

**BAŐKENT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

**GAZ SENSÖRLERİNDE ÇOK DEĐİŐKENLİ
KALİBRASYON YÖNTEMLERİNİN İNCELENMESİ**

GÜLNUR BEGÜM ERGÜN

YÜKSEK LİSANS TEZİ

2018

**GAZ SENSÖRLERİNDE ÇOK DEĞİŞKENLİ
KALİBRASYON YÖNTEMLERİNİN İNCELENMESİ**

**INVESTIGATION OF THE MULTIVARIATE CALIBRATION
METHODS FOR GAS SENSORS**

GÜLNUR BEGÜM ERGÜN

Başkent Üniversitesi
Lisansüstü Eğitim Öğretim ve Sınav Yönetmeliğinin
ELEKTRİK-ELEKTRONİK Mühendisliği Anabilim Dalı İçin Öngördüğü
YÜKSEK LİSANS TEZİ olarak hazırlanmıştır.

2018

“Gaz sensörlerinde çok deęişkenli kalibrasyon yöntemlerinin incelenmesi” başlıklı bu çalışma, jürimiz tarafından, 17 / 08 / 2018 tarihinde, **ELEKTRİK-ELEKTRONİK MÜHENDİSLİĐİ ANABİLİM DALI 'nda YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Başkan : Doç. Dr. Mustafa DOĐAN

Üye (Danışman) : Dr. Öğr. Üyesi Selda GÜNEY

Üye : Dr. Öğr. Üyesi Derya YILMAZ

ONAY

.. / 08 / 2018

Prof. Dr. Faruk ELALDI
Fen Bilimleri Enstitüsü Müdürü



BAŞKENT ÜNİVERSİTESİ FEN BİLİMLERİ ENSTİTÜSÜ
YÜKSEK LİSANS TEZ ÇALIŞMASI ORJİNALLİK RAPORU

Tarih: 20 / 08 / 2018

Öğrencinin Adı, Soyadı : Gülnur Begüm ERGÜN

Öğrencinin Numarası : 21610309

Anabilim Dalı : Elektrik-Elektronik Mühendisliği

Programı : Yüksek Lisans Programı

Danışmanın Unvanı/Adı, Soyadı : Dr. Öğr. Üyesi Selda GÜNEY

Tez Başlığı : Gaz sensörlerinde çok değişkenli kalibrasyon yöntemlerinin incelenmesi

Yukarıda başlığı belirtilen Yüksek Lisans tez çalışmamın; Giriş, Ana Bölümler ve Sonuç Bölümünden oluşan, toplam 27 sayfalık kısmına ilişkin, 20 / 08 / 2018 tarihinde şahsım/tez danışmanım tarafından Turnitin adlı intihal tespit programından aşağıda belirtilen filtrelemeler uygulanarak alınmış olan orijinallik raporuna göre, tezimin benzerlik oranı %7'dir.

Uygulanan filtrelemeler:

1. Kaynakça hariç
2. Alıntılar hariç
3. Beş (5) kelimedenden daha az örtüşme içeren metin kısımları hariç

“Başkent Üniversitesi Enstitüleri Tez Çalışması Orijinallik Raporu Alınması ve Kullanılması Usul ve Esaslarını” inceledim ve bu uygulama esaslarında belirtilen azami benzerlik oranlarına tez çalışmamın herhangi bir intihal içermediğini; aksinin tespit edileceği muhtemel durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.

Öğrenci İmzası:.....

Onay

20 / 08 / 2018

Öğrenci Danışmanı Unvan, Ad, Soyad,

TEŐEKKÜR

Sayın tez danıřmanım Dr. Öğr. Üyesi Selda GÜNEY...

Sayın jüri üyelerim Doç. Dr. Mustafa DOĞAN ile Dr. Öğr. Üyesi Derya YILMAZ...

Sevgili eřim Gürkan ERGÜN...

Biricik ailem...

Her ne kadar kapakta tek bir isim yazıyor olsa da, bu alıřmanın arkasında böyle bir dev kadro var ve alıřmalarım sırasında hiçbir desteęini esirgemeyen bu dev kadroya sonsuz teőekkürlerimi iletiyorum. Tüm başarılar sizlerle anlamlı.



ÖZ

GAZ SENSÖRLERİNDE ÇOK DEĞİŞKENLİ KALİBRASYON YÖNTEMLERİNİN İNCELENMESİ

Gülnur Begüm ERGÜN

Başkent Üniversitesi Fen Bilimleri Enstitüsü

Elektrik-Elektronik Mühendisliği Anabilim Dalı

Gaz sensörlerinin uzun süreli kullanıldığı elektronik burun (e-nose) uygulamalarında, sensörler üzerinde istenmeyen bir sapma (drift) etkisi meydana gelir. Bu etki verilerin analizinde çok etkin olup ölçüm sonuçlarının doğru analizini de engellemektedir. Gaz sensör dizilerinden oluşan farklı cihazlar arasında da benzer sapma sorunları meydana gelmektedir. Bu etkilerin azaltılması, yani gelecek verilerin daha iyi tahmin edilmesi, amacı ile kalibrasyon yöntemi geliştirilerek sensörden gelecek yeni verilere karşı iyi bir tahmin performansı elde etmek mümkün olabilir.

Bu çalışma kapsamında, çevrimiçi olarak erişilebilen öznitelikleri çıkarılmış hazır bir veri seti üzerinde çok değişkenli kalibrasyon yöntemlerinden olan doğrudan standartlaştırma (direct standardization-DS), ortogonal sinyal düzeltimi (orthogonal signal correction-OSC) ile parçalı doğrudan standartlaştırma (piecewise direct standardization-PDS) yöntemleri uygulanmıştır. Ayrıca yine bu yöntemler; aynı veri setinin öznitelik seçimi yapılmış haline uygulanmış ve sonuçların tümü karşılaştırılarak en başarılı yöntem, sınıflandırma başarı oranlarına bakılarak bulunmaya çalışılmıştır.

Sonuçlara göre; öznitelik seçimi yapılan veri setine uygulanan kalibrasyon yöntemleri, ham veri setine oranla daha başarılı olmuştur. Böylece daha az veri ile kalibrasyonun başarılı bir şekilde gerçekleşebileceği gözlemlenmiştir. Bu çıktının gelecek çalışmalarda yeni bir perspektif yaratacağı düşünülmektedir.

ANAHTAR SÖZCÜKLER: Gaz sensörleri, kalibrasyon transferi; öznitelik seçimi; sapma düzeltme; standartlaştırma yöntemleri.

Danışman: Dr. Öğr. Üyesi Selda GÜNEY, Başkent Üniversitesi, Elektrik-Elektronik Mühendisliği Bölümü.

ABSTRACT

INVESTIGATION OF THE MULTIVARIATE CALIBRATION METHODS FOR GAS SENSORS

Gülnur Begüm ERGÜN

Baskent University Institute of Science and Engineering
Department of Electrical and Electronics Engineering.

In many electronic nose applications where gas sensors utilizing for a long time, there is an undesirable drift effect on the sensors, which affects the classification quality negatively. Although the sensor drift is inevitable, it is possible to reduce this effect with the calibration transfer methods. This paper presents a comparison study of various multivariate standardization methods to facilitate an effective calibration way on a comprehensive dataset, which is reachable on-line. In this study, three methods applied: direct standardization (DS), orthogonal signal correction (OSC) and piecewise direct standardization (PDS). In addition, these three methods are applied the data, which consisted of selected features.

The results have shown that the classification success has increased with multivariate calibration technique applied to the selected features. The results also demonstrate that using the best features in the signal processing part can play an important role for the calibration success. This outcome may lead to a new perspective for the future works.

KEYWORDS: calibration transfer; drift correction; feature selection; gas sensors; multivariate drift correction; standardization methods

Advisor: Dr. Selda GÜNEY, Baskent University, Department of Electrical and Electronics Engineering.

İÇİNDEKİLER LİSTESİ

	<u>Sayfa</u>
ÖZ.....	i
ABSTRACT	ii
İÇİNDEKİLER LİSTESİ.....	iii
ŞEKİLLER LİSTESİ.....	iv
ÇİZELGELER LİSTESİ.....	v
SİMGELER VE KISALTMALAR LİSTESİ.....	vi
1 GİRİŞ.....	1
1.1. Gaz Sensörlerinde Sapma Etkisi.....	1
1.1.1. Kısa Süreli (Birincil) Sapma.....	4
1.1.2. Uzun Süreli (İkincil) Sapma.....	4
1.2. Gaz Sensörlerinde Kalibrasyon.....	5
2 KALİBRASYON YÖNTEMLERİ.....	7
2.1. Doğrudan Standartlaştırma (Direct Standardization).....	7
2.2. Parçalı Doğrudan Standartlaştırma (Piecewise Direct Standardization).....	7
2.3. Ortogonal Sinyal Düzeltimi (Orthogonal Signal Correction).....	9
3 YAPILAN ÇALIŞMALAR.....	10
3.1. Çalışmalarda Kullanılan Veri Seti.....	10
3.2. Veri Setine DS, PDS ve OSC Yöntemlerinin Uygulanması.....	13
3.3. Veri Setine Önerilen Yöntemin Uygulanması.....	15
3.3.1. Öznitelik seçimi.....	16
3.3.1.1. <u>Geriye doğru öznitelik seçimi (BFS)</u>	16
3.3.1.2. <u>İleriye doğru öznitelik seçimi (FFS)</u>	17
4 SONUÇ	22
KAYNAKLAR LİSTESİ.....	24
EK 1	27

ŞEKİLLER LİSTESİ

Sayfa

Şekil 1.1	MOS tipi gaz sensörlerinin gaz algılama aşamaları [6]	2
Şekil 1.2	TGS2620 kodlu gaz sensörü için basit elektrik devresi	2
Şekil 1.3	MOS tipi gaz sensörlerinin iç yapısı [7]	3
Şekil 1.4	Gaz sensörlerinde sapma etkisi (A: Olması gereken kararlı sensör tepkisi. B: Sensörlerde zaman içinde meydana gelen sapma etkili tepki) [8]	4
Şekil 2.1	PDS Yöntemi ile birincil cihazdaki herhangi bir i'ninci dalga boyunun ikincil cihazdaki i-j ile i+k arasındaki dalga boyları ile ilişkilendirilmeye çalışılması.....	8
Şekil 3.1	Bir sensörün gaz algılama aşamaları [23]	11
Şekil 3.2	Sensörün yükselen ve düşen kısımları için 3 farklı düzleme parametresi ($\alpha= 0,1$, $\alpha= 0,01$ ve $\alpha= 0,001$) için ayrı ayrı alınan en yüksek ve en düşük değerler [23]	12

ÇİZELGELER LİSTESİ

Sayfa

- Çizelge 3.1** 36 ay boyunca alınan tüm ölçümleri gösteren çizelge [23] 13
- Çizelge 3.2** Tüm veri setinin ikinci sütunda belirtilen ölçüm ay/ayları ile oluşturulan 10 adet grup bilgisini içeren çizelge 13
- Çizelge 3.3** Ham veri setinin herhangi bir yöntem uygulanmadan elde edilen başarı yüzdesi ile aynı veri setine DS, PDS ve OSC yöntemlerinin uygulanmasıyla elde edilen sınıflandırma başarı yüzdeleri.....15
- Çizelge 3.4** FFS yönteminin uygulanması sonucunda seçilen öznitelikler.....17
- Çizelge 3.5** En iyi öznitelik seçilerek oluşturulmuş veri setinin herhangi bir yöntem uygulanmadan elde edilen başarı yüzdesi ile aynı öznitelikli veri setine DS, PDS ve OSC yöntemlerinin uygulanmasıyla elde edilen sınıflandırma başarı yüzdeleri.....18
- Çizelge 3.6** Ham veri setinden oluşturulmuş 7. grup eğitim verisi ile ortaya çıkan hata matrisi.
(i.Etanol, j.Etilen, k.Amonyak, l.Asetaldehit, m.Aseton, n.Tolün).....19
- Çizelge 3.7** Öznitelik seçimi ile oluşturulmuş veri setinden 7. grup eğitim verisi ile ortaya çıkan hata matrisi.
(i.Etanol, j.Etilen, k.Amonyak, l.Asetaldehit, m.Aseton, n.Tolün)19
- Çizelge 3.8** Ham veri seti ile yapılan analiz sonucunda tüm gazların sahip olduğu özgünlük ve hassasiyet yüzde değerleri.20
- Çizelge 3.9** Öznitelik seçimi ile yapılan analiz sonucunda tüm gazların sahip olduğu özgünlük ve hassasiyet yüzde değerleri.....20

SİMGELER VE KISALTMALAR LİSTESİ

AC	Alternative current
ANN	Artificial neural network
BFS	Back feature selection
DC	Direct current
Diag.	Diagonal
DS	Direct standardization
FFS	Forward feature selection
Inc.	Incorporation
k-NN	k- nearest neighbours
MLPCA	Maximum likelihood principal component analysis
NIPALS	Nonlinear iterative partial least squares
MOS	Metal oxide semiconductor
MU	Model updating
OSC	Orthogonal signal correction
PCA	Principal component analysis
PDS	Piecewise direct standardization
PLS	Partial least squares
PMF	Positive matrix factorization
PRS	Piecewise reverse standardization
RS	Reverse standardization
SBC	Slope and bias correction
S_1	Temel cihaza ait transfer örneklerinin oluşturduğu alt küme matrisi
S_2	İkincil cihaza ait transfer örneklerinin oluşturduğu alt küme matrisi
S_2^+	İkincil cihaza ait transfer örneklerinin oluşturduğu alt küme matrisinin sözde (pseudo) tersi
F	Dönüşüm matrisi
\underline{x}	DS yönteminin uygulanması sonucu oluşan veri matrisi

X_i	PDS yönteminde seçilen pencere arasında kalan ölçümleri içeren matris
b_i	PDS yönteminde i . ölçüm için regresyon vektörü
$r_{1,i}$	PDS yönteminde temel cihaza ait i . ölçüm
$r_{2,i}$	PDS yönteminde ikincil cihaza ait i . ölçüm
t_{osc}	OSC bileşeni
w_{osc}	OSC bileşeni
p_{osc}	OSC bileşeni
T	X verisinin skor matrisi
U	Y verisinin skor matrisi
P	X verisinin yük matrisi
Q	Y verisinin yük matrisi
E	X verisinin kalıntı matrisi
F	Y verisinin kalıntı matrisi
$r [k]$	Zaman serisi
$y [k]$	$r [k]$ zaman serisinin üstel hareketli ortalaması
α	Düzleme parametresi
ΔR	Sensörlerin kalıcı durum cevabı

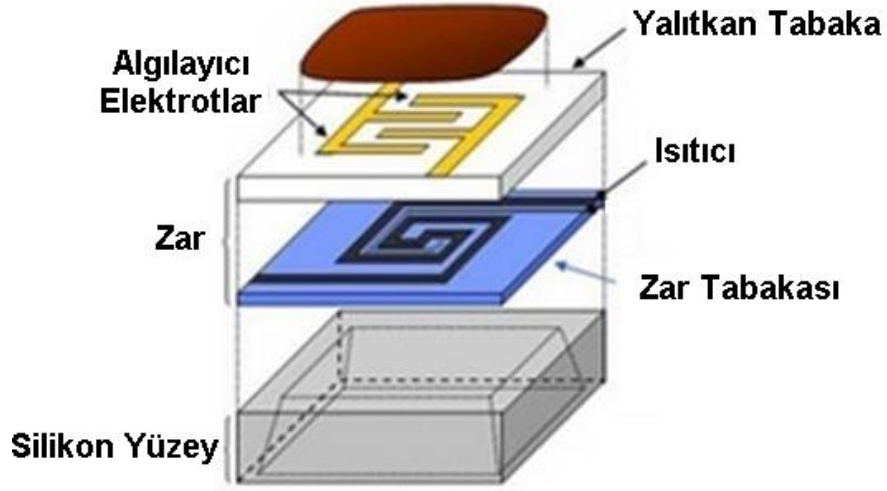
1. GİRİŞ

1.1. Gaz Sensörlerinde Sapma Etkisi

Gaz sensörleri, ortamdaki gazların yoğunluğunu ve cinsini algılayan elektro-kimyasal cihazlardır. Son yıllardaki teknolojik ilerlemeler ile kullanımı oldukça yaygınlaşan gaz sensörleri, yangın alarm sistemlerinde [1], yiyecek ve içecek kalitesinin belirlenmesinde [2,4], hasta tanısı koyma [3,4] gibi pek çok uygulama alanlarında kullanılabilirler.

Günümüzde gaz sensörleri sınırlı teknolojiye sahiptir ve her gaz sensörü her uygulamaya uygun olmayabilir. Dolayısıyla piyasada bulunan pek çok çeşit gaz sensöründen ihtiyaca en uygun olan çeşidi kullanılmalıdır. Elektro-kimyasal, kızılötesi, pelistör, kalorimetrik, metal oksit gibi çeşitleri olan gaz sensörleri içinde en yaygın olarak kullanılan çeşidi Metal oksit sensörler (MOS)'dir. MOS tipi gaz sensörler her ne kadar yavaş tepkileri, fazla hassasiyetleri ve güç kullanımının çok olması gibi dezavantajlara sahip olsa da, düşük maliyetleri sebebiyle çalışmalarda en çok tercih edilen gaz sensörlerindendir. Bu çalışmada da bu nedenle bu sensörlere ait veri tabanı tercih edilmiştir.

MOS tipi gaz sensörlerinin yapısında kalaydioksit (SnO_2) gibi yarı iletken malzemeler kullanılır. Bu sensörlerin çalışma prensipleri kabaca dört aşamada incelenebilir. İlk aşamada sensörler hazır durumda değil, yani ısıtılmamış durumda iken, sensör yüzeylerine herhangi bir gaz bağlanması, dolayısıyla da gaz algılaması söz konusu değildir. Sonraki aşamada yarı iletken kalaydioksit tabakalarının sensörlere gerilim uygulanması sonucu ısıtma bobinleri tarafından ısıtılmasıyla sensörler gaz algılamaya uygun hale getirilir. Temiz havadaki oksijen atomları, sensörlerin bu yarı iletken tabakası tarafından tutulurlar ve yalıtkan durumdaki yarı iletken malzeme, üzerinden elektrik akımı geçişine izin vermez. Ancak son aşamada ortama indirgeyici (yanıcı) bir gaz girdiği zaman, bu oksijen atomları indirgenir ve yarı iletken malzeme içinde serbest kalan elektronların hareketiyle iletken durumuna geçen malzeme üzerinden elektrik akımı akmaya başlar. Sensörlerin iletkenliği havadaki gaz konsantrasyonu ile orantılı olarak artar ve basit bir elektrik devresiyle bu iletkenlik çıkışa gerilim olarak aktarılarak ortamdaki gazların varlığı tespit edilir [5]. Bu dört aşama Şekil 1.1'de görselleştirilerek anlatılmıştır.

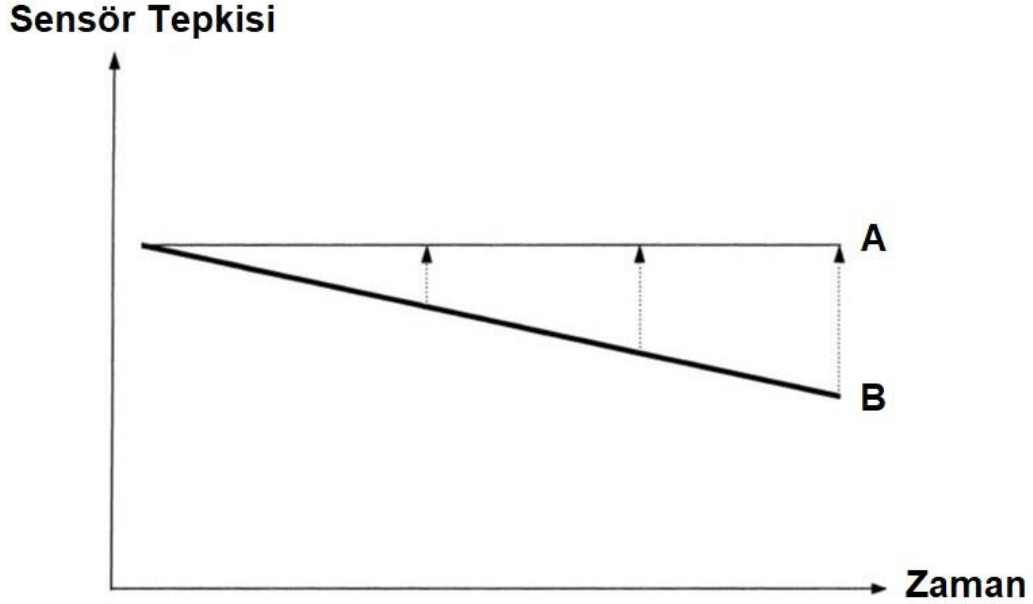


Şekil 1.3. MOS tipi gaz sensörlerinin iç yapısı [7]

Sensörlerin çalışabilmesi için ısıtılması şarttır. Bu işlem dışarıdan verilen AC ya da DC besleme sayesinde gerçekleştirilir. Isıtma işlemi süresi sensör tipine bağlı olarak değişir. Sensörlerden doğru sonuç alabilmek için bu süre üretici firmasının belirlediğinden az olmamalıdır.

Sensörlerin aktif konuma geçtikten sonra verdiği ilk tepki ile aynı şartlarda uzun süreli kullanımları sonucu verdikleri tepki aynı olmaz. Bu, sensörlerin doğasından kaynaklanan bir durumdur ve bu durum doğru analiz sonuçlarına ulaşmada problem teşkil eder. Gaz sensörlerindeki bu problem literatüre 'sapma etkisi' ya da 'drift etkisi' olarak geçmiştir.

Sensörlerdeki sapma etkisini Şekil 1.4'deki gibi gözlemleyebiliriz. Şekilde A eğrisi ile gösterilen kararlı tepki, aynı şartlarda aynı tip gaz sensörlerinde olması gereken tepkidir. Sensör aynı cins ve yoğunluktaki gaza, ilk kullanımında da son kullanımında da aynı tepkiyi vermelidir. Ancak bu durum gerçekte pek olası değildir. B ile gösterilen eğri gerçeğe daha uygun olup sensörlerin kullanımından kaynaklı olarak, zamanla kararlı durumlarının bozulmasını göstermektedir. Şekildeki sensör tepkisi her ne kadar zamanla azalan bir eğri ile ifade edilse de, artan bir eğilim de gösterebilir.



Şekil 1.4. Gaz sensörlerinde sapma etkisi (A: Olması gereken kararlı sensör tepkisi. B: Sensörlerde zaman içinde meydana gelen sapma etkili tepki) [8]

Gaz sensörlerinin tepkilerinde zamana karşı kontrol edilemeyen bu artış ya da azalış durumu, yani sapma etkisi, iki farklı şekilde incelenebilir.

1.1.1. Uzun süreli (birincil) sapma

Sensör yüzeyinde meydana gelen kimyasal tepkimeler sayesinde sensörlerin ortamdaki gazı algıladığı giriş kısmında detaylı olarak anlatılmıştı. Bilindiği üzere, tüm kimyasal tepkimeler kalıcıdır; yani geri dönüştürülebilir değildir. Bu nedenle oluşan her bir kimyasal tepkime zamanla sensör yüzeyinde kalıcı bir değişiklik yaratarak sensörlerde yaşlanma (aging) adı verilen etkiyi meydana getirir. Bundan dolayı sensörler uzun süreli kullanımlarında aynı yoğunluktaki aynı gaza ilk kullanımları gibi tepki veremezler. Bu etkiyle oluşan sapma, uzun süreli ya da birincil sapma olarak adlandırılır [9]. Uzun süreli sapma etkileri kontrol altına alınamadığı için önüne geçmek için çeşitli kalibrasyon yöntemleri ortaya atılmıştır.

1.1.2. Kısa süreli (ikincil sapma)

Kısa süreli ya da diğer adıyla ikincil sapma etkisi, uzun süreli sapmanın aksine kontrol altına alınabilir. Çünkü bu tip etkiler sensörlerin bulunduğu çevresel faktörler sebebiyle meydana gelirler. Gaz sensörleri sıcaklık, nem ve benzeri çevresel etkilere oldukça duyarlı cihazlar olduklarından, ortamdaki bu faktörleri işlem süresince sabit

tutmak gerekir. Ayrıca ölçüm işlemlerinde, yeni gaz ortama enjekte edilmeden önce, önceki ölçümden herhangi bir gaz kalıntısının ortamda kalmamasına dikkat edilmelidir. Bu da bellek etkisi (memory affect) olarak isimlendirilir ve önceki ölçümden kalan gaz molekülleri bir sonraki ölçümü etkileyerek sensörlerde yine bir sapma etkisi meydana getirebilir [10].

1.2. Gaz Sensörlerinde Kalibrasyon

Sensörlerde zamana bağlı olarak meydana gelen kalıcı kimyasal tepkimeler ile çevresel faktörlerin sapma etkisi meydana getirdikleri belirtilmişti. Kalibrasyondaki asıl hedef de, zamanla değişen bu tepkileri çeşitli matematiksel manipülasyonlar uygulayarak olması gereken duruma yaklaştırmaktır. Bununla ilgili literatürde uygulanmış pek çok çeşit matematiksel manipülasyonlar; yani kalibrasyon yöntemleri mevcuttur. Bu çalışmada, aynı sensörlerde meydana gelen sapma düzeltme (drift correction) irdelendiği gibi; aynı sensörleri kullanan iki farklı cihaz için de kullanılacak kalibrasyon yöntemleri araştırılmıştır. Bu iki konuda asıl amaç, verilerin kalibrasyonunu yaparak sapmaya uğramış veri için yeniden bir eğitim kümesi oluşturmadan harcanan zamandan ve deney maliyetinden tasarruf edilmesidir.

Kalibrasyon yöntemleri çok değişkenli (multivariate) ve tek değişkenli (univariate) kalibrasyon yöntemleri olmak üzere ikiye ayrılarak incelenebilir. Tek değişkenli kalibrasyon yöntemlerinde, incelenen sensörün (slave) her bir değişkeni, hedef sensörün (master) her bir değişkeni ile doğrusal bir eğriye oturtulur. Eğrinin kesim noktası ile eğimi hesaplanarak incelenen sensör tepkileri hedef sensör tepkilerine yaklaştırılmaya çalışılır. Bu yöntemlerde sapmanın doğrusal olduğu var sayılır. Bu sebeple yöntemler her ne kadar kolay uygulanabilir olsa da uygulama sonucu istenilen sensör performansı elde edilemeyebilir. Çok değişkenli kalibrasyon yöntemleri ise tek bir değişkene bağlı kalmaksızın aynı anda pek çok değişkeni göz önünde bulundurarak genel bir model çıkartmaya yöneliktir [11].

Bu konu hakkında günümüz literatürünün en popüler yöntemlerinin doğrudan standartlaştırma (direct standardization-DS), parçalı doğrudan standartlaştırma (piecewise direct standardization-PDS) ve orthogonal sinyal düzeltimi (orthogonal signal correction-OSC) olduğu söylenebilir. Öyle ki, yapılan en kapsamlı çalışmalardan biri olan makalede, DS, PDS, yapay sinir ağları (artificial neural network-ANN), en yüksek olasılıklı temel bileşen analizi (maximum likelihood

principal component analysis-MLPCA) ve pozitif matris ayırımı (positive matrix factorization-PMF) yöntemleri irdelenmiş; PDS yönteminin karmaşık sistemler için en iyi çözüm olduğundan bahsedilmiştir. Ayrıca makalede, en iyi kalibrasyon tekniği için kesin bir bilgi olmadığı, veri setine göre en iyi yöntemin değişebileceği yorumlanmıştır [12].

Başka bir çalışmada, benzin kalitesini tayin etmek maksadıyla kullanılan üç farklı cihazda kalibrasyon transferi yapılmıştır [13]. DS, PDS, OSC, ters standartlaştırma (reverse standardization-RS), parçalı ters standartlaştırma (piecewise reverse standardization-PRS), eğim ve sapma düzeltimi (slope and bias correction-SBC) ve model yenileme (model updating-MU) yöntemleri 156 gaz örneği üzerinde kullanılmıştır Bu çalışma sonucunda en düşük hata oranının RS ile PDS yöntemleri ile alındığı sonucuna varılmıştır.

Bazı yöntemler her ne kadar sinyallerin ön işleme için kullanılıyor olsa da, kalibrasyon transferi uygulamalarında da sıkça kullanılır. OSC, temel bileşen analizi (principal component analysis-PCA), parçalı en küçük kareler (partial least squares-PLS) bu yöntemlere örnek olarak gösterilebilir. Dört farklı gazı algılayabilen 39 gaz sensörü ile kurulan bir düzenekten örnek alınan bir çalışmada, PCA ile PLS yöntemleri karşılaştırılmış, PLS'in daha başarılı sonuçlar verdiği gözlemlenmiştir [14].

Bu tez çalışmasında da günümüz literatüründe yoğunlukla çalışılan DS, PDS ve OSC yöntemleri incelenmiştir. Bu konu hakkında daha detaylı örnek çalışmalar ise kaynaklar bölümünde okuyucuların bilgisine sunulmuştur [15-19].

2. KALİBRASYON YÖNTEMLERİ

2.1. Doğrudan Standartlaştırma Yöntemi (Direct Standardization)

Doğrudan standartlaştırma (DS) yöntemi 1991 yılında Wang ve ekibi tarafından ortaya atılmıştır [20]. Bu yöntem ile temel (master) cihazdaki tepkiler ile ikincil (slave) cihazdaki tepkiler basit bir matematiksel denklem ile doğrudan ilişkilendirilebilir.

$$S_1 = S_2 \cdot F \quad (2.1)$$

Buradaki S_1 temel cihaza ait transfer örneklerinin oluşturduğu alt küme matrisi, S_2 ikincil cihaza ait transfer örneklerinin oluşturduğu alt küme matrisi olmak üzere; iki cihaz arasındaki ilişki bir F dönüşüm matrisi ile ifade edilmektedir. S_2^+ , S_2 matrisinin sözde (pseudo) tersini göstermek üzere (eşitlik 2.2); F kare matrisi eşitlik 2.3'teki gibi ifade edilebilir.

$$S_2^+ = (S_2^T \cdot S_2)^{-1} \cdot S_2^T \quad (2.2)$$

$$F = S_2^+ \cdot S_1 \quad (2.3)$$

F dönüşüm matrisinin hesaplanmasının ardından ikincil cihaza ait yeni verileri temsil eden x matrisinin devriği (transpose) alınır ve F matrisiyle çarpılarak manipüle edilmeye çalışılır. Yani amaç, var olan verilerle oluşturulmuş model ile yeni gelen test verilerini tahmin etmektir. Denklemden yer alan altı çizili x , modelin uygulandığı verileri temsil eder. Anlatılanlar eşitlik ile 2.4'teki gibi ifade edilebilir.

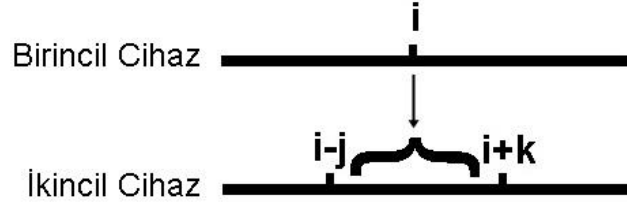
$$\underline{x} = x^T \cdot F \quad (2.4)$$

Bu model her ne kadar uygulaması kolay olsa da, bazı dezavantajlara sahiptir. Bu modeli oluşturmak için, hem birincil hem de ikincil cihazın tepki verilerine ihtiyaç vardır. Eğer iki cihazdan da yeterli sayıda örnek yok ise model başarılı bir kalibrasyon tekniği olarak kullanılamaz.

2.2. Parçalı Doğrudan Standartlaştırma Yöntemi (Piecewise Direct Standardization)

Parçalı doğrudan standartlaştırma yöntemi (PDS) de yine Wang ve ekibi tarafından literatüre kazandırılmıştır [20]. DS yönteminde, birincil cihazdaki tüm ölçümler, ikincil cihazdaki tüm ölçümler ile ilişkilendiriliyordu. Ancak gerçek uygulamalarda spektral

örtüşmeler küçük bölgelerde meydana gelir ve DS yöntemi standartlaştırmada yeterli başarıyı yakalayabilir. Bu nedenle DS'in hemen ardından PDS yöntemi önerilmiş ve ikincil cihazdaki her bir ölçümü birincil cihazdaki ölçümler ile ilişkilendirmek yerine, birincil cihazın i'ninci dalga boyundaki ölçümünü ikincil cihazın i'ninci dalga boyuna yakın olan dalga boylarındaki ölçümleri ile oluşturulan bir pencerede ilişkilendirilmesi önerilmiştir. Anlaşılması için hazırlanan görsel Şekil 2.1'de verilmiştir.



Şekil 2.1. PDS Yöntemi ile birincil cihazdaki herhangi bir i'ninci dalga boyunun ikincil cihazdaki i-j ile i+k arasındaki dalga boyları ile ilişkilendirilmeye çalışılması

İkincil cihazdaki i-j ile i+k indisleri arasında kalan tüm ölçümler eşitlik 2.5'teki gibi bir X matrisinde toplanmak üzere; b ile sembolize edilen bir regresyon vektörü, X matrisi kullanılarak eşitlik 2.6'dan hesaplanır.

$$X_i = [r_{2,i-j}, r_{2,i-j+1}, r_{2,i-j+2}, \dots, r_{2,i+k-1}, r_{2,i+k}] \quad (2.5)$$

$$r_{1,i} = X_i \cdot b_i \quad (2.6)$$

Bulunan tüm b_i değerleri de eşitlik 2.7'de verilen köşegen F matrisini oluşturur. Denklemden gösterilen n indisi, spektral kanal sayısı olarak geçer.

$$F = \text{diag} (b_1^T, b_2^T, \dots, b_{n-1}^T, b_n^T) \quad (2.7)$$

F matrisinin oluşturulması ile model meydana getirilmiş olup aynı DS yönteminde olduğu gibi, iki cihaz arasındaki farklılıkları gidermede kullanılabilir [20].

PDS en çok kullanılan kalibrasyon tekniklerinden biridir ve diğer yöntemlere göre literatürde daha başarılı bilinir. Yapılan çalışmalara göre PDS, daha yüksek kaliteli bir cihaza örnek transfer edilme amacı taşıyorsa [20] ya da dalga boylarında kayma meydana gelen bir durum söz konusuysa [12] diğer metotlara göre daha başarılı bir yöntem olarak karşımıza çıkmaktadır.

2.3. Ortogonal Sinyal Düzeltimi (Orthogonal Signal Correction)

Ortogonal sinyal düzeltimi (OSC) yöntemi her ne kadar karşımıza bir sinyal ön işleme uygulaması olarak çıksa da, literatürde kalibrasyon amacıyla da kullanılmıştır [21]. OSC yönteminde temel amaç, veri setinde gürültü olarak kabul edilen bileşenleri veriden ayırmaktır. Bunu yaparken de gürültü bileşenlerinin veriye ortogonal olduğu yaklaşımından faydalanır. Kalibrasyon işleminde de yapılan iş aynıdır: her iki cihazdan da elde edilen veri matrislerine dik olan bileşenleri ortadan kaldırarak iki cihaz arası farklılıkları azaltmaktır. OSC için kullanılan algoritma, PCA ve PLS'de kullanılan NIPALS (nonlinear iterative partial least squares) algoritması ile çok benzerdir [12].

Bu algoritmaya göre; X ve Y birbirinden bağımsız iki ayrı veri matrisi olmak üzere, PLS tabanlı OSC ile X verisiyle tamamen alakasız olan Y verisi, X'den ayıklanır. Bu işlemde Y, X verisinin filtresi gibi kullanılır. X ve Y veri matrislerinin matematiksel çıkarımları eşitlik 2.8 ve eşitlik 2.9'da verilmiştir. Bu denklemlerde T ve U skor matrislerini; P ve Q yük matrislerini, E ve F de kalan matrisleri ifade eder.

$$X = T \cdot P^T + E \quad (2.8)$$

$$Y = U \cdot Q^T + F \quad (2.9)$$

X ile sembolize edilen tek bileşenli veri için oluşturulan OSC modeli, eşitlik 2.10 ve 2.11 sağlanmak koşuluyla eşitlik 2.12'de verilmiştir.

$$Y^T \cdot t_{osc} = 0 \quad (2.10)$$

$$t_{osc} = X \cdot w_{osc} \quad (2.11)$$

$$X = t_{osc} \cdot p_{osc}^T + X' \quad (2.12)$$

Eşitliklerde kullanılan w, t ve p değişkenleri OSC bileşenleri, X' ise OSC filtreli matris olarak geçer [21,22].

3. YAPILAN ÇALIŞMALAR

3.1. Çalışmalarda Kullanılan Veri Seti

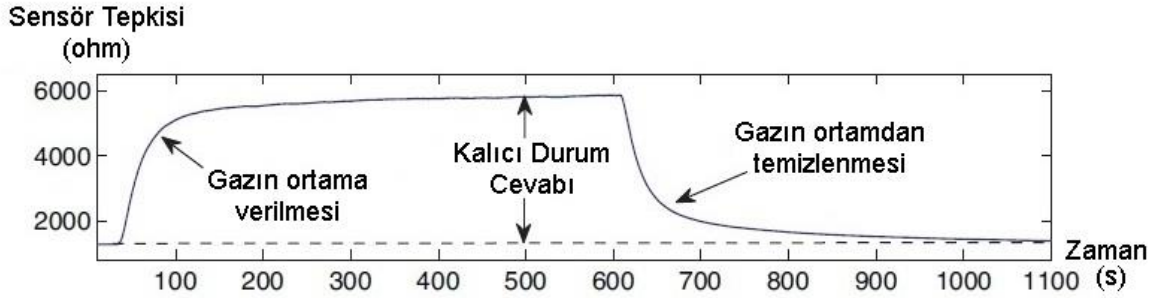
Çalışmalarda Vergara ve ekibi tarafından oluşturulmuş, çevrimiçi ücretsiz olarak erişilebilen kapsamlı bir veri seti kullanılmıştır [23]. Bu veri seti, üç yıl boyunca on altı adet sensör kullanılarak altı farklı gazın ölçümleri ile oluşturulmuştur. Çalışmada Figaro Inc. firması tarafından üretilen MOS çeşit sensörler kullanılmış ve toplamda 13.910 adet ölçümün kaydı tutulmuştur. Sensörler ve ölçüm için kurulan deney düzeneği hakkındaki dataylı bilgiye Ek 1’de yer verilmiştir. Bu çalışmadaki amaç, altı adet farklı gazı yoğunluklarına bakılmaksızın sensörler vasıtasıyla başarılı bir şekilde ayırt edebilmektir. Kullanılan gaz çeşitleri: amonyak, asetaldehit, aseton, etilen, etanol ve tolüendir.

Çalışma kapsamında tüm bu verilerden sekiz tane öznelik çıkartılmıştır. Bunları şu şekilde sıralayabiliriz:

1. Sensörlerin kalıcı durum cevabı.
2. Kalıcı durum cevabının normalize edilmiş versiyonu.
3. Sensör cevabının yükselen kısmı için, düzleme katsayısı (smoothing parameter) 0,1 iken alınan üstel hareketli ortalamanın (exponential moving average) en yüksek değeri.
4. Sensör cevabının düşen kısmı için, düzleme katsayısı 0,1 iken alınan üstel hareketli ortalamanın en düşük değeri.
5. Sensör cevabının yükselen kısmı için, düzleme katsayısı 0,01 iken alınan üstel hareketli ortalamanın en yüksek değeri.
6. Sensör cevabının düşen kısmı için, düzleme katsayısı 0,01 iken alınan üstel hareketli ortalamanın en düşük değeri.
7. Sensör cevabının yükselen kısmı için, düzleme katsayısı 0,001 iken alınan üstel hareketli ortalamanın en yüksek değeri.
8. Sensör cevabının düşen kısmı için, düzleme katsayısı 0,001 iken alınan üstel hareketli ortalamanın en düşük değeri.

Şekil 3.1’ten bir sensörün gaz algılamasına karşı verdiği tepki incelenerek çıkarılan öznelikler daha iyi anlaşılabilir. Grafik incelendiğinde; ortama gaz enjekte edilmiş iken sensör tepkisinin hızla arttığını, gazın ortamdan temizlendiği zaman da tepkisinin daha

yavaş bir şekilde düştüğü gözlemlenmektedir. Bu şekilde kalıcı durum cevabı olarak belirtilen değer yukarıda belirtilen özniteliklerden biridir ve sensör tepkisinin en yüksek değeri ile en düşük değeri arasındaki farka eşittir.

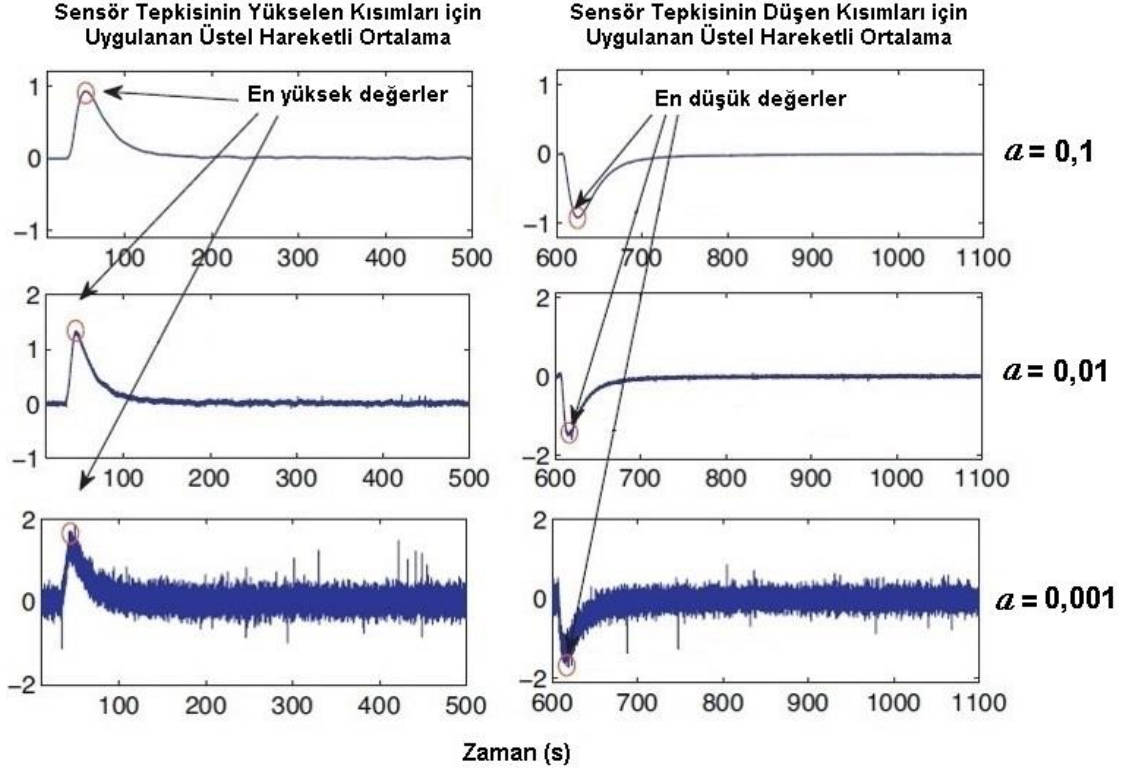


Şekil 3.1. Bir sensörün gaz algılama aşamaları [23]

Çıkarılan diğer öznitelikler ise Şekil 3.2'de daha detaylı incelenebilir. Çalışma kapsamında üstel hareketli ortalama, üç farklı düzleme parametresi ile sensörlerin hem yükselen, hem düşen tepkileri için ayrı ayrı uygulanmış ve gaz algılanırken elde edilen değerlerin en yükseği ile gaz ortamdan çekildiğinde elde edilen değerlerin en düşüğü alınarak öznitelikler çıkartılmıştır.

Üstel hareketli ortalama hesabında, r zaman serisini göstermek üzere; başlangıç koşulları ($y[0]=0$) alınarak eşitlik 3.1 kullanılmıştır.

$$y[k] = (1 - \alpha) \cdot y[k - 1] + \alpha(r[k] - r[k - 1]), \quad k = 1 \dots T \text{ ve } \alpha \in (0, 1) \quad (3.1)$$



Şekil 3.2. Sensörün yükselen ve düşen kısımları için 3 farklı düzleme parametresi ($\alpha = 0,1$, $\alpha = 0,01$ ve $\alpha = 0,001$) için ayrı ayrı alınan en yüksek ve en düşük değerler [23]

36 ay boyunca 6 farklı gazdan alınan ölçümler, düzgün bir dağılım göstermediğinden dolayı, yapılan çalışmada toplam veri seti benzer örnek sayılarına sahip olması adına 10 gruba bölünmüştür. Çizelge 3.1 üç yıl boyunca tüm gazlardan alınan ayrı ayrı ölçümleri gösterirken, Çizelge 3.2 ise veri setinin belirtilen ay ya da aylardaki ölçümleri ile 10 gruba bölünmesini göstermektedir.

Çizelge 3.1. 36 ay boyunca alınan tüm ölçümleri gösteren çizelge [23]

Ölçüm Yapılan Ay Bilgisi	Toplanan Örnek Sayısı						Toplam
	Amonyak	Asetaldehit	Aseton	Etilen	Etanol	Tolün	
1.	76	0	0	88	84	0	248
2.	7	30	70	10	6	74	197
3.	0	0	7	140	70	0	217
4.	0	4	0	170	82	5	261
8.	0	0	0	20	0	0	20
9.	0	0	0	4	11	0	15
10.	100	105	525	0	1	0	731
11.	0	0	0	146	360	0	506
12.	0	192	0	334	0	0	526
13.	216	48	275	10	5	0	554
14.	0	18	0	43	52	0	113
15.	12	12	12	0	12	0	48
16.	20	46	63	40	28	0	197
17.	0	0	0	20	0	0	20
18.	0	0	0	3	0	0	3
19.	110	29	140	100	264	9	652
20.	0	0	466	451	250	458	1625
21.	360	744	630	662	649	568	3613
22.	25	15	123	0	0	0	163
23.	15	18	20	30	30	18	131
24.	0	25	28	0	0	1	54
30.	100	50	50	55	61	100	416
36.	600	600	600	600	600	600	3600

Çizelge 3.2. Tüm veri setinin ikinci sütunda belirtilen ölçüm ay/ayları ile oluşturulan 10 adet grup bilgisini içeren çizelge

Grup No	Ölçüm Yapılan Ay
1. Grup	1. Ay, 2. Ay
2. Grup	3. Ay, 4. Ay, 8. Ay, 9. Ay, 10. Ay
3. Grup	11. Ay, 12. Ay, 13. Ay
4. Grup	14. Ay, 15. Ay
5. Grup	16. Ay
6. Grup	17. Ay, 18. Ay, 19. Ay, 20. Ay
7. Grup	21. Ay
8. Grup	22. Ay, 23. Ay
9. Grup	24. Ay, 30. Ay
10. Grup	36. Ay

3.2. Veri Setine DS, PDS ve OSC Yöntemlerinin Uygulanması

Yapılan bu çalışmada, anlatılan veri setinin 3613 adet örnek içeren 7 numaralı grubu ve 3600 adet örnek içeren 10 numaralı grubu kullanılmıştır. 7 numaralı grup, eğitim seti; 10 numaralı grup da test grubu olarak düşünülmüştür. Bu gruplar seçilirken göz önünde bulundurulmuş iki husus olmuştur:

1. Grupta her bir çeşit gaz örneğinin yeterli miktarda olması.
2. İki grubun ölçümleri arasında sapma etkisini gözlemleyecek kadar zaman farkının olması.

7 ve 10 numaralı grupların Çizelge 3.1'de görüleceği üzere, altı farklı gazdan da yeterli miktarlarda örnek barındırdığı görülmektedir. Ayrıca iki grup ölçümleri arasında 15 aylık bir zaman farkı vardır ve bu süre sapma etkisini incelemek için yeterli bir süre zarfıdır.

Analoji yapılacak olursa; 7. grup örnekleri birincil cihaz, 10. grup örnekleri de ikincil cihaz olarak düşünülebilir. Burada temel amaç 15 aylık süre zarfında yaşanan, yani ölçümlerinde sapma etkisi meydana gelen 10. grup ile güvenilirlik açısından daha iyi konumda olduğu düşünülen, 7. grup ölçümleri arasında bir ilişki yakalayıp sensörlerin sapma etkisini veriden kaldırmaktır.

Yapılan tez çalışması kapsamında, ilk olarak ham veri setine literatürde sıkça kullanılan DS, PDS ve OSC yöntemleri uygulanmıştır. Buradaki başarı ölçütü, en yakın k komşu (k-NN) algoritmasıyla altı farklı gazın ne kadar doğru sınıflandırılabilirliği. Sınıflandırma algoritmasında en yakın komşuluk parametresi, denenen her yöntem için en başarılı sınıflandırmayı sergileyen değer olarak alınmıştır. Yani literatürde geçen adıyla, çapraz doğrulama (cross validation) yapılarak k değerlerine karar verilmiştir.

İlk aşamada, ham veri setinin hiçbir yöntem uygulamadan elde edeceği sınıflandırma başarısı test edilmek istenmiştir. Bu bağlamda 7. grup eğitim, 10. grup da test seti olarak seçilmesiyle alınan sınıflandırma başarısı %64,12 olmuştur. Eğitim seti olarak 7. grup yerine; hem eğitim hem de test seti 10. gruptan seçilirse başarı oranı %87,68 olarak bulunmuştur. Eğitim ve test verileri aynı gruptan alındığında sınıflandırma performansının yüksek olduğu; ancak bunun beraberinde zaman kaybı ve maliyet yarattığı unutulmamalıdır.

Sonraki aşamada ham veri setine DS yöntemi uygulanmıştır. 7. grup eğitim seti olduğunda başarı %63,33; hem eğitim hem de test verisi 10. gruptan seçilirse başarı %60,83 olarak hesaplanmaktadır. Aynı şekilde PDS ve OSC yöntemleri de aynı veri setine uygulanmış ve eğitimde 7. grup kullanıldığında sırasıyla %67,91 ve %64,12; hem test hem eğitim seti için sadece 10. grup kullanıldığında ise sırasıyla %95,41 ile

%87,73 oranlarında başarı yakalamışlardır. Sonuçların tablolatırılmış hali Çizelge 3.3'de verilmiştir.

Çizelge 3.3. Ham veri setinin herhangi bir yöntem uygulanmadan elde edilen başarı yüzdesi ile aynı veri setine DS, PDS ve OSC yöntemlerinin uygulanmasıyla elde edilen sınıflandırma başarı yüzdeleri.

Veri Seti		Başarı (%)			
Eğitim Seti	Test Seti	Ham Veri	DS Yöntemi	PDS Yöntemi	OSC Yöntemi
10. Grup	10. Grup	87,68	60,83	95,41	87,73
7. Grup	10. Grup	64,12	63,33	67,91	64,12

Çizelge 3.3'e bakarak söylenebilir ki; ham veri seti ile kalibrasyon işlemi, tek bir cihaz (aynı gruptan alınan eğitim ve test veri setleri) verileri ile yapılsa da, iki farklı cihaz (7. grup verileri ile eğitim ve 10.grup verileri ile test) verileri arasında yapılsa da en iyi başarı oranı PDS yöntemi ile alınmıştır. Veriye sadece PDS yöntemi uygulamak; var olan başarıyı yaklaşık %4 ile %8 oranında artırmıştır. Ayrıca tabloda dikkat çeken bir diğer durum; ham veri başarısının referans olarak alınması durumunda, DS yönteminin başarıyı düşürdüğüdür. OSC yöntemi ise başarıyı yok denecek kadar az bir oranda artırmıştır.

3.3. Veri Setine Önerilen Yöntemin Uygulanması

Ham veri setine sırasıyla uygulanan DS, PDS ve OSC arasında en başarılı olan yöntemin PDS olduğu Çizelge 3.3'e bakarak yorumlanmış idi. Ancak bu tez kapsamında; literatüre yenilik kazandırması adına bu yöntemler, öncesinde öznitelik seçimi yapılarak kalibrasyona en uygun özniteliklerin belirlenmesinin ardından uygulanmıştır.

3.3.1. Öznitelik seçimi

Öznitelik, veri matrisinde her bir satırı betimleyen bir sütundur denilebilir. Daha kısa ve öz bir tanımla da veri setinin sahip olduğu özellik olarak geçer. Bir veri setinde kaç adet öznitelik olacağı o veri setine göre değişkenlik gösterir. Basit bir örnek verilecek olursa; hava durumu ile oluşturulan kapsamlı bir veri setinde; sıcaklık, nem, rüzgar, basınç gibi ölçümler o veri setinin birer özneliği olabilir [24].

Kapsamlı veri setlerinde, genellikle matris bazında işlemler uygulandığı için, matris boyutları oldukça büyük olur ve bu da işlem zamanı ile maliyetini artırır. Ayrıca veri setinin sahip olduğu tüm öznitelikler, çalışmalarda kullanılmak istenmiyor olabilir. Bu durumda çözüm öznitelik seçimi yapmaktan geçer. Çünkü öznitelik çıkarımı ile sadece incelenmek, yani işlem yapılmak istenen, öznitelikler dikkate alınır; geri kalanları veriden ayıklanır. Bu şekilde işlem zamanı ve maliyetten kazanç sağlanabilir. Ayrıca analizler sırasında bazı öznitelikler, incelenmek istenen diğer öznitelikleri baskılıyor olabilir. Bu durum da, incelenmek istenen değişken ya da değişkenlerin hatalı yorumlanmasına sebebiyet verebilir. Kısacası öznitelik çıkarımı, veriyi daha anlaşılır hale getirir.

Öznitelik seçimi için hali hazırda pek çok yöntem vardır. İleriye doğru öznitelik seçimi (forward feature selection-FFS) ve geriye doğru öznitelik seçimi (back feature selection-BFS) bunlardan en çok kullanılanlarıdır. Bu tez kapsamında yapılan çalışmada da ileriye doğru öznitelik seçimi (FFS) yöntemi kullanılmıştır.

3.3.1.1. Geriye doğru öznitelik seçimi (BFS)

Geriye doğru öznitelik seçimi (BFS); FFS ile aynı amacı taşır: veriden en uygun eğitim kümesini oluşturmaktır. FFS'den farkı, öznitelik eleme işleminin geriye doğru, yani çoktan aza olacak şekilde, yapılmasıdır. İlk aşamada veri setinin sahip olduğu tüm öznitelikler ile işleme başlanır ve ardından birer birer bu öznitelikler çıkartılarak hiçbir öznitelik kalmayana kadar devam ettirilir ve başarı değerlendirmesi yapılır. En yüksek başarı elde edilen küme, öznitelik kümesi olarak seçilir ve analizlerde eğitim seti olarak kullanılmaya hazır hale gelir.

3.3.1.2. İleriye doğru öznitelik seçimi (FFS)

Öznitelik seçme işlemi azdan çoğa doğru yapıldığı için ileriye doğru olarak adlandırılmıştır. İleriye doğru öznitelik seçimi (FFS); boş bir öznitelik seti ile başlayıp, o sete veriden birer birer öznitelik ekleyerek öznitelik kümesini oluşturan ve bu şekilde de en iyi eğitim setini oluşturmayı amaçlayan bir yöntem olarak geçer [25]. Eklenen her bir öznitelik ile başarı kıyaslaması yaparak, en yüksek başarıyı veren öznitelik kümesini, veri setinin sahip olduğu tüm öznitelikleri teker teker ekleyerek seçer. Yöntemin dezavantajı; öznitelik kümesine eklenen bir özneliğin, çıkartılarak geriye yönelik işlem yapılamamasıdır.

Tez kapsamında FFS yöntemi, çok öznitelik barındıran veri setleri için zaman kazandırması avantajı sayesinde tercih edilmiştir. Yöntemin uygulanması sonucunda seçilen özniteliklerin bilgisi Çizelge 3.4'te sergilenmiştir. Veri setini en çok etkileyen sensörlerin 2., 4., 9., 10., 11., 14. ve 16. sensörlerin olduğu ve tüm sensörler için en iyi özneliğin de ΔR ile gösterilen kalıcı durum cevabı olduğu gözlemlenmiştir.

Çizelge 3.4. FFS yönteminin uygulanması sonucunda seçilen öznitelikler

Seçilen Sensör	2. Sensör	4. Sensör	9. Sensör	10. Sensör	11. Sensör	14. Sensör	16. Sensör
Seçilen Öznitelik	ΔR	ΔR	ΔR	ΔR	ΔR	ΔR	ΔR

Basit bir sınıflandırma başarısı ölçen algoritmayla 7. grup verilerini içeren matris ve 10. grup verilerini içeren matris ile çalışmada incelenmek üzere iki ayrı eğitim kümesi meydana getirilmiştir. Oluşturulan eğitim kümesi ile 3.2. bölümünde incelenen tüm yöntemler, öznitelik seçimi yapılmış daha küçük boyuttaki bu matrislere uygulanmıştır. Öznitelik seçimi yapılan veriye yine sırasıyla DS, PDS ve OSC yöntemleri uygulanarak sonucunda elde edilen başarı oranlarına Çizelge 3.4'te detaylı olarak yer verilmiştir. Sınıflandırma başarısı ölçütü yine k-NN algoritmasıyla elde edilen değerlerdir.

Çizelge 3.5. En iyi öznitelik seçilerek oluşturulmuş veri setine herhangi bir yöntem uygulanmadan elde edilen başarı yüzdesi ile aynı öznitelikli veri setine DS, PDS ve OSC yöntemlerinin uygulanmasıyla elde edilen sınıflandırma başarıları yüzdeleri.

Veri Seti		Başarı (%)			
Eğitim Seti	Test Seti	Öznitelikli Veri	DS Yöntemi	PDS Yöntemi	OSC Yöntemi
10. Grup	10. Grup	88,14	88,70	94,81	88,14
7. Grup	10. Grup	72,87	77,17	75,69	72,87

Çizelge 3.5'ten de anlaşılıyor ki; öznitelik seçilmiş veri seti ile uygulanan yöntemler arasında en başarılı performansı, 10. grup eğitim verisiyle PDS; 7. grup eğitim verisiyle DS vermiştir. OSC yöntemi başarı oranında bir değişiklik yapmazken, özellikle PDS yönteminin başarıyı tatmin edici bir oranda artırdığı gözlemlenmiştir.

Bu sonuçların daha anlamlı bilgi teşkil etmesi adına, hata matrisleri Çizelge 3.6 ve Çizelge 3.7'de sırasıyla ham veri ve öznitelik seçimi yapılmış veri analizleri için oluşturulmuştur. Ayrıca her bir gazın özgünlük (specificity) ve hassaslık (sensitivity) tablolarına da Çizelge 3.8 ve Çizelge 3.9'da yer verilmiştir.

Çizelge 3.6. Ham veri setinden oluşturulmuş 7. grup eğitim verisi ile ortaya çıkan hata matrisi.
(i.Etanol, j.Etilen, k.Amonyak, l.Asetaldehit, m.Aseton, n.Tolün)

	Tahmin Değerleri						
		A^i	B^j	C^k	D^l	E^m	F^n
Gerçek Değerler	A^i	358	0	0	22	0	0
	B^j	0	331	0	0	1	0
	C^k	0	51	118	0	0	0
	D^l	51	40	0	444	0	3
	E^m	0	0	2	0	324	2
	F^n	0	0	0	38	65	323

Çizelge 3.7. Öznitelik seçimi ile oluşturulmuş veri setinden 7. grup eğitim verisi ile ortaya çıkan hata matrisi.
(i.Etanol, j.Etilen, k.Amonyak, l.Asetaldehit, m.Aseton, n.Tolün)

	Tahmin Değerleri						
		A^i	B^j	C^k	D^l	E^m	F^n
Gerçek Değerler	A^i	296	0	0	22	0	0
	B^j	0	338	2	0	0	0
	C^k	0	21	357	0	0	0
	D^l	64	0	0	337	0	0
	E^m	0	0	0	0	360	0
	F^n	0	1	1	1	0	360

Çizelge 3.8. Ham veri seti ile yapılan analiz sonucunda tüm gazların sahip olduğu özgünlük ve hassasiyet yüzde değerleri.

Gaz Çeşidi	Özgünlük (%)	Hassasiyet (%)
Etanol	97,16	94,21
Etilen	95,06	99,70
Amonyak	99,90	69,82
Asetaldehit	96,33	82,53
Aseton	96,42	98,78
Tolün	99,71	75,82

Çizelge 3.9. Öznitelik seçimi ile yapılan analiz sonucunda tüm gazların sahip olduğu özgünlük ve hassasiyet yüzde değerleri.

Gaz Çeşidi	Özgünlük (%)	Hassasiyet (%)
Etanol	96,52	93,08
Etilen	98,79	99,41
Amonyak	99,83	94,44
Asetaldehit	98,69	84,03
Aseton	100	100
Tolün	100	99,17

Hata matrisleri, sınıflandırma performansını ölçmek amacıyla kullanılan bir ölçüttür. Doğru ve yanlış tahmin edilen kaç adet örnek olduğunu tablo halinde görselleştirmeye yarar. Bu tablolara göre de özgünlük ve hassasiyet tabloları oluşturulmuştur. Hassasiyet, verilerin doğru sınıflandırma ölçütü olarak; özgünlük ise olmayan veriyi doğru tahmin etme ölçütü olarak değerlendirilir. Çizelge 3.8 ile Çizelge 3.9

kıyaslandığında, özellikle amonyak ve tolün gazlarının öznelik seçimi ile daha doğru sınıflandırıldığı fark edilir.



4. SONUÇ

Çizelge 3.3 ile Çizelge 3.5 karşılaştırıldığında; sadece öznitelik seçiminin bile başarı oranını dikkate alınacak miktarda artırdığı; yöntemler bazında bakılacak olursa da PDS ve DS yöntemlerinin ön plana çıktığı sonucuna rahatlıkla varılabilir.

Eğitim setini 7.grup; test setini 10. grup; yani önceki bölümlerde yapılan analoji ile, eğitim setini birincil cihazdan; test setini ikincil cihazdan alınarak yapılan karşılaştırma çalışmadaki esas amacı oluşturur. Çünkü cihazlar arası kalibrasyon transferi, ikincil cihazda eğitim kümesinin tekrar oluşturulmaya gerek kalmaması sayesinde zaman ve maliyet konusunda büyük avantaj sağlar. Tek bir cihaz verileriyle kalibrasyon işlemi ise çalışmada yöntem başarısını görmede yan kriter olarak kullanılmış ve başarının o kriterde de sağlandığı sonucuna varılmıştır. Bu bağlamda, Çizelge 3.3'te ham veriye hiç bir yöntem uygulamadan elde edilmiş başarı oranı %64.12 iken; FFS yöntemiyle öznitelik seçimi yapılmış veri ile yine hiç bir kalibrasyon yöntemi uygulamadan elde edilen başarı oranı %72,87 olmuştur. Yani kalibrasyon yöntemine gerek duyulmaksızın yalnızca en iyi öznitelik seçimi bile sistem performansını artırabilir. Yöntemler uygulanarak inceleme yapıldığında; PDS'in ham veri işlendiğinde en iyi; DS'in ise öznitelik seçilmiş verinin işlenmesiyle en iyi performansı sağladığı görülmektedir.

FFS yöntemiyle yapılan öznitelik seçiminde, seçilen tüm özniteliklerin sensörlerin kalıcı durum cevaplarının (ΔR) olması, çalışmada göze çarpan diğer bir detaydır. Bu detay ile sapma etkisinden en az etkilenen özniteliğin ΔR olduğu yorumlanabilir. Ayrıca bu öznitelik tüm sensörler için değil; en iyi durumda olan sensörler için alınmış ve bu sayede seçilen daha küçük boyuttaki veri setiyle sınıflandırma başarısı artırılmıştır.

Tez kapsamında yapılan bu çalışma, sapma etkisinden en az etkilenen öznitelikler ile daha yüksek başarılar elde etmenin mümkün olduğunu göstermektedir. Elde edilen bu bulgular ışığında, özellikle makine öğrenmesi alanında daha yüksek başarı oranları ve daha doğru analizler ile gelecek çalışmalar literatüre kazandırılabilir. Ancak unutulmamalıdır ki; tez kapsamında öznitelikleri çıkarılmış, tek bir veri setiyle çalışılmıştır. Gelecek çalışmalarda önerilen yöntemin farklı veri setleri ile de denenmesi tez analiz sonuçlarını daha gerçekçi kılacaktır.

Bu tez çalışmasının ardından gelecek çalışmalarımız arasında, sınıflandırıcı çeşidinin sapma ve kalibrasyon transferine nasıl etki ettiği konusu yer almaktadır. Çünkü bu çalışma kapsamında yöntemler sadece k-NN sınıflandırıcısı ile irdelenmiştir. Ayrıca bu veri setine farklı öznitelik seçim algoritmalarının uygulanmasıyla seçilecek olan öznitelik çeşidinin incelenmesi de yakın zaman hedeflerimizden arasında yer almaktadır.



KAYNAKLAR LİSTESİ

- [1] SOLÓRZANO, A., Fonollosa J. and Marco S., Improving Calibration of Chemical Gas Sensors for Fire Detection Using Small Scale Setups, Eurosensors 2017, Paris-FRANCE, Proceedings 1, no: 4: p. 453, 2017.
- [2] RODRÍGUEZ-MÉNDEZ, M.L, De Saja, J.A, González-Antón, R, et al., Electronic Noses and Tongues in Wine Industry, Frontiers in Bioengineering and Biotechnology, vol.4, p.81, 2016.
- [3] AKSEBZECİ, B.H., Mikroorganizmaların Elektronik Burun Teknolojisi Kullanılarak Tanımlanması ve Sınıflandırılması, Doktora Tezi, Erciyes Üniversitesi Fen Bilimleri Enstitüsü Elektrik-Elektronik Mühendisliği Anabilim Dalı, Kayseri-Türkiye, 2011.
- [4] CASALINUOVO, I. A, Di Pierro, D, Coletta, M, Di Francesco, P. Application of Electronic Noses for Disease Diagnosis and Food Spoilage Detection, Sensors, vol: 6, no:11, pp.1428-1439, 2006.
- [5] FIGARO ENGINEERING INCORPORATION, Operating Principle, <http://www.figaro.co.jp/en/technicalinfo/principle/mos-type.html>, 2018.
- [6] PETERSON, P. J. D., Aujla, A., Grant, K. H., Brundle, A. G., Thompson M. R., Hey, J. V., and Leigh R. J., Practical Use of Metal Oxide Semiconductor Gas Sensors for Measuring Nitrogen Dioxide and Ozone in Urban Environments, vol.17, no: 7, p. 1653, 2017.
- [7] KAKOTY, P. and Bhuyan, M., Design and Electro-Thermal Analysis of a Meander Shaped Platinum Microheater for SnO₂ Gas Sensors, IEEE Sponsored 3rd International Conference On Electronics And Communication Systems, vol: 78, p.1836, 2016.
- [8] HAUGEN, J., Tomic, O., Kvaal, K., A Calibration Method For Handling The Temporal Drift Of Solid State Gas-Sensors, Anal. Chim. Acta, vol. 407, pp. 23-39, 2000.

- [9] PEARCE, T.C., Schiffman s.s, Nagle, H.T. and Gardner J.W., Handbook of Machine Olfaction. 1st ed., 2002.
- [10] MARCO, S., Gutierrez-Galvez, A., Signal and Data Processing for Machine Olfaction and Chemical Sensing: A Review, vol:12, pp. 3189 – 3214, 2012.
- [11] ERGÜN, E.G.C., Bakır (II) ve Çinko (II)'nin Birlikte Tayininde Potansiyometrik ve Kemometrik Yöntemlerin Uygulanması, Yüksek Lisans Tezi, Ankara Üniversitesi Fen Bilimleri Enstitüsü, Ankara-Türkiye, 2010.
- [12] FEUDALE, R.N., Woody, N.A., Tan, H., Myles, A.J., Brown, S.D., and Ferre, J., Transfer of multivariate calibration models: A review, Chemomet. Intell. Lab. Syst., vol. 64, pp. 181-192, 2002.
- [13] PEREIRA, C.F., Pimentel, M.F., Galvao, R.K.H., Honorato, F.A., Stragevitch, L., and Martins, M.N., A Comparative Study of Calibration Transfer Methods for Determination of Gasoline Quality Parametres in Three Different Near Infrared Spectrometers, Anal. Chim. Acta, vol. 611, pp. 41-47, 2008.
- [14] ARTURSSON, T., Eklöv, T., Lundström, I., Martensson, P., Sjöström, M., and Holmberg, M., Drift Correction for Gas Sensors Using Multivariate Methods,” J. Chemomet., vol. 14, pp. 711-723, 2000.
- [15] FERNANDEZ, L., Guney, S., Gutierrez-Galvez, A., and Marco, S., Calibration Transfer in Temperature Modulated Gas Sensor Arrays, Sensors and Actuat. B, Chem., vol. 231, pp. 276-284, 2016.
- [16] GALVÃO, R.K.H., Soares, S.F.C., Martins, M.N., Pimentel, M.F., and Araújo, M.C.U., Calibration Transfer Employing Univariate Correction and Robust Regression, Anal. Chim. Acta, vol. 864, pp. 1-8, 2015.
- [17] MALLI, B., Birlutiu, A., and Natschläger, T., Standard-Free Calibration Transfer - an Evaluation of Different Techniques, Chemomet. Intell. Lab. Syst., vol.161, pp. 49-60, 2017.
- [18] PANCHUK, V., Kirsanov, D., Oleneva, E., Semenov, V., Legin, A., Calibration Transfer Between Different Analytical Methods, Talanta, vol. 170, pp. 457-463, 2017.

- [19] ZHANG, F., Chen, W., Zhang, R., Ding, B., Yao, H., Ge, J., Ju, L., Yang, W., and Du, Y., Sampling Error Profile Analysis for Calibration Transfer In Multivariate Calibration, *Chemomet. Intell. Lab. Syst.*, vol. 171, pp. 234-240, 2017.
- [20] WANG, Y., Veltkamp, D.J., and Kowalski, B.R., Multivariate Instrument Standardization," *Anal. Chem.*, vol. 63, pp. 2750–2756, 1991.
- [21] WOLD, S., Antti, H., Lindgren, F., Öhman, J., Orthogonal Signal Correction of Near-Infrared Spectra, *Chemomet. Intell. Lab. Syst.*, vol. 44, pp. 175-185, 1998.
- [22] NIAZI, A., AZIZI, A., Orthogonal Signal Correction – Partial Least Squares Method for Simultaneous Spectrophotometric Determination of Nickel, Cobalt, and Zinc, *Turk J. Chem.*, vol. 32, pp. 217-228, 2008.
- [23] VERGARA, A., Vembu, S. Ayhan, T., Ryan, M.A., Homer, M.L., Huerta, R., Chemical Gas Sensor Drift Compensation Using Classifier Ensembles, *Sensors and Actuat. B, Chem.*, vol. 166-167, pp. 320-329, 2012.
- [24] ALAN, K., Veri Madenciliğine Giriş, <https://medium.com/@kadiralan021/veri-madencili%C4%9Ffine-giri%C5%9F-fa0a63e655fb>, 2017.
- [25] CHANDRASHEKAR, G., and Sahin, F., A Survey on Feature Selection Methods, *Computers and Electrical Engineering*, vol. 40, pp. 16-28, 2014.

EK 1 Veri toplamada kullanılan deneysel set düzeneği ve ölçüm hakkında detaylı bilgiler

Çalışmada, Figaro Inc. tarafından üretilmiş TGS2600, TGS2602, TGS2610, TGS2620 kodlarına sahip toplam 16 adet (her birinden dörder adet olmak üzere) gaz sensörü kullanılmıştır. 36 ay boyunca alınan ölçümlerde ana metinde de bahsedildiği üzere 6 farklı gaz tespit edilmeye çalışılmıştır. Bahsedilen sensörler, 60 mililitre hacmindeki bir gaz odacığına yerleştirilmiş ve gaz akışını kontrol eden bir aparat yardımıyla arzu edilen miktarda gaz, odacığa enjekte edilerek sensör tepkileri veri toplama kartı vasıtasıyla bilgisayar üzerinde kaydedilmiştir. Ölçümler sırasında sıcaklık 25 ± 1 °C seviyesinde, nem ise %10 oranında da tutulmaya çalışılmıştır. Sensörler, ölçüm sonuçları alınmadan önce üretici firması tarafından belirlenen sıcaklık olan 400°C'de ısıtılmıştır. Kurulan deney düzeneği aşağıdaki görselde verilmiştir.

