon süreçlerinde kullanılmak üzere geliştirilen tahminleme modellerinin doğruluğu hakkında bir çıkarım yapmanın yanlış model tercihlerine götürebileceğini göstermektedir. Makine öğrenmesi algoritmaları toplanan veriler hakkında doğru tahminlerde bulunabilirken, bu tahminler optimizasyon süreci boyunca toplanan verilerin dışına çıktığı zaman yanlış tahminleme sonuçlarını doğurabilmektedir. Bu sebeple, hızlı bir ön optimizasyon yöntemi ile toplanan veri dışında bir optimizasyon verisinin tahminleme modelleri üzerinde kullanılması, daha doğru model seçimini beraberinde getirebilmektedir. Bunun yanı sıra birden fazla istatistik yöntem ile modellerin değerlendirilmesi yine büyük önem arz etmektedir. Farklı makine öğrenmesi algoritmalarının farklı bina performansı kriterlerinde göstermiş oldukları davranış sonrası, ileriki çalışmalarda bir binadaki çoklu performans tahminlerinde hibrit makine öğrenimi süreçlerine yer verilebilir.



## KAYNAKLAR

- Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., & Isard, M. (2016). Tensorflow: a system for large-scale machine learning. Osd Savannah, GA, USA.
- Chatzikonstantinou, I., & Sariyildiz, S. (2016). Approximation of simulation-derived visual comfort indicators in office spaces: a comparative study in machine learning. *Architectural Science Review*, 59(4), 307-322.
- Chollet, F. (2015). keras, <https://keras.io/>
- Costa, A., & Nannicini, G. (2018). RBFOpt: an open-source library for black-box optimization with costly function evaluations. *Mathematical Programming Computation*, 10, 597-629.
- Cubukcuoglu, C., Ekici, B., Tasgetiren, M. F., & Sariyildiz, S. (2019). OPTIMUS: self-adaptive differential evolution with ensemble of mutation strategies for grasshopper algorithmic modeling. *Algorithms*, 12(7), 141.
- Ekici, B., Turkcan, O. F., Turrin, M., Sariyildiz, I. S., & Tasgetiren, M. F. (2022). Optimising high-rise buildings for self-sufficiency in energy consumption and food production using artificial intelligence: Case of Europoint complex in Rotterdam. *Energies*, 15(2), 660.
- McNeel, R. (2014). Grasshopper 3D. URL: <https://www.grasshopper3d.com/>.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., & Dubourg, V. (2011). Scikit-learn: Machine learning in Python. *the Journal of machine Learning research*, 12, 2825-2830.
- Roman, N. D., Bre, F., Fachinotti, V. D., & Lamberts, R. (2020). Application and characterization of metamodels based on artificial neural networks for building performance simulation: A systematic review. *Energy and Buildings*, 217, 109972.
- Roudsari, M. S., Pak, M., & Smith, A. (2013). Ladybug: a parametric environmental plugin for grasshopper to help designers create an environmentally-conscious design. Proceedings of the 13th international IBPSA conference held in Lyon, France Aug.
- Sariyıldız, S. (2012). Performative Computational Design. Keynote speech in: Proceedings of ICONARCH-I: International congress of architecture-I, Konya, Turkey, 15-17 November 2012.
- UN. (2022). *The Sustainable Development Goals Report 2022* (<https://unstats.un.org/sdgs/report/2022/The-Sustainable-Development-Goals-Report-2022.pdf>).

- Westermann, P., & Evins, R. (2019). Surrogate modelling for sustainable building design–A review. *Energy and Buildings*, 198, 170-186.
- Wortmann, T. (2017). Opossum-introducing and evaluating a model-based optimization tool for grasshopper. *Proceedings of the CAADRIA*, vol 17, 283-292.