



Trafik Mikro-Simülasyon Model Kalibrasyonu için Özellik Seçim Algoritmalarının Karşılaştırılması

Comparison of Feature Selection Algorithms for Traffic Micro-Simulation Model Calibration

Erdem Doğan¹

¹Kırıkkale Üniversitesi Yerleşkesi, Mühendislik Fak. İnşaat Müh. Böl., Yahşihan/71450, Kırıkkale, Türkiye

Başvuru/Received: 30/03/2022

Kabul / Accepted: 29/07/2022

Çevrimiçi Basım / Published Online: 31/07/2022

Son Versiyon/Final Version: 31/07/2022

Öz

Günümüzde trafik simülasyonlarını gerçekleştirmek için önemli sayıda gelişmiş mikro simülasyon modeli geliştirilmiştir, ancak bu modeller tüm trafik koşullarını modellemek için kalibre edilmesi gereken çok sayıda parametre içermektedir. Tüm bu parametreleri kalibre etmeye çalışmak maliyetli olabilir ve hatta kalibrasyon doğruluğunu azaltabilir. Bu çalışmada, çeşitli Özellik Seçim Algoritmalarının (ÖSA) kalibrasyon doğruluğu üzerindeki etkilerinin bir analizi yapılmış ve uygun ÖSA tipinin belirlenmesi için bir yaklaşım önerilmiştir. Önerilen yaklaşım kapsamında, SUMO'nun araç tipi, araç takip ve şerit değiştirme model parametrelerinden bir model parametre seti oluşturulmuş ve Latin Hiper Küpü kullanılarak deney seti oluşturulmuştur. Deneyler, yüksek zaman çözünürlüğünde veri toplama yeteneğine sahip detektörlerle donatılmış 9,2 km uzunluğundaki bir karayolu kesimi için gerçekleştirilmiştir. Sonuç olarak, ÖSA kullanımının kalibrasyon performansını önemli ölçüde iyileştirebileceği gözlemlenmiştir. Ayrıca bu çalışmada önerilen kalibrasyon yönteminin trafik simülasyonu uygulayıcıları ve araştırmacılar için fonksiyonel olacaktır.

Anahtar Kelimeler

“Trafik, kalibrasyon, simülasyon, özellik seçimi”

Abstract

A significant number of advanced microsimulation models have now been developed to perform traffic simulations, but these models contain a large number of parameters that must be calibrated to model all traffic conditions. Attempting to calibrate all of these parameters can be costly and even reduce calibration accuracy. In this study, an analysis of the effects of various Feature Selection Algorithms (FSA) on calibration accuracy is conducted and an approach is proposed to determine the appropriate FSA type. As part of the proposed approach, a model parameter set was created from SUMO's vehicle type, car following and lane change model parameters, and an experimental set was created utilizing the Latin Hyper Cube. The experiments were carried out for a 9.2 km long road section equipped with detectors capable of collecting high time resolution data. As a result, it was observed that using FSA can significantly improve the calibration performance. In addition, the calibration method proposed in this study can be functional for traffic simulation practitioners and researchers.

Key Words

“Traffic, calibration, simulation, feature selection”

1. Giriş

Trafik simülasyon modelleri, bir tasarımın değerlendirilmesi, sürüş davranışının incelenmesi, taşıtlar arasındaki etkileşimlerin ve yönetim alternatiflerinin araştırılması için uygun maliyetli ve esnek bir yaklaşım sağlamaktadır (Balakrishna et al., 2007; Rakha et al., 1996). Diğer taraftan, belirli düzeyde faydanın elde edilebilmesi için model sonuçları güvenilir olmalıdır. Simülasyonun güvenilirliği, model parametrelerinin incelenen bölgedeki hâkim koşullara uyarlanmasıyla, yani modelin kalibrasyonu ile mümkündür.

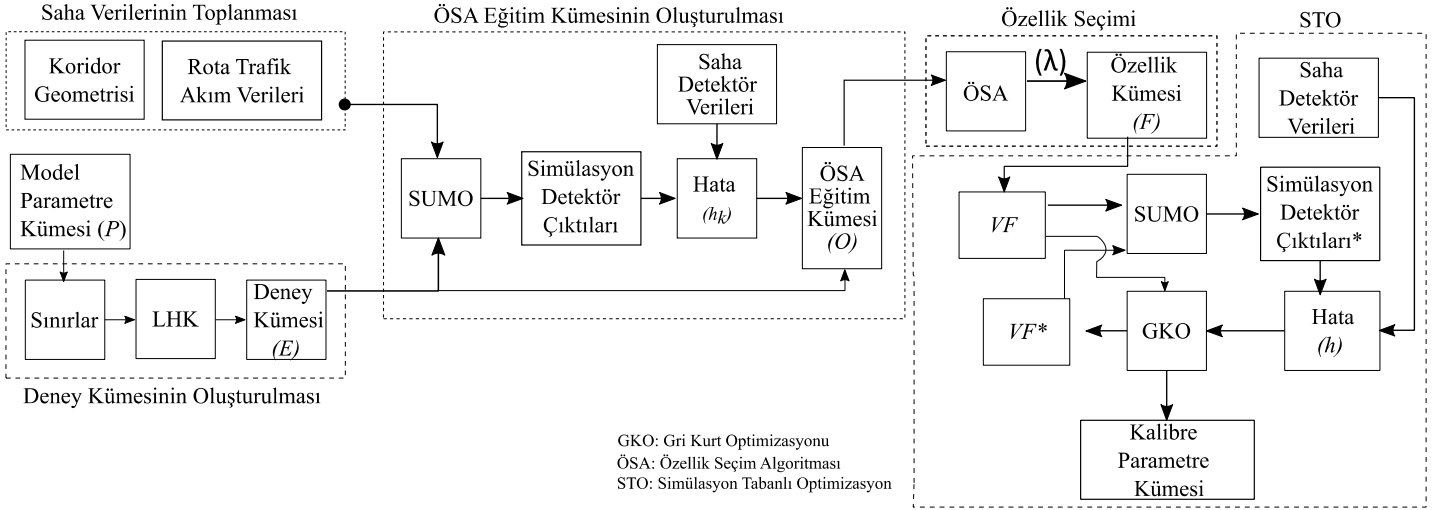
Bir trafik modelinin kalibrasyonu, belirlenen trafik değişkeni için yerel saha ölçümü ile simülasyon değerleri arasındaki farkı kabul edilebilir bir düzeye indirmek için model parametrelerin ayarlanması prosedürüdür. Bir mikro-simülasyon model kalibrasyonu genellikle Simülasyon Tabanlı Optimizasyon (STO) problemi olarak ele alınır, yani amaç fonksiyonun değerlendirilmesi için bir simülasyon modeline ihtiyaç duyulur. Bu nedenle gerekli işlem gücü ve süresi yüksek bir yaklaşımdır, fakat bu durum uygun optimizasyon yönteminin seçilmesiyle hafifletilebilir. Diğer taraftan bir bölgenin detaylı modellenmesi için kullanılan mikro-simülasyon modelleri karmaşık taşıt davranışlarını modelleyebilmek için önemli sayıda parametre barındırmakta ve parametre sayısının artması kalibrasyon işlem süresini katlanarak artırma neden olmaktadır. Sonuç olarak, en son teknoloji sezgisel algoritmalar bile bu geniş arama uzayı nedeniyle uygun çözümü bulmakta zorlanır (Ciuffo et al., 2014). Bu nedenle, ele alınan trafik değişkenine uygun bir alt parametre kümesinin kalibre edilmesi, model güvenliğini artırmak ve aşırı işlem sürelerinden kaçınmak için kullanılabilir.

Özellik Seçim Algoritmaları (ÖSA) model geliştirme aşamasında kullanılan değişen sayısını azaltmak için kullanılan araçlardır. Bu algoritmalar, ölçülen yanıtları en iyi şekilde ifade eden bir parametre/özellik alt kümesi önerirler. Bu alt kümeler modellerin geliştirilmesinde kullanılır. Böylece modeller bir taraftan basitleştirilirken diğer taraftan tahmin performansının iyileştirilmesi sağlanır. Buradan yola çıkarak, kalibrasyon işleminde etkisi yüksek parametrelerden kullanılmasıyla, kalibrasyon etkinliği artırılabilir.

Bu çalışmada, trafik simülasyon model kalibrasyonu isabetini arttırmak ve kullanılan parametre sayısını azaltılmasını sağlayan bir metod geliştirilmesi amaçlandı. İlave olarak, farklı ÖSA yaklaşımlarının bu metodun performansına etkilerinin araştırıldı. Belirlenen amaca ulaşmak için ilk olarak, gerçek bir otoyol kesiminin modeli, açık kaynak kodlu ve ücretsiz bir simülasyon programı olması nedeniyle giderek artan kullanıcı sayısına sahip olan Simulation of Urban Mobility (SUMO) (Lopez et al., 2018) ortamında oluşturuldu. Ardından taşıt takip, şerit değiştirme ve taşıt tipini modellemek için kullanılan parametrelerden, bir parametre ana kümesi (P) belirlendi. Sonraki aşamada, Latin Hiper Küpü (LHK) kullanılarak bir deney kümesi oluşturuldu. Bu deney kümesi ÖSA'ların eğitilmesi için kullanıldı ve ÖSA'ların parametrelere atadıkları ağırlıklar yardımıyla P 'nin bir alt kümesi olan özellik kümesi (F) oluşturuldu. Son aşamada, F elemanları Gri Kurt Optimizasyonu (GKO) kullanılarak kalibre edildi. Bu süreç sonunda, farklı tip ÖSA'ların SUMO için önerdikleri özellik kümeleri ve bu kümelerin kalibrasyon performansına etkileri analiz edildi.

ÖSA'ları kalibrasyon sürecine dâhil eden farklı alanlarda çalışmalar (Kimlik doğrulaması (Vigneau & Thomas, 2012), kızılötesi spektroskopisi analizi (Li et al., 2009), vb.) olmakla birlikte, bir ÖSA'nın trafik mikro-simülasyon kalibrasyonuna uygulayan bir çalışmaya imkânlar dâhilinde rastlanmamıştır. Diğer taraftan bazı çalışmalarda Duyarlılık Analizi (DA) tekniğinin kullanıldığı ve parametre ağırlıklarının araştırıldığı anlaşılmaktadır. Örneğin, Ge ve Menendez (2014) OTEE ismi verilen yöntemle kalibrasyon işlemini ele alan ve etkin parametreleri duyarlılık indekslerine göre sıralayan bir yöntem önerdi. Çalışmada VISSIM programı kullandılar ve bu yöntemin hesaplama karmaşıklığı yüksek mikroskobik modeller için faydalı olacağını belirttiler. Yine VISSIM ortamında çalışan Azam et al. (2019) bir elmas tip farklı düzey kavşak için simülasyon ve kalibrasyon işlemleri yaptı. Kalibre edilecek parametreleri DA kullanarak belirlendiler. Araştırmacılar parametre seçim işlemi için DA'nın yardımcı olduğu belirttiler. Ciuffo et al. (2014) varyans tabanlı DA tekniğini akıllı sürücü araç takip modeline uyguladılar. Farklı parametre kombinasyonlarını deneyen araştırmacılar DA'nın etkin olmayan parametrelerin belirlenmesinde etkin bir rol oynadığını ve farklı takip modelleri için farklı kombinasyonların etkin olduğunu belirttiler. (Arkatkar et al., 2016) VISSIM kullanarak heterojen trafiğe sahip bir otoyol yol için mikro-simülasyon kalibrasyonu yaptılar, fakat bu analiz sadece etki düzeyini anlayabilmek için kullanıldı. Essa & Sayed (2015) ise DA kullanarak simülasyon sonuçlarına etkisi yüksek VISSIM parametrelerini belirledi ve bu parametreleri genetik algoritmayla optimize ettiler. Bu geçmiş çalışmaların çoğunluğunda DA metodu kullanıldığı ve bazılarında ise DA'nın kalibrasyon sürecine dâhil edilmediği anlaşılıyor. Bu çalışma da ise DA'nın yerine ÖSA'lar kalibrasyon sürecine dâhil edildi, ilave olarak geçmiş kalibrasyon çalışmalarında bu alanda kullanılmayan SUMO trafik modeli kullanıldı. Hem yeni ÖSA'ların kalibrasyon sürecine dâhil edilmesi ve etkinliklerinin belirlenmesi hem de SUMO'nun bu çalışmada kullanılması, bu çalışmanın kalibrasyon konusundaki sürece katkı verme potansiyeli olduğunu göstermektedir.

İlerleyen bölümlerde öncelikle önerilen yöntem bileşenleri olan Latin Hiper Küpü, ele alınan ÖSA'lar, GKO ve diğer yöntem bileşenleri hakkında bilgiler verildi. Ardından, deneylerin yapıldığı yol kesimi ve kullanılan trafik akış değerleri hakkında bilgiler sunuldu. Sonuç bölümünde ise önerilen ÖSA'ların kalibrasyon performansına etkileri detaylı olarak tartışıldı.



Şekil 1. Önerilen kalibrasyon yaklaşımının genel akış diyagramı.

2. Metodoloji

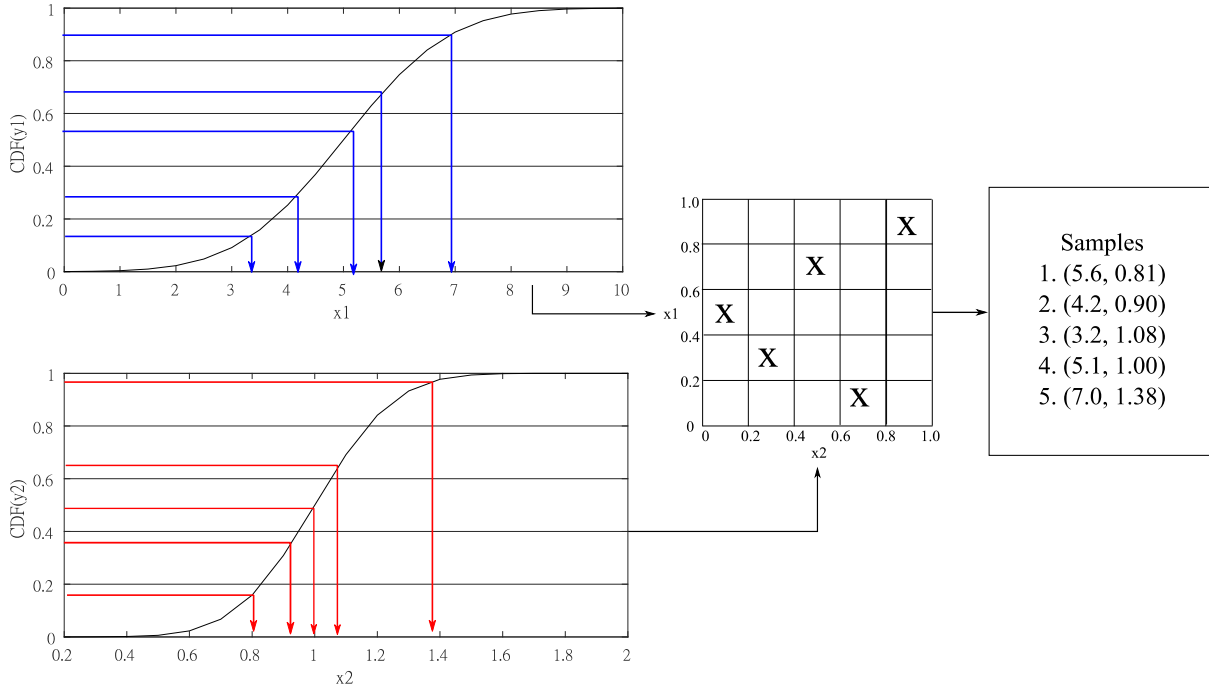
2.1. Genel Bakış

Bu bölümde önerilen trafik model kalibrasyon yaklaşımının Şekil 1’ de şematik olarak resmedilen beş ana aşaması hakkında detaylı bilgiler verildi. (i) Saha verilerinin toplanması: Bu aşamada modelin oluşturulması için gerekli veriler ölçüm veya sayım yöntemleriyle elde edilir. (ii) Deney kümesinin (E) oluşturulması: Bu küme kalibre edilecek parametrelere ait değerlerin kombinasyonudur ve Latin-Hiper Küpü (LHK) kullanılarak oluşturulur. (iii) ÖSA eğitim kümesinin (O) oluşturulması: SUMO ve E kümesi kullanılarak oluşturulmuş trafik modelinden elde edilen simülasyon detektör çıktıları ile gerçek çıktılar arasındaki farklarla hata kümesi tespit edilir. Hata kümesi ve deney kümesi birleştirilmesiyle sonucunda O kümesi elde edilir. (iv) ÖSA’nın O kümesiyle eğitilmesi ve parametreler için ağırlık değerlerinin hesaplanması ve eşik değeri (λ) yardımıyla özellik kümesinin (F) tespit edilmesi: F , P kümesinin alt kümesi olup, ÖSA tarafından seçilen parametreleri içeren bir kümedir. (v) özellik kümesi elemanlarının, yani seçilmiş parametrelerin, Simülasyon Tabanlı Optimizasyon (STO) ile kalibre edilmesi: Diğer bir deyişle, F kümesi içindeki model parametrelerin değerleri bu aşamada kalibre edilir. Bu kısımda SUMO ve Gri Kurt Optimizasyonu (GKO) birlikte çalışarak trafik akım değerleri arasındaki farkı minimize edecek parametre değerlerini araştırır ve bir Kalibre Parametre Kümesi önerir.

2.2. Latin Hiperküpe ve Deney Veri Kümesi

Bir olayın başarılı biçimde modellenebilmesi için, olayla ilgili parametre değerleri ve bu değerlere karşılık oluşan çıktı değerlerini içeren yeterli sayıda elemana sahip bir eğitim kümesi gereklidir. Fakat parametre sayısının fazla ve parametrelerin alacağı değer sınırlarının geniş olması durumlarında önemli sayıda deney yapma gerekliliği oluşur. İlk olarak McKay et al. (1979) tarafından önerilen LHK bir girdinin başlangıç dağılımına yakın bir dağılımı daha az sayıda numune ile ederek kaliteli numuneler içeren bir deney kümesi oluşturmayı amaçlar. Böylece gerekli deney sayısı uygulanabilir sayılara indirgenbilir.

LHK’nın çalışmada prensibi girdilerin kümülatif olasılık dağılımını (CDF) nE adet eşit aralıklara bölmesi (nE aynı zamanda deney kümesi eleman sayısı olmaktadır) ve her bir aralıktan rastgele bir örnek toplamasıdır. Şekil 2’de iki değişkenli (x_1 ve x_2) bir problemde öncelikle her değişkenin, $nE = 5$ olması için, CDF’leri 0.2 aralıklarla bölünmüş ve her aralık içinden rastgele CDF değerleri ve karşılık gelen değişken değerleri belirlenmiştir. Örneğin x_1 değişkenine ait CDF aralığı 0.8-1.0’de, x_1 değişkenin değeri 7.0’dır. Bir istatistiksel örnek toplamanın Latin Hiperküpe olması için her bir kolon ve satırdan sadece bir adet örnek alınmalıdır. Bu şartın sağlandığı Şekil 2’deki (x_1 , x_2) ızgara tablosunda görülmektedir. LHK’nın son aşamasında bu düzene göre 5 eleman içeren deney kümesini oluşturulur.



Şekil 2. Latin Hiper küpü kullanarak örnek bir deney kümesinin oluşturulması.

Simülasyon modeli kapasitene bağlı olarak, taşıt davranışlarının modellenmesinde kullandığı önemli sayıda parametre içerir. Amaç fonksiyonu ile ilgisi yüksek parametrelerin kalibrasyonun yapılması sürece katkı sağlayacaktır. Örneğin, çoğu simülasyon programı kavşaklardaki taşıt davranışlarını modellemek için çeşitli parametreler barındırır fakat bir kesintisiz akım koşullarının geçerli olduğu bir otoyol kesiminin kalibrasyonuna bu parametrelerin dâhil edilmesi süreci zorlaştıracaktır. Bu nedenle parametre seçim işlemine geçilmeden önce kaliteli bir deney kümesinin oluşturulması sürecin performansını arttıracaktır.

Şekil 2’de görülen deney kümesinin (E) oluşturulma süreci, simülasyon programının içerdiği parametreler arasından elemanları seçilen ve $P = \{p_i: i \leq np \wedge i \in \mathbb{N}\}$ olarak ifade edilen Model Parametre Kümesi belirlenmesiyle başlar. Ardından, p_i parametresinin değeri vp_i ile gösterilecek olursa, alt $L = \{l_i: l_i \leq vp_i\}$ ve üst $U = \{u_i: u_i \geq vp_i\}$ sınır kümeleri tespit edilir. Bu sınır değerleri taşıt dinamikleri ile ilgili fizik kurallarına bağlı uç değerler göz önüne alınarak belirlenebilir. Sürücü davranışlarıyla ilgili sınırların belirlenmesi ise daha karmaşık bir süreç olup, bazı saha gözlemleri veya modelin varsayılan sınırları kullanılabilir. LHK’nın uygulanmasının ardından $E = \{VP_k: vp_{i(k)} \in VP_k, (k \leq nE), (k \in \mathbb{N})\}$ olarak ifade edilen deney kümesi elde edilir. Yukarıda LHK’nın çalışma prensibinde de anlatıldığı gibi nE , E kümesinin eleman sayısını, yani deney sayısını belirler ve işlem öncesinde kullanıcı tarafından belirlenmelidir.

2.3. ÖSA Eğitim Kümesinin Oluşturulması

Bu aşama, VP_k elemanlarının SUMO parametreleri olarak kullanılması sonucunda, simülasyon ile saha detektörleri arasındaki değer farklarının, yani simülasyon hatalarının tespit edilmesi ve ÖSA eğitim kümesinin (O) oluşturulması işlemlerini kapsar.

İncelenen yol ağı üstünde nd adet detektör bulunsun ve j detektörünün $t \leq nt$ periyodunda ölçtüğü simülasyon ve gerçek trafik akım değerleri, sırasıyla $q_{j(t)}$ ve $\hat{q}_{j(t)}$ olsun. Bu çalışmada trafik akım değerlerine (q) bağlı kalibrasyon yapılmıştır. Farklı bir değişken için bu kısmın değiştirilmesi yeterlidir. Bu durumda bu yol ağının t periyodu için hata değeri Denklem 1 ile hesaplanır.

$$h_{(t)} = \frac{1}{nd} \sum_{j=1}^{nd} (q_{j(t)} - \hat{q}_{j(t)})^2 \quad (1)$$

Denklem 1 diğer periyotlar için tamamlanıp, hata değerlerinin aritmetik olarak toplanmasıyla h_k elde edilir. Böylece, her V_k kümesine karşılık gelen $h_k \in H$ değerleri tespit edilir. Sonuç olarak ÖSA eğitim kümesi (O) Denklem 2 tarif edilen biçimde oluşturulur.

$$O = \{(VP_k, h_k): VP_k \in \mathbb{R}^{np}, h_k \in \mathbb{R}\} \quad (2)$$

2.4. Özellik Seçimi Algoritmaları ve Eşik Değeri

Özellik seçim işlemi, önemli sayıda parametre (özellik) tarafından etkilendiği düşünülen bir olayda, belirli adımlarla etkisi yüksek parametreleri belirlemeyi ve düşük ağırlıklı parametreleri elemeyi amaçlar. Böylece karmaşıklığı azaltılmış ve hatta daha yüksek performanslı modellerin geliştirilmesi sağlanır. Farklı yaklaşımlarla geliştirilmiş önemli sayıda ÖSA'lar bulunmakta ve bunlar filtre, sarıcı ve gömülü olarak üç grupta incelenmektedir. Bu çalışmada filtre ve gömülü tip ÖSA'lar kullanılmış ve özet bilgileri referanslar ile birlikte Tablo 1'de sunulmuştur.

Tablo 1. ÖSA'lara ait bilgiler

ÖSA	Özet Bilgi
F-RANK (Saeys et al., 2007)	F-RANK her bir özelliğin önemini F-testi kullanarak ayrı ayrı inceler. Öncelikle, her bir özellik ile tepki (bu çalışmada için hata) değerleri arasında doğrusal regresyon modeli kurarak korelasyon değerleri hesaplar. Ardından bu korelasyonlardan faydalanılarak F-skorları ve p-değerlerini bulur ve özelliklerin önem dereceleri elde edilen p-değerlerine göre sıralar.
Komşuluk Bileşenleri Analizi (NCA) (Yang et al., 2012)	k -en yakın komşular algoritmasının k parametresinin belirlenme problemine karşı geliştirilmiştir. Bir noktayı rastgele seçerek bir olasılık fonksiyonu içinde noktalar arası uzaklıkları kullanır ve nokta ile ilgili karar verir. Orijinalinde sınıflandırma için geliştirilmiş olsa da regresyon problemleri içinde kullanılmaktadır.
RELIFF (Kira & Rendell, 1992)	Her bir tahmin edici vektör içindeki özellikleri karşılaştırır. Öncelikle hedef bir vektör ve ardından hedef vektör tepkisine yakın ve uzak değerlere sahip iki vektör belirler. Eğer özellik değeri değiştiğinde tahminde değişkenlik oluyorsa, o özelliğin ağırlığı ödüllendirilir. Aksi durumda cezalandırılır. Böylece incelenen özelliklere ait ağırlıklar hesaplanır.
Gauss Süreci Regresyonu (GSR)(Rasmussen, 2003)	Tahmin edici ve tepki verilerine GSR uygulanır. Regresyon sonucunda elde edilen özellik ağırlıkları aynı zamanda özelliklerin önemini gösterir. Düşük ağırlıklı özellikler modelden elenir.
Lasso Regülasyonu (LASSO)(Tibshirani, 1996)	Gözlem ve model çıktısı arasındaki hatayı minimize etmek için bir amaç fonksiyonu kullanılır. LASSO bu amaç fonksiyona bir ceza terimi ekler. Ceza terimi içinde yer alan α katsayılarının büyütülmesiyle model performansına katkısı düşük olan özellik katsayılarının küçülterek modelden çıkarır.

Bu ÖSA algoritmalarının her biri O kümesiyle eğitildikten sonra, $W = \{w_i: i \leq nP, i \in \mathbb{N}\}$ biçimde ifade edilen özellik ağırlık kümesi ortaya çıkar. ÖSA'lar ele alınan parametreler için sonuç olarak bir ağırlık değeri önermektedir. Fakat algoritmaların probleme yaklaşımları farklı olduğu için önerdikleri ağırlıkların sınır değerleri birbirinden farklı olmaktadır. Bu nedenle, bu çalışmada ağırlıkların anlaşılır biçimde karşılaştırılabilmesi ve parametre seçimi yapabilmek için öncelikle bir normalizasyon işlemi uygulanmış ardından bir belirlenen eşik değerini (λ) geçenler F kümesi elemanı olarak seçilmiştir.

Normalize edilmiş ağırlık değeri, $w'_s \in W'$ Denklem 5 yardımıyla hesaplanır. Burada l' ve u' sırasıyla normalizasyon için belirlenen alt ve üst sınırlardır.

$$w'_s = l' + \frac{w_s - \min(W)}{\max(W) - \min(W)} * (u' - l') \quad (5)$$

Böylece, her bir özelliğe atanan ağırlıklar, W' kümesini oluşturur ve hangi parametrenin kalibrasyona dâhil edileceği önceden belirlenen λ kullanılarak belirlenir. Eğer $w'_s \geq \lambda$ ise, f_s kalibre edilir, aksi durumda ise çıkartılır ve $F = \{f_s: s \leq nP, s \in \mathbb{N}, nP \leq nP\}$ olarak ifade edilen Özellik Kümesi oluşturulmuş olur. Böylece ağırlığı yüksek parametreler kalibre edilmek için seçilirken, optimizasyon işlemi için karar değişkeni sayısı azaltılmış olur.

2.5. Simülasyon Tabanlı Optimizasyon

Süreç F elemanlarına başlangıç değerinin atanması ve Denklem 3 de tanımı verilen VF' nin oluşturulmasıyla başlar. Denklemde kullanılan l_s ve u_s , sırasıyla U ve L kümelerindeki ilgili parametreye ait alt ve üst sınır değerleri ifade eder.

$$VF = \{vf_s: l_s \leq vf_s \leq u_s, vf_s \in \mathbb{R}\} \quad (3)$$

Ardından, VF kullanılarak simülasyon detektörlerinin ölçtüğü trafik akım değerleri ve gerçek değerler arasındaki fark, yani hata değerleri (h) tespit edilir. GKO, VF ve h değerlerini kullanarak optimizasyon işlemi gerçekleştirir ve bir sonraki iterasyonu için bir VF^* kümesi önerilir. Bu işlem bir optimizasyon problemi olarak şu şekilde ifade edilebilir:

$$\text{Min} \quad f(VF) = \frac{1}{nd} \sum_{t=0}^{nt} \sum_{j=1}^{nd} (q_{j(t)}(VF) - \hat{q}_{j(t)})^2 \quad (4)$$

Denklem 3'de verilen VF' 'ye ait kısıtlar Denklem 4 içinde geçerlidir.

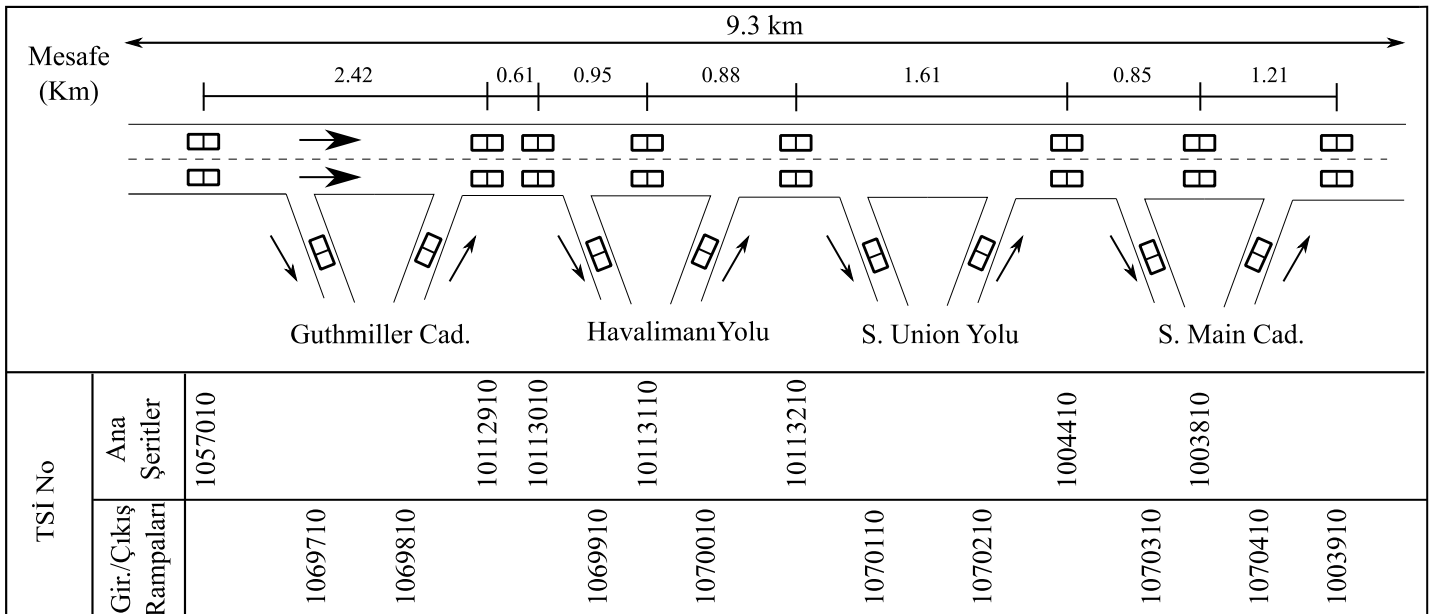
GKO algoritmasının çözüm arayışı önceden belirlenen bir hata değerine veya iterasyon sayısına ulaşıldığında sonlanır ve son iterasyonda elde edilen VF^* elemanları, kalibre parametre kümesi olarak tanımlanır.

GKO, gri kurt sürülerinin avlanma stratejilerini taklit eden bir yaklaşım ile çözüm arayan bir optimizasyon algoritmasıdır. Gri kurtlarda bulunan hiyerarşiyi dört seviye kullanarak taklit eder. Bunlar: alfa, beta, delta ve omega seviyeleridir ve alfa kurt hiyerarşinin tepesinde yer alır. GKO' da, problemin çözümü bir avın pozisyonuna benzetilir ve her bir hiyerarşinin amacı çeşitli teknikler ile av pozisyonuna yaklaşımdır. Bu algoritmaya ait detaylı bilgi orijinal çalışmadan elde edilebilir (Mirjalili et al., 2014).

3. Deneysel Düzeninin Kurulması

3.1. Trafik Veri Kümesi ve Test Otoyol Kesimi

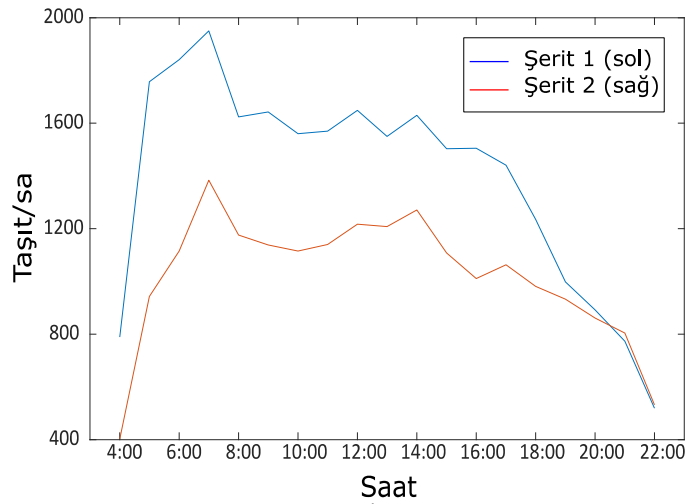
Önerilen yöntemin test edilmesi için California Eyaleti, Manteca Kenti içinde yer alan ve Şekil 3'de resmedilen otoyol bölümü (Batı-Doğu istikameti) seçilmiştir. Bu otoyol üstünde Performance Measurement System (PeMS) sistemine bağlı Taşıt Sayım İstasyonları (TSİ) bulunmakta olup 5 dk. periyotlarla bilgi akışı sunabilmektedir. Ana yol ile giriş ve çıkış rampalarında ayrı detektörler sayesinde kalibrasyon için önemli büyüklükte veri sağlayabilmesi, bu yol kesiminin analizler için seçilmesinde etkili olmuştur. Ana yollarda aynı kesit içinde her şerit için ayrı detektörler bulunmakta olup, bu iki detektör bir TSİ olarak anılmaktadır. Çalışmada kullanılan TSİ'lerin ara mesafeleri ve kod numaraları Şekil 3' de verilmektedir.



Şekil 3. Otoyol planı ve detektör numaraları

Analizlerin yapılacağı zaman kesiminin seçimi için detektörlerin sağlık durumları incelenmiş ve veri kaybının olmadığı 17 Mayıs 2019 tarihinde 04:30-08:30 saatleri arası kalibrasyon işlemleri için seçilmiştir. Böylece trafik akımının en düşük ve en yüksek olduğu (zirve

saat) saat aralığı kapsamıştır. Şekil 4’de incelenen tarihte 1013110 nolu TSİ tarafından ölçülen trafik akımı saatlik olarak verilmiştir. Zirve saatin 07:00 olduğu ve analiz periyodunun hem en düşük akım durumlarını hem de zirve saati içerdiği net olarak görülmektedir.



Şekil 4. 1013110 nolu TSİ'ye ait trafik akış grafiği

3.2. Model Parametre Kümesi ve Sınırlar

SUMO, taşıt tipi, kavşak içindeki taşıt davranışı, şerit değiştirme ve taşıt takip durumlarını modellemek için önemli sayıda parametre sunmaktadır. Bu çalışmada Tablo 2’de açıklamaları ile verilen dokuz adet parametre P elemanları olarak seçilmiştir.

Tablo 2. Model parametre kümesine ait bilgiler

Parametre	l	u	Açıklama
$speedFactor$	0.2	2	Sürücünün, izin verilen maksimum hızın ne oranda üstünde veya altında gideceğini belirleyen katsayıdır. Taşıt tipi parametreleri içinde yer alır
$speedDev$	0.1	0.8	Trafik akımı içindeki taşıt hızlarını heterojen hale getirmek için kullanılır. Taşıt tipi parametreleri içinde yer alır
$minGap$	1	10	Taşıtların birbirini izleyeceği minimum takip mesafesini ifade eder. Taşıt tipine ait bir katsayıdır.
$lcStrategic$	0	10	Belirlenen taşıt rotasını takip edebilmek için gerçekleştirilen şerit değiştirme işlemini modellemek için kullanılır. Daha yüksek değerler taşıtın daha erken şerit değiştirilmesiyle sonuçlanır. Şerit değiştirme modeline ait bir parametredir.
$lcSpeedGain$	0	10	Hız kazanmak için şerit değiştirme isteğidir. Daha yüksek değerler daha fazla şerit değiştirmeye neden olur. Şerit değiştirme modeline ait bir parametredir.
$accel$	1	3	Taşıtların hızlanma ivmesini ifade eden değer. Taşıt takip modeline ait bir parametredir.
$decel$	1	5	Taşıtların yavaşlama ivmesini ifade eden değer. Taşıt takip modeline ait bir parametredir.
σ	0	1	Sürücü kusuru (0, mükemmel sürüşü ifade eder). Taşıt takip modeline ait bir parametredir.
τ	1	10	Sürücünün takip ettiği araç ile arasındaki olmasını istediği sn. cinsinden minimum mesafedir. Taşıt takip modeline ait bir parametredir.

Tablo 2' de belirtilen parametreler SUMO programında kullanılan orijinal isimleri değiştirilmeden İngilizce olarak bırakılmıştır. Alt ve üst sınır değerleri (l, u), fizik kuralları ve olabilecek makul uç sınırlar gözetilerek belirlenmiştir. P kümesi dışında kalan parametreler için SUMO'nun varsayılan değerleri kullanılmıştır.

3.3. Algoritmalar için başlangıç ayarları

Kullanılan ÖSA algoritmaları seçim işlemlerini gerçekleştirebilmek için çeşitli parametreler kullanılmaktadır. Bu çalışmaya ait ÖSA çözümleri Matlab 2020b ortamında gerçekleştirilmiş olup, genel olarak programın önerdiği varsayılan değerler kullanılmıştır. GKO ve ÖSA parametrelerine ait değerler Tablo 3'de sunulmuştur.

Tablo 3. Model parametre kümesine ait bilgiler

Algoritma	Parametreler
F-RANK	Grup Adedi=10 Optimizasyon Algoritması= lbfgs
NCA	Yaklaşma Metodu = exact İterasyon sayısı = 1000
RELIEFF	Sigma = 50 Kernel Fonksiyonu = Karesel eksponansiyel
GSR	Fit Metodu = exact Optimizasyon Algoritması= Quasi-newton
LASSO	Lamda=100 10-kat çapraz doğrulama
GKO	İterasyon sayısı = 100 Popülasyon = 30

4. ÖSA Hatalarının Karşılaştırılması

Bu bölümde çalışmadan elde edilen iki temel sonuç sunulup tartışılmıştır. Bunlar: (i) incelenen ÖSA'ların kalibrasyon için önerdikleri simülasyon modeli parametreleri ve (ii) önerilen parametreler ile kalibre edilen simülasyonların kalibrasyon hatalarıdır.

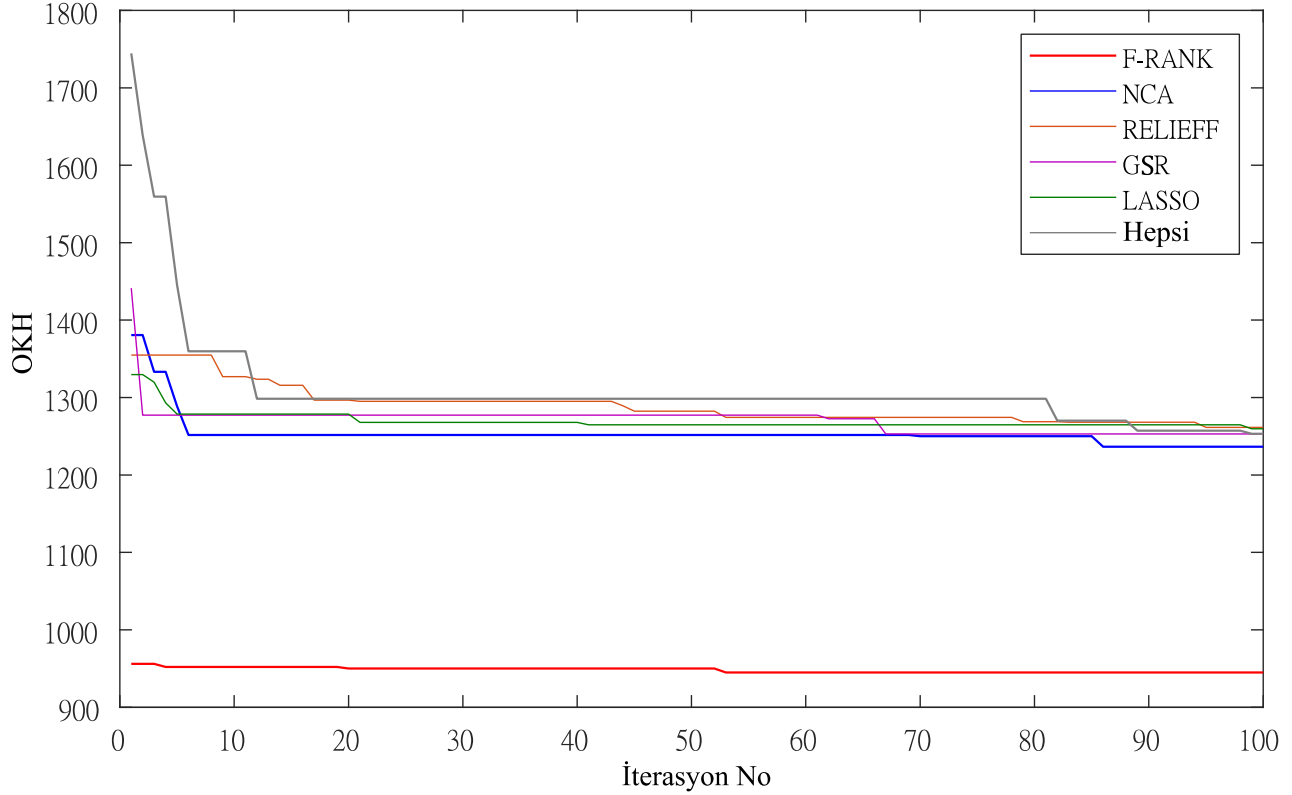
O kümesi ile eğitilen ÖSA'lar her parametre için belirli ağırlıklar tanımlamakta ve ağırlığı yüksek olan parametrelerin F kümesine seçilme şanslarını artmaktadır. Tablo 3'de, ÖSA'ların parametreler için atadıkları ağırlık değerlerinin 0-100 arasına normalize edilmiş değerleri sunulmaktadır. Normalizasyon farklı sınırlarda üretilen parametre ağırlıklarının karşılaştırılabilir hale getirmek için uygulanmış bir işlemdir.

Kalibrasyon hatasını düşürmek için en etkin parametrenin (özelliğin), *speedFactor* olduğu ve NCA dışında tüm yöntemlerin bu parametrenin ağırlığını 100 olarak tespit ettiği Tablo 3'den görülmektedir. *speedFactor*, yani hız faktörü taşıt hız dağılımları üstünde doğrudan etkiye sahip olduğu sebebiyle SUMO kalibrasyon işlemlerinde ilk ele alınması gereken parametre olarak öne çıkmaktadır. Diğer parametrelerin ağırlıkları incelendiğinde, *speedDev*'in (hız dağılımı) ve τ parametrelerinin genel olarak öne çıktığı anlaşılmaktadır. NCA *speedDev* için ağırlığı 100 olarak belirlediği fakat diğer ÖSA'ların önemli derecede düşük olduğu, hatta F-RANK'ın 1.93 e kadar ağırlığı düşürdüğü görülmektedir. Diğer taraftan τ parametresine önerilen ağırlıkların *speedDev*'e kıyasla daha uniform dağıldığı anlaşılmaktadır. Kalan parametreler incelendiğinde en yüksek ağırlığın, NCA tarafından *accel*'e atandığı görülmekte ve diğerlerinin dikkate değer ağırlıklara sahip olmadığı belirlenmiştir. Tablo 4 sonuçları genel olarak yorumlandığında, NCA dışındaki algoritmaların yaklaşık benzer eğilimler gösterdiği, hız faktörünün, hız dağılımının ve τ 'nun öne çıkan parametreler olduğu sonuçlarına ulaşılmıştır.

Tablo 4. ÖSA'ların önerdikleri parametre ağırlıkları

ÖSA	<i>speedFactor</i>	<i>speedDev</i>	<i>minGap</i>	<i>lcStrategic</i>	<i>lcSpeedGain</i>	<i>minGap</i>	<i>accel</i>	<i>decel</i>	σ	τ
F-RANK	100.00	1.93	0.16	2.11	0.76	0.16	0.33	0.09	0.00	22.42
NCA	88.36	100.00	6.12	2.16	7.17	6.12	23.13	0.00	35.70	24.50
RELIEFF	100.00	11.79	2.25	3.98	0.18	2.25	5.79	0.00	2.31	38.32
GSR	100.00	37.84	0.00	0.00	0.00	0.00	0.98	0.00	0.03	4.63
LASSO	100.00	27.78	1.99	0.00	0.00	1.99	0.00	1.08	0.00	16.12

Farklı ÖSA'ların önerdikleri F kümesinin eleman çeşitliliği sağlanması ve aynı zamanda algoritmaların belirli düzeyde önemli gördüğü parametrelerin analiz edilmesi amaçlanarak $\lambda=5$ olarak belirlenmiştir. Bu değerin belirlenmesi normalizasyon sonucunda oluşan Tablo 4'ün incelenmesi sonucunda bu çalışmada için belirlenmiş bir değerdir. Bu ÖSA'ların belirledikleri ağırlık değerlerinin $\lambda=5$ eşliğine göre filtrelenmesi sonucunda seçilen parametreler Tablo 4'de gri renk ile boyanarak gösterilmiştir. Örneğin, F-RANK için $F = \{speedFactor, \tau\}$ olarak tespit edilmiştir. En fazla parametre öneren algoritma NCA olmuş ve *lcStrategic* haricindeki bütün parametreleri F kümesine dâhil etmiştir.



Şekil 5 Optimizasyon sürecinde ÖSA hata değerleri

Şekil 5'de ÖSA'lar tarafından önerilen parametrelerin kalibre edilmesi sürecinde, GKO iterasyonlarına bağlı OKH'ların değişimi resmedilmiştir. Optimizasyon süreci sonunda oluşan hatalar incelendiğinde, en düşük hatanın F-RANK algoritmasının önerdiği parametrelerin kalibre edilmesiyle elde edildiği ve diğer algoritma hatalarının F-RANK'tan önemli miktarda fazla olduğu görülmüştür. Optimizasyonun başlangıcında en yüksek hatanın ise bütün parametrelerin, yani P kümesinin, (Şekil 5 Gri çizgi) kalibre edilmesi sürecinde oluştuğu anlaşılmıştır. F-RANK dışındaki ÖSA'lar incelendiğinde önemli bir farkın oluşmadığı sadece NCA hata değerinin bir miktar pozitif ayrıştığı gözlenmiştir.

Bu sonuçlara göre, simülasyon modelinde yer alan tüm parametrelerin kalibre edilmesinin hem işlem karmaşıklığı doğurduğu hem de kalibrasyon performansını düşürdüğü anlaşılmaktadır. Diğer taraftan OHK en düşük olan F-RANK'ın kalibre edilecek parametre sayısını önemli miktarda düşürdüğü böylece işlem karmaşıklığını azalttığı, dahası OHK hatalarda önemli düzeyde iyileşme yarattığı anlaşılmaktadır.

5. Sonuç

Trafik simülasyon modelleri, büyük ve detaylı yol ağlarını düşük maliyetle detaylı biçimde analiz edilmesini sağlayan önemli trafik mühendisliği araçlarıdır. Fakat incelenen bölge için kalibre edilmemeleri halinde, ürettikleri sonuçlar hatalı kararların alınmasına sebep olacaktır. Diğer taraftan, simülasyon modellerinin sağladığı çok sayıda parametre bulunmakta ve bunların tümünü kalibre etmek kalibrasyon sürecini karmaşık hale getirmektedir. Bu çalışmada, kalibre edilecek SUMO model parametre sayısını azaltmak için farklı özellik seçim algoritmaları kullanılmıştır. Bu algoritmalar LHK ile oluşturulmuş veri kümesiyle eğitilmiş ve parametre ağırlıkları rasyonel bir yaklaşımla tespit edilmiştir. Son aşamada, detaylı trafik verisine sahip bir yol koridoru için kalibrasyon işlemleri GKO kullanılarak gerçekleştirilmiştir. Kalibrasyon için seçilen zaman periyodunun düşük ve yüksek trafik akım değerlerini içermesine de dikkat edilmiştir. Analizler sonucunda, F-RANK özellik seçim algoritmasının önerdiği hız faktörü ve τ parametreleriyle SUMO'nun eğitilmesinin kalibrasyon performansını önemli düzeyde artırdığı kanıtlanmıştır. Diğer bir ifadeyle, kalibrasyon işlemi öncesinde ÖSA

kullanılması simülasyon doğruluğunu önemli düzeyde arttıracak potansiyele sahip bir işlem olduğu gösterilmiştir. Önerilen model parametreleri SUMO için kısıtlı olsa da diğer simülasyon programlarının kalibrasyonu için bu yaklaşımın kullanılması yararlı olacaktır. İleri çalışmalarda, diğer önemli simülasyon programları için benzer deneyler yapılabilir, farklı tip yol kesimleri için (Kavşaklar, şehir içi yollar, vb.) kalibrasyon çalışmaları yapılabilir. Böylece, farklı bölgelerde çalışan uygulamacı ve araştırmacılar için kalibrasyon sürecinin kısaltılmasına katkı sağlanabilir.

Referanslar

- Arkatkar, S., Velmurugan, S., Puvvala, R., Ponnu, B., & Narula, S. (2016). Methodology for simulating heterogeneous traffic on expressways in developing countries: A case study in India. *Transportation Letters*, 8(2), 61–76. <https://doi.org/10.1179/1942787515Y.0000000008>
- Azam, M., Puan, O. C., Hassan, S. A., & Mashros, N. (2019). Calibration of microsimulation model for tight urban diamond interchange under heterogeneous traffic. *IOP Conference Series: Materials Science and Engineering*, 527(1). <https://doi.org/10.1088/1757-899X/527/1/012077>
- Balakrishna, R., Antoniou, C., Ben-Akiva, M., Koutsopoulos, H. N., & Wen, Y. (2007). Calibration of microscopic traffic simulation models: Methods and application. *Transportation Research Record*, 1999(1), 198–207. <https://doi.org/10.3141/1999-21>
- Ciuffo, B., Punzo, V., & Montanino, M. (2014). Global sensitivity analysis techniques to simplify the calibration of traffic simulation models. Methodology and application to the IDM car-following model. *IET Intelligent Transport Systems*, 8(5), 479–489.
- Essa, M., & Sayed, T. (2015). Simulated traffic conflicts: Do they accurately represent field-measured conflicts? In *Transportation Research Record* (Vol. 2514, pp. 48–57). <https://doi.org/10.3141/2514-06>
- Ge, Q., & Menendez, M. (2014). An efficient sensitivity analysis approach for computationally expensive microscopic traffic simulation models. *International Journal of Transportation*, 2(2), 49–64.
- Kira, K., & Rendell, L. A. (1992). A practical approach to feature selection. In *Machine learning proceedings 1992* (pp. 249–256). Elsevier.
- Li, G.-Z., Meng, H.-H., Yang, M. Q., & Yang, J. Y. (2009). Combining support vector regression with feature selection for multivariate calibration. *Neural Computing and Applications*, 18(7), 813–820.
- Lopez, P. A., Behrisch, M., Bieker-Walz, L., Erdmann, J., Flötteröd, Y.-P., Hilbrich, R., Lücken, L., Rummel, J., Wagner, P., & Wießner, E. (2018). Microscopic traffic simulation using sumo. *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, 2575–2582.
- McKay, M. D., Beckman, R. J., & Conover, W. J. (1979). A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. *Technometrics*, 21(2), 239–245. <https://doi.org/10.2307/1268522>
- Mirjalili, S., Mirjalili, S. M., & Lewis, A. (2014). Grey Wolf Optimizer. *Advances in Engineering Software*, 69, 46–61. <https://doi.org/https://doi.org/10.1016/j.advengsoft.2013.12.007>
- Rakha, H., Hellinga, B., Van Aerde, M., Perez, W., Aerde, M. Van, Perez, W., Van Aerde, M., & Perez, W. (1996). Systematic verification, validation and calibration of traffic simulation models. *75th Annual Meeting of the Transportation Research Board, Washington, DC*.
- Rasmussen, C. E. (2003). Gaussian processes in machine learning. *Summer School on Machine Learning*, 63–71.
- Saeyns, Y., Inza, I., & Larrañaga, P. (2007). A review of feature selection techniques in bioinformatics. *Bioinformatics*, 23(19), 2507–2517. <https://doi.org/10.1093/bioinformatics/btm344>
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 267–288.
- Vigneau, E., & Thomas, F. (2012). Model calibration and feature selection for orange juice authentication by 1H NMR spectroscopy. *Chemometrics and Intelligent Laboratory Systems*, 117, 22–30.
- Yang, W., Wang, K., & Zuo, W. (2012). Neighborhood component feature selection for high-dimensional data. *Journal of Computers*, 7(1), 162–168. <https://doi.org/10.4304/jcp.7.1.161-168>