StatSci.org / Home

OzDASL

Carbon Monoxide from a Freeway

Keywords: multiple regression, daily trend, harmonics.

Description

Hourly carbon monoxide (CO) averages were recorded on summer weekdays at a measurement station in Los Angeles. The station was established by the Environmental Protection Agency as part of a larger study to assess the effectiveness of the catalytic converter. It was located about 25 feet from the San Diego Freeway, which in this particular area is located at 145 degrees north. It was located such that winds from 145 to 325 degress (which in the summer are the prevalent wind directions during the daylight hours) transport the CO emissions from the highway toward the measurement station. Aggregate measurements were recored for each hour of the day 1 to 24.

Hour - hour of the day, from midnight to midnight
CO - average summer weekday CO concentration (parts per million)
TD - average weekday traffic density (traffic count/traffic speed)
WS - average perpendicular wind-speed component,
wind speed x cos(wind direction - 235 degrees)

It would be interesting to have wind speed and direction recorded separately.

Download

Data File (tab-delimited text)

Source

Ledolter, J., and Tiao, G. C. (1979). Statistical methods for ambient air pollutants with special reference to the Los Angeles Catalyst Study (LACS) Data. Environmental Science and Technology, 13, 1233-1240.

Hogg, R. V., and Ledolter, J. (1992). Applied Statistics for Engineers and Physical Scientists, Second Edition. Macmillan, New York. Exercise 1.5-6.

Analysis

- CO depends nearly linearly on traffic intensity.
- The quadratic dependence of CO on wind speed suggests that wind has a diffusion effect as well as a transport effect.
- There are unexplained daily trends, to which one can use sin and cos harmonics to fit a periodic trend.

```
> summary(lm.co,cor=F)

Call: lm(formula = CO ~ Traffic + Wind + Wind^2 + sin((2 * pi)/24 * Hour) +
        cos((2 * pi)/24 * Hour) + sin((4 * pi)/24 * Hour) + cos((4 * pi)/24 * Hour))
Residuals:
     Min      1Q Median     3Q    Max
 -0.3477 -0.1078 0.0059 0.1407 0.3994

Coefficients:
                          Value Std. Error  t value Pr(>|t|)
            (Intercept)   1.7523   0.7793     2.2484   0.0390
                Traffic   0.0179   0.0013    13.7864   0.0000
```

```
                    Wind   0.3178   0.3447     0.9220   0.3702
               I(Wind^2)  -0.0774   0.0268    -2.8851   0.0108
sin((2 * pi)/24 * Hour)  -0.5437   0.9324    -0.5831   0.5680
cos((2 * pi)/24 * Hour)  -0.7300   0.6574    -1.1106   0.2832
sin((4 * pi)/24 * Hour)   0.3664   0.4331     0.8460   0.4100
cos((4 * pi)/24 * Hour)   0.3110   0.1424     2.1838   0.0442

Residual standard error: 0.2298 on 16 degrees of freedom
Multiple R-Squared: 0.9918
F-statistic: 277.5 on 7 and 16 degrees of freedom, the p-value is 1.776e-015

> anova(lm.co)
Analysis of Variance Table

Response: CO

Terms added sequentially (first to last)
                       Df Sum of Sq  Mean Sq  F Value      Pr(F)
Traffic                 1  95.87547 95.87547 1815.139 0.00000000
Wind                    1   2.35721  2.35721   44.627 0.00000528
I(Wind^2)               1   1.40893  1.40893   26.674 0.00009402
sin((2 * pi)/24 * Hour) 1   0.82150  0.82150   15.553 0.00116174
cos((2 * pi)/24 * Hour) 1   0.87639  0.87639   16.592 0.00088470
sin((4 * pi)/24 * Hour) 1   1.01972  1.01972   19.306 0.00045302
cos((4 * pi)/24 * Hour) 1   0.25191  0.25191    4.769 0.04420810
Residuals              16   0.84512  0.05282
```
