

ÇEV 715
Hava Kirliliđi Modellemesi ve Uygulamaları

Model Performansının Belirlenmesi

Özgür ZEYDAN (PhD.)

<http://www.ozgurzeydan.com/>

Kantil - Kantil (Q-Q) Grafikleri

- Q-Q grafiklerinde, ölçülen değerler “x” ekseninde ve model ile tahmin edilen değerler de “y” ekseninde yer alacak şekilde bir x-y dağılım grafiği oluşturulur.
- Bu noktalar kullanılarak çizilen en iyi çizgini $y=x$ doğrusunun altında kalması modelin gerçekteki verileri daha düşük oranda tahmin ettiğini (under-estimate), çizginin $y=x$ doğrusunun üzerinde yer alması ise modelin verileri daha yüksek oranda tahmin ettiğini (over-estimate) gösterir.
- En iyi performans gösteren modeller için de grafikteki noktalar $y=x$ doğrusu üzerinde veya çok yakınındadır.

Taşma Faktörü

- Modelin az veya fazla tahmin etme miktarının belirlenebilmesi için “taşma faktörü” (factor of exceedance - FOEX) değeri hesaplanır.
- FOEX değeri -%50 ile +%50 arasında değişmekte olup en iyi model %0 FOEX değerine sahiptir.

$$FOEX = \left[\frac{n_{(P_i > O_i)}}{n} - 0.5 \right] \times (\%100)$$

- $n_{(P_i > O_i)}$: Tahmin edilen değer ölçüm değerini geçtiği verilerin sayısı
- n : Toplam veri sayısı

Kalan (Residual)

- Atmosferik konsantrasyonların belirlendiği hava kalitesi modellerinin performansları belirleyebilmek için de “Kalan” (residual) miktarının hesaplanması gerekmektedir.
- “Kalan” değeri, sahada ölçülen kirleticinin konsantrasyon ile model tarafından tahmin edilen konsantrasyon arasındaki fark olarak tanımlanır.

$$d = O - P$$

- d: Kalan
- O: Ölçülen değer (observed)
- P: Tahmin edilen değer (predicted)

Ortalama Hata

- Farklı zamanlar için hesaplanan kalan değerlerinin ortalamaları alındığında “Ortalama Hata” (mean error - ME) belirlenmiş olur.

$$ME = \bar{d} = \frac{1}{n} \sum_{i=1}^n (O_i - P_i)$$

- ME: Ortalama hata
- n: Veri sayısı

Yanlılık (Bias)

- Bazı kaynaklarda “Yanlılık” (bias) olarak da belirtilen ortalama hata, en basit ifadeyle modelin konsantrasyonları az veya fazla olarak tahmin ettiğinin göstergesidir.
- $(-\infty, +\infty)$ aralığında olabilen ortalama hata için ideal değer 0’dır.
- Ortalama hata değeri, pozitif ve negatif değerlerin birbirini götürmesi nedeniyle performans belirlemede uygun sonuç vermemektedir.
- Dolayısıyla, daha farklı istatistiksel parametrelerin türetilmesine ihtiyaç duyulmuştur.

Ortalama Mutlak Hata (Mean Absolute Error - MAE)

$$MAE = \frac{1}{n} \sum_{i=1}^n |O_i - P_i|$$

- \bar{O} : Ölçülen değerlerin ortalaması
- \bar{P} : Tahmin edilen değerlerin ortalaması
- n: Veri sayısı
- Ortalama mutlak hata (MAE) da ise ölçülen ve tahmin edilen konsantrasyonlar arasındaki farkın mutlak değeri alındığından hatanın büyüklüğünü miktar olarak vermekte fakat yanlışlığın yönü yani modelin az veya fazla tahmin etmesi hakkında herhangi bir bilgi vermemektedir.

Ortalama Hatanın Karesi (Mean Square Error - MSE)

$$MSE = \frac{1}{n} \sum_{i=1}^n (O_i - P_i)^2$$

- \bar{O} : Ölçülen değerlerin ortalaması
- \bar{P} : Tahmin edilen değerlerin ortalaması
- n : Veri sayısı
- Ortalama hatanın karesi'nde (MSE) ise ortalama alınmadan önce farkların karesi alındığından hata miktarı da büyümektedir. Hatta uç değerlerin (outliers) varlığı MSE değerinin daha da artmasına sebep olacaktır.

Hata Kareleri Ortalamasının Karekökü (Root Mean Square Error - RMSE)

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (O_i - P_i)^2}$$

- \bar{O} : Ölçülen değerlerin ortalaması
- \bar{P} : Tahmin edilen değerlerin ortalaması
- n : Veri sayısı
- Hata kareleri ortalamasının karekökü (RMSE), MSE değerinin karekökünü aldığından, hatanın boyutları verilerin boyutları ile aynı büyüklükte ifade edilebilmektedir. Aynı zamanda, veri setindeki uç değerlerin atılması durumunda RMSE değeri diğer parametrelere göre daha fazla etkilenmektedir.

Normalleştirilmiş Ortalama Hatanın Karesi (Normalized Mean Square Error - NMSE)

$$NMSE = \frac{\frac{1}{n} \sum_{i=1}^n (O_i - P_i)^2}{\bar{O} \times \bar{P}}$$

- \bar{O} : Ölçülen değerlerin ortalaması
- \bar{P} : Tahmin edilen değerlerin ortalaması
- n: Veri sayısı
- Normalleştirilmiş ortalama hatanın karesi (NMSE) tüm veri setinin dağılımına odaklanır. En iyi model NMSE değeri 0 için elde edilirken NMSE ≤ 0.5 olduğu durumlarda modelin başarılı olduğu kabul edilmektedir. Paydada yer alan ortalamaların çarpım terimi ile normalleştirilen NMSE, modelin az veya fazla tahmin etmesinden dolayı yanlışlık göstermeyecektir.

Kesirli Sapma (Fractional Bias - FB)

$$FB = \frac{(\bar{O} - \bar{P})}{0.5 \times (\bar{O} + \bar{P})}$$

- \bar{O} : Ölçülen değerlerin ortalaması
- \bar{P} : Tahmin edilen değerlerin ortalaması
- n: Veri sayısı
- Kesirli sapma (FB) parametresi (-2, 2) aralığında değer almakta ve en iyi model için de 0 sonucunu vermektedir. Negatif FB değeri modelin konsantrasyonları fazla tahmin ettiğini, pozitif FB değeri de konsantrasyonların model tarafından daha düşük olarak tahmin edildiğini gösterir.

Korelasyon Katsayısı (Correlation Coefficient - r)

$$r = \frac{\frac{1}{n} \sum_{i=1}^n (O_i - \bar{O})(P_i - \bar{P})}{\sigma_O \times \sigma_P}$$

- \bar{O} : Ölçülen değerlerin ortalaması
- \bar{P} : Tahmin edilen değerlerin ortalaması
- n: Veri sayısı
- σ_O : Ölçülen değerlerin standart sapması
- σ_P : Tahmin edilen değerlerin standart sapması

- Korelasyon analizi, hem nümerik hem de grafiksel hesaplamaları içermektedir.
- Nümerik sonuçlar, ölçülen ve model tarafından tahmin edilen değerler arasındaki ilişkiyi kantitatif olarak, grafiksel sonuçlar da bu ilişkiyi kalitatif olarak ifade ederler.
- Korelasyon katsayısı, en iyi sonuç veren model için 1 değerini alır.

Uyum İndeksi (Index of Agreement - IA)

$$IA = 1 - \frac{\sum_{i=1}^n (P_i - O_i)^2}{\sum_{i=1}^n (|P_i - \bar{O}| + |O_i - \bar{O}|)^2}$$

- \bar{O} : Ölçülen değerlerin ortalaması
- \bar{P} : Tahmin edilen değerlerin ortalaması
- n: Veri sayısı

- Uyum indeksi (IA), model tahminlerindeki hata derecesinin standartlaştırılmış ölçümü olarak tanımlanabilir.
- IA, 0 ile 1 arasında değer alır.
- 1 değeri ölçülen ve tahmin edilen konsantrasyonlar arasındaki uyumun mükemmel olduğunu belirtir.

2'nin Faktörü (Factor of 2 – FAC2)

- FAC2 = Veri seti içerisinde $0.5 \leq \frac{P}{O} \leq 2.0$ koşulunu sağlayan veri oranı.
- Uç değerlere karşı en dirençli parametre 2'nin Faktörü (Factor of 2 – FAC2)'dür.
- $FAC2 \geq 0.8$ olan modeller başarılı kabul edilmektedir.

Geometrik Ortalama Hata (MG) ve Geometrik Varyans (VG)

- Atmosferik kirleticilerin konsantrasyon dağılımları genellikle log-normaldir. Bu nedenle, doğrusal olan parametreler (FB ve NMSE gibi), seyrek olarak ölçülen veya tahmin edilen yüksek konsantrasyonlardan daha fazla etkilenebilmektedirler.
- Fakat, Geometrik Ortalama Hata (MG) ve Geometrik Varyans (VG) gibi logaritmik parametreler, çok yüksek ve çok düşük konsantrasyon değerlerinden doğrusal parametreler kadar etkilenmemektedir.

$$MG = \exp(\overline{\ln O - \ln P}) \quad VG = \exp(\overline{(\ln O - \ln P)^2})$$

- Ancak logaritmik parametreler 0'dan büyük değerler için tanımlı olduklarından, ölçüm cihazlarının eşik değerlerinden daha düşük konsantrasyonların olması durumunda bu parametreler tanımsız olacaktır.

Model Performansının Belirlenmesi

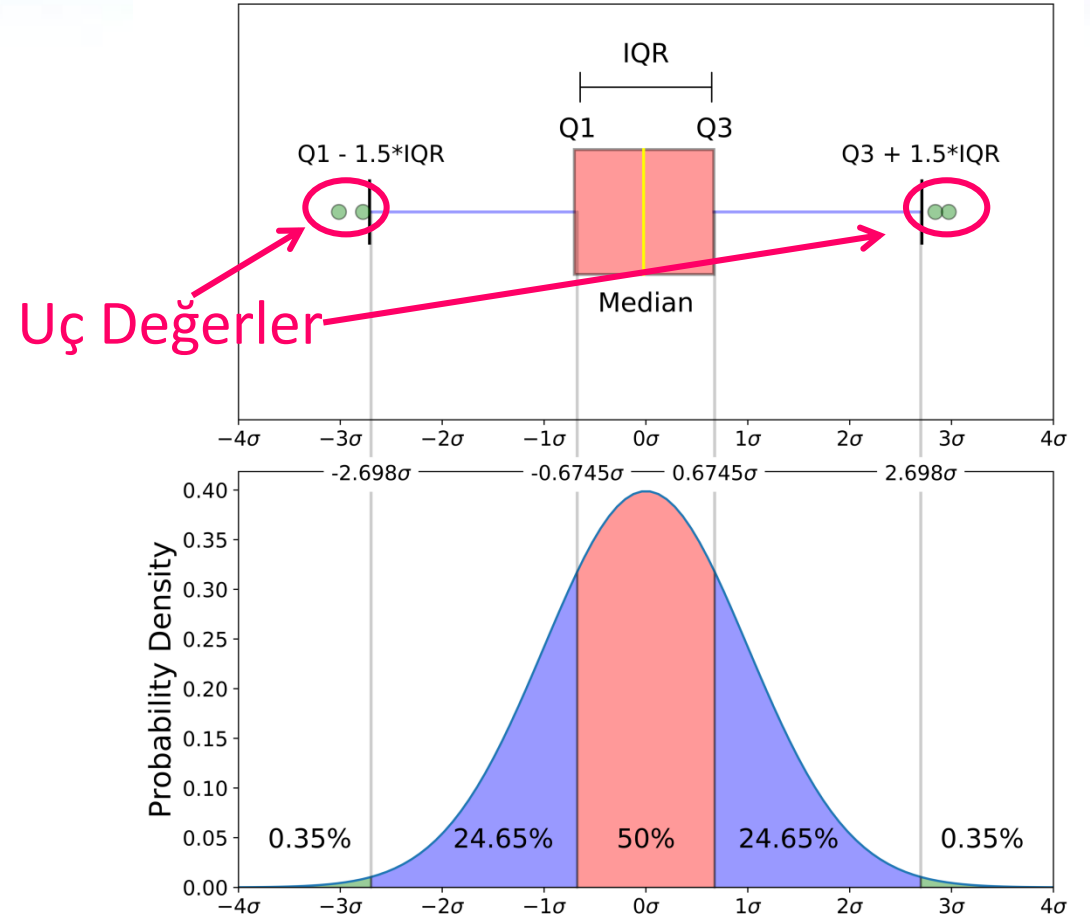
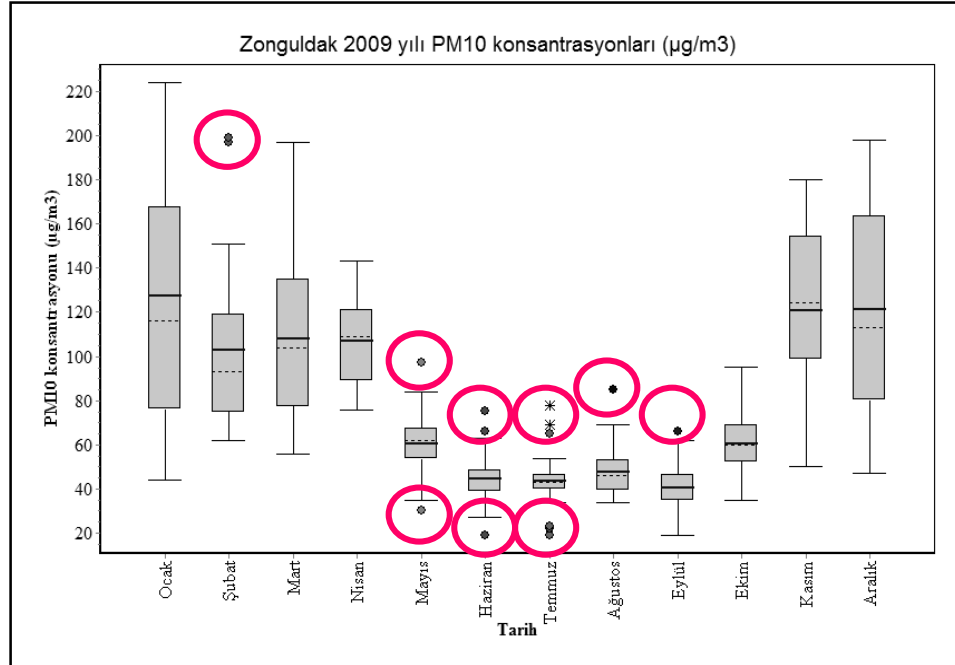
Parametre	Değer aralığı	En iyi değer	Kabul değeri
FOEX	(-%50, +%50)	%0	
ME	$(-\infty, +\infty)$	0	
MAE	$(0, +\infty)$	0	
MSE	$(0, +\infty)$	0	
RMSE	$(0, +\infty)$	0	
NMSE	yok	0	$NMSE \leq 0.5$
FB	$(-2, 2)$	0	$-0.3 \leq FB \leq 0.3$ (kırsal), $-0.67 \leq FB \leq 0.67$ (kentsel)
r	$(-1, 1)$	1	
IA	$(0, 1)$	1	$IA \geq 0.5$
MG	> 0	1	$0.7 < MG < 1.3$
VG	> 0	1	$VG < 1.6$
FAC2	$(0.5, 2)$	1	$FAC2 \geq 0.8$

Hava Kalitesi Modellerindeki Belirsizlikler

- Hava kalitesi modellerindeki belirsizlikler sistematik ve sistematik olmayan belirsizlikler olarak ikiye ayrılır.
- Sistematik belirsizlikler, modelin yapısı ve modele girilen verilerden kaynaklanırken, sistematik olmayan belirsizliklerin kaynağı ise atmosferik ve ölçümler sırasında meydana gelen rasgele türbülanslardır.
- FB ve MG parametreleri sadece sistematik hataları ölçebilirken, RMSE, NMSE ve VG hem sistematik hem de sistematik olmayan hataları belirleyebilmektedir.

Veri Setindeki Uç Değerlerin Belirlenmesi

- Ölçüm veya model sonuçları aylık olarak kutu grafiği (box plot) ile gösterildiğinde uç değerler veri setinden ayrılabilir.



<https://towardsdatascience.com/understanding-boxplots-5e2df7bcbd51>

Atıf

- Bu slayta aŖağıdaki Ŗekilde atıf yapabilirsiniz:
- Zeydan, Ö. (2014) Zonguldak Bölgesi PM_{10} Konsantrasyonu Dağılımının Modellenmesi, YayınlanmamıŖ Doktora Tezi, Kocaeli Üniversitesi, Fen Bilimleri Enstitüsü, Kocaeli.

Kaynaklar

- Behera S. N., Sharma M., Dikshit O., Shukla S. P., GIS-Based Emission Inventory, Dispersion Modeling and Assessment for Source Contributions of Particulate Matter in an Urban Environment, *Water Air Soil Pollution*, 2011, **218**, 423-436.
- Bennett N. D., Croke B. F. W., Guariso G., Guillaume J. H. A., Hamilton S. H., Jakeman A. J., Marsili-Libelli S., Newham L. T. H., Norton J. P., Perin C., Pierce S. A., Robson B., Seppelt R., Voinov A. A., Fath B. D., Andreassian V., Characterising Performance of Environmental Models, *Environmental Modelling & Software*, 2013, **40**, 1-20.
- Borrego C. Monteiro A., Ferreira J., Miranda A.I., Costa A.M., Carvalho A.C., Lopes M., Procedures for Estimation of Modelling Uncertainty in Air Quality Assessment, *Environment International*, 2008, **34**, 613-620.
- Boubel R.W., Fox D.L., Turner D. B., Stern A.C., *Fundamentals of Air Pollution*, 3rd ed., Academic Press, U.S.A., 1994.

Kaynaklar

- Chang J. C., Hanna S. R., Air Quality Model Performance Evaluation, *Meteorol Atmos Phys*, 2004, **87**, 167-196.
- Donnelly R. P., Lyons T. J., Flassak T., Evaluation of Results of a Numerical Simulation of Dispersion in an Idealised Urban Area for Emergency Response Modelling, *Atmospheric Environment*, 2009, **43**, 4416-4423.
- Hanha S. R., Air Quality Model Evaluation and Uncertainty, *JAPCA*, 1988, **38(4)**, 406-412.
- Jeong S. J., CALPUFF and AERMOD Dispersion Models for Estimating Odor Emissions From Industrial Complex Area Sources, *Asian Journal of Atmospheric Environment*, 2011, **5(1)**, 1-7.
- Moriasi D. N., Arnold J. G., Van Liew M. W., Bingner R. L., Harmel R. D., Veith T. L., Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations, *Transactions of the ASABE*, 2007, **50(3)**, 885-900.

Kaynaklar

- Paschalidou A. K., Kassomenos P. A., On the Evaluation of Box Model Results: the Case of BOXURB Model, *Environ Monit Assess*, 2009, **155**, 103-118.
- Righi S., Luciali P., Pollini E., Statistical and Diagnostic Evaluation of the ADMS-Urban Model Compared with an Urban Air Quality Monitoring Network, *Atmospheric Environment*, 2009, **43**, 3850-3857.
- Sokhi R. S., San Jose R., Kitwiroon N., Fragkou E., Perez J. L., Middleton D. R., Prediction of Ozone Levels in London Using the MM5–CMAQ Modelling System, *Environmental Modelling & Software*, 2006, **21**, 566-576.
- Trapp W., Paciorek M., Paciorek M. K., Juda-Rezler K., Warchałowski A., Reizer M., Modeling of PM₁₀ and PM_{2.5} Particulate Matter Air Pollution in Poland, Editors: Pawlowski L., Dudzinska M. R., Pawlowski A., *Environmental Engineering III*, Taylor & Francis Group, London, 97-104, 2010.
- Yasin M. F., Al-Awadhi M. M., Impact of Sulfur Dioxide Emissions of Power Stations on Ambient Air Quality, *Environmental Engineering Science*, 2011, **28(7)**, 469-475.