



ÇEV 715
Hava Kirliliği Modellemesi ve Uygulamaları

İstatistiksel Modeller

Özgür ZEYDAN (PhD.)
<http://www.ozgurzeydan.com/>

İstatistiksel Hava Kalitesi Modelleri

- Kirletici kaynakları ve emisyon envanterleri yetersiz olduğunda istatistiksel modellerden faydalanılabilir.
- Kısa süreli hava kalitesi tahmini için kullanılabilir (örneğin alarm durumları).
- Kolay elde edilen parametreler (ör. meteorolojik parametreler gibi) yardımı zor ölçülen parametreler (ör. kirletici konsantrasyonları) istatistiksel yöntemler (ör. Regresyon modeli) tahmin edilebilir.

Hava Kirliliği Modellemede İstatistiksel Yöntemler

- Zaman serileri analizleri
 - Box-Jenkins modelleri
- Reseptör modelleri
 - EPA Chemical Mass Balance Model
 - https://www3.epa.gov/scram001/receptor_cmb.htm
 - EPA Positive Matrix Factorization (PMF)
 - <https://www.epa.gov/air-research/positive-matrix-factorization-model-environmental-data-analyses>
- Regresyon Modelleri

Box-Jenkins Modelleri

- Doğrusal durağan stokastik modeller
 - Otoregresif Model (AR)
 - Hareketli Ortalama Modeli (MA)
 - Otoregresif Hareketli Ortalama Modeli (ARMA)
- Durağan olmayan doğrusal stokastik modeller
 - ARIMA
- Mevsimlik modeller



Available online at www.alphanumericjournal.com

alphanumeric journal

The Journal of Operations Research, Statistics, Econometrics and Management Information Systems

Volume 3, Issue 2, 2015



2015.03.02.STAT.07

TIME SERIES ANALYSIS AND FORECASTING FOR AIR
POLLUTION IN ANKARA: A BOX-JENKINS APPROACH

Duygu TURGUT *
İzzetin TEMİZ †

<http://dx.doi.org/10.17093/aj.2015.3.2.5000148347>

Regresyon Modeli

$$Y = \beta_0 + \beta_1 X_1 + \varepsilon \quad (\text{Tek deęişkenli})$$

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \varepsilon \quad (\text{Çok deęişkenli})$$

- Y: baęımlı deęişken (hava kirleticisi konsantrasyonu)
- β_i : regresyon katsayısı
- X_i : baęımsız deęişken
- ε : hata

Regresyon Modeli

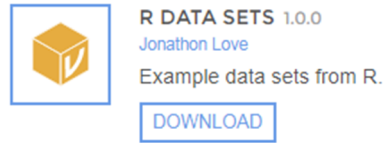
Table 4. Transformation of the nonlinear regression and linear regression.

Types	Nonlinear Function	Do Transformation	Linear Function
Hyperbolic function	$Y = a + b\frac{1}{x}$	$x' = \frac{1}{x}$	$Y' = a + bx'$
Power function	$Y = ax^b$	$Y' = \ln Y, x = \ln x, A = \ln a$	$Y' = A + bx'$
Exponential function	$Y = ae^{bx}$	$Y' = \ln Y, A = \ln a$	$Y' = A + bx'$
	or $Y = ae^{\frac{b}{x}}$	or $Y' = \ln Y, x = \frac{1}{x}, A = \ln a$	or $Y' = A + bx'$
Logarithmic function	$Y = a + b \ln x$	$x' = \ln x$	$Y' = a + bx'$
S curve type	$Y = \frac{1}{a+be^{-x}}$	$Y' = \frac{1}{Y}, x = e^{-x}$	$Y' = a + bx'$
Parabolic type	$Y = a + bx + cx^2$	$x_1 = x, x_2 = x^2$	$Y' = a + bx_1 + cx_2$

(Bai et al., 2018)

Regresyon Modeli – Örnek Uygulama

- Yazılım ve Veri Seti
- Jamovi (<https://www.jamovi.org/>) (ücretsiz)
- R veri seti (<https://www.jamovi.org/library.html>)



Örnek Hava Kalitesi Verisinin Açılması

← jamovi

↑ Data Library > r-datasets

1 Open

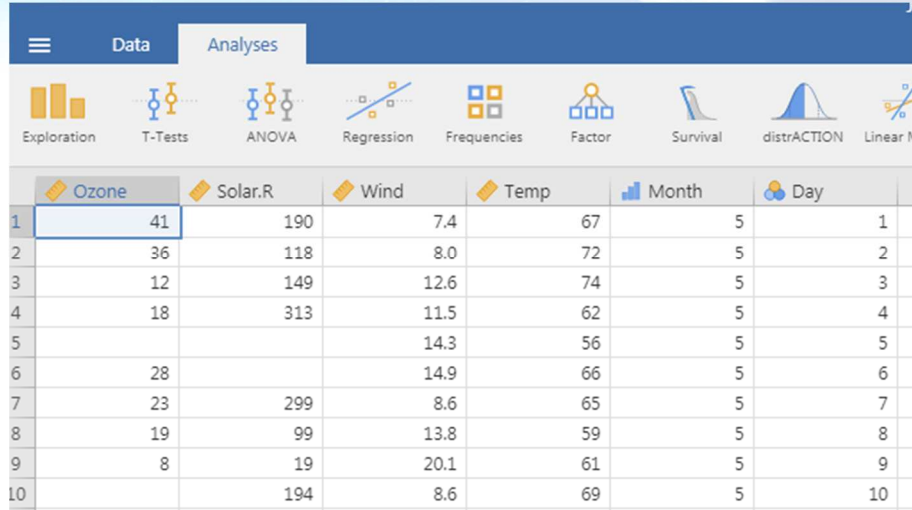
2 Data Library

3

4

Dataset Name	Description	Correlation	Linear regression	Descriptives
airquality	New York Air Quality Measurements	Correlation	Linear regression	
anscombe	Anscombe's Quartet of 'Identical' Simple Linear Regressions			Descriptives
attitude	The Chatterjee-Price Attitude Data	Correlation	Linear regression	
cars	Speed and Stopping Distances of Cars	Correlation	Linear regression	
ChickWeight	Weight versus age of chicks on different diets			

Örnek Hava Kalitesi Verisi

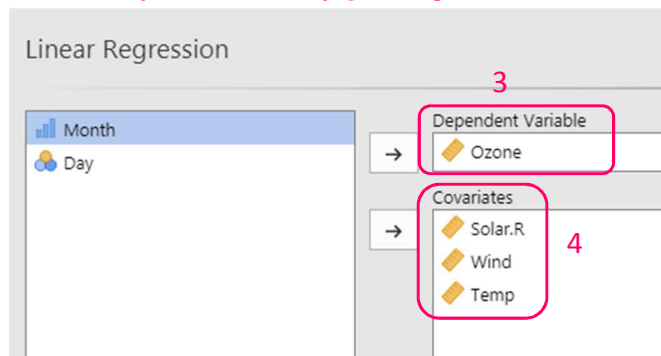
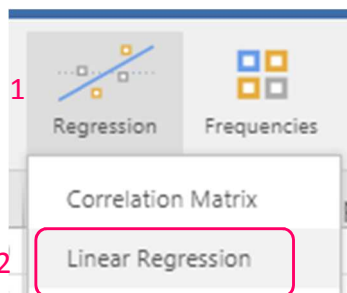


	Ozone	Solar.R	Wind	Temp	Month	Day
1	41	190	7.4	67	5	1
2	36	118	8.0	72	5	2
3	12	149	12.6	74	5	3
4	18	313	11.5	62	5	4
5			14.3	56	5	5
6	28		14.9	66	5	6
7	23	299	8.6	65	5	7
8	19	99	13.8	59	5	8
9	8	19	20.1	61	5	9
10		194	8.6	69	5	10

Çok Değişkenli Doğrusal Regresyon

- Bağımlı değişken: Ozon konsantrasyonu
- Bağımsız değişkenler: Solar radyasyon, rüzgar ve sıcaklık

$$Ozon = \beta_0 + \beta_1 Solar.R + \beta_2 Wind + \beta_3 Temp$$



Test Parametrelerinin Seçilmesi

Linear Regression

> Model Builder

> Reference Levels

Assumption Checks

Assumption Checks

Autocorrelation test

Collinearity statistics

Q-Q plot of residuals

Residual plots

Data Summary

Cook's distance

Model Fit

Fit Measures

R

R²

Adjusted R²

AIC

BIC

RMSE

Overall Model Test

F test

Model Sonuçlarının Değerlendirilmesi

Linear Regression

Model, bağımsız değişkenler ile bağımlı değişkeni %60 oranında açıklayabiliyor.

Model Fit Measures

Model	R	R ²	Adjusted R ²	RMSE	Overall Model Test			
					F	df1	df2	p
1	0.778	0.606	0.595	20.8	54.8	3	107	<.001

Sabit ve Katsayılar

Model Coefficients

Predictor	Estimate	SE	t	p
Intercept	-64.3421	23.0547	-2.79	0.006
Solar.R	0.0598	0.0232	2.58	0.011
Wind	-3.3336	0.6544	-5.09	<.001
Temp	1.6521	0.2535	6.52	<.001

Model anlamlı
(p<0.05 olduğu için)

Model parametrelerinin her biri anlamlı
(p<0.05 olduğu için)

$$\text{Ozon} = -64.3421 + 0.0598 \times \text{Solar.R} - 3.3336 \times \text{Wind} + 1.6521 \times \text{Temp}$$

Otokorelasyon ve Kolinearite

Assumption Checks

Durbin-Watson Test for Autocorrelation

Autocorrelation	DW Statistic	p
0.0315	1.94	0.708

[3]

Otokorelasyon için Durbin-Watson değeri (0 ile 4 arasında değişir). Otokorelasyon olmaması için DW değerinin 2'ye yakın olması istenir.

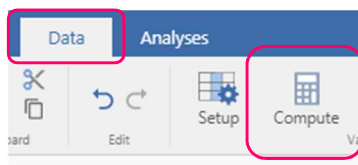
Kolinearite için VIF değerleri. Kolinearite olmaması için VIF değerinin 10'dan küçük olması istenir.

Collinearity Statistics

	VIF	Tolerance
Solar.R	1.10	0.913
Wind	1.33	0.752
Temp	1.43	0.699

Çok Değişkenli Doğrusal Olmayan Regresyon

- Doğrusal olmayan model geliştirmek için her terimin Ln değerini hesaplayalım.



COMPUTED VARIABLE

Ln(Solar.R)

Description

Formula $f_x = \ln(\text{Solar.R})$

Retain unused levels

Month	Day	Ln(Ozon)	Ln(Solar.R)
5	1	3.714	5.247
5	2	3.584	4.771
5	3	2.485	5.004
5	4	2.890	5.746

Çok Değişkenli Doğrusal Olmayan Regresyon

$$\ln(\text{Ozon}) = \beta_0 + \beta_1 \ln(\text{Solar.R}) + \beta_2 \ln(\text{Wind}) + \beta_3 \ln(\text{Temp})$$

Linear Regression

Dependent Variable: Ln(Ozon)

Covariates: Ln(Solar.R), Ln(Wind), Ln(Temp)

Regresyon Modeli Sonuçları

Linear Regression

Model Fit Measures

Model	R	R ²	Overall Model Test					
			Adjusted R ²	RMSE	F	df1	df2	p
1	0.829	0.687	0.679	0.482	78.4	3	107	<.001

Durbin-Watson Test for Autocorrelation

Autocorrelation	DW Statistic	p
0.123	1.74	0.138

Model ve parametreler anlamlı!

Model Coefficients

Predictor	Estimate	SE	t	p
Intercept	-10.559	2.0895	-5.05	<.001
Ln(Solar.R)	0.305	0.0587	5.20	<.001
Ln(Wind)	-0.663	0.1377	-4.81	<.001
Ln(Temp)	3.205	0.4604	6.96	<.001

Collinearity Statistics

	VIF	Tolerance
Ln(Solar.R)	1.18	0.850
Ln(Wind)	1.35	0.741
Ln(Temp)	1.54	0.651

Doğrusal ve Doğrusal Olmayan Model Karşılaştırması

$$Ozon = -64.3421 + 0.0598 \times Solar.R - 3.3336 \times Wind + 1.6521 \times Temp$$

➤ Doğrusal Model: Adj. $R^2 = 0.595$, RMSE = 20.8

$$\ln(Ozon) = -10.559 + 0.305 \times \ln(Solar.R) - 0.663 \times \ln(Wind) + 3.205 \times \ln(Temp)$$

➤ Doğrusal Olmayan Model: Adj. $R^2 = 0.679$, RMSE = 0.482

İleri Okuma...



atmosphere



Review

Statistical Modeling Approaches for PM₁₀ Prediction in Urban Areas; A Review of 21st-Century Studies

Hamid Taheri Shahraini^{1,2,*} and Sahar Sodoudi¹

¹ Institut für Meteorologie, Freie Universität Berlin, Carl-Heinrich-Becker-Weg 6-10, Berlin 12165, Germany; sodoudi@zedat.fu-berlin.de

² Remote Sensing Research Center, Sharif University of Technology, Tehran 1458889694, Iran

* Correspondence: hamid.taheri@met.fu-berlin.de; Tel.: +49-30-8385-4366; Fax: +49-30-8387-1160

Academic Editor: Pasquale Avino

Received: 29 September 2015; Accepted: 12 January 2016; Published: 26 January 2016

<https://doi.org/10.3390/atmos7020015>

Kaynaklar

- Bai, L., Wang, J., Ma, X., Lu, H. (2018). Air Pollution Forecasts: An Overview. *Int. J. Environ. Res. Public Health*, 15, 780.
- Taheri Shahraini, H.; Sodoudi, S. (2016). Statistical Modeling Approaches for PM₁₀ Prediction in Urban Areas; A Review of 21st-Century Studies. *Atmosphere*, 7, 15.
- Turgut, D., Temiz, İ. (2015). Ankara'daki Hava Kirliliği İçin Zaman Serileri Analizi ve Tahmin: Box-Jenkins Yaklaşımı. *Alphanumeric Journal*, 3(2), 131-138.