

STAT452/652 Solution to HW05

Copyright 2020 William Ruth and Thomas Loughin
Distribution without their permission is illegal and may be prosecuted

Due on Oct 09, 2020

1 Applications

1.1 Question 1

The variables included in the best model of each size and their corresponding BIC values are as follows.

##	Vars	BIC
## 1	Intercept, TW-Ratio	-73.94
## 2	Intercept, Temp, TW-Ratio	-97.48
## 3	Intercept, Solar.R, Temp, TW-Ratio	-100.40
## 4	Intercept, Solar.R, Wind, Temp, TW-CrossProd	-97.92
## 5	Intercept, Solar.R, Wind, Temp, TW-CrossProd, TW-Ratio	-96.15

The best model appears to be the one with four predictors: solar radiation, temperature, temperature/wind speed, and an intercept.

1.2 Question 2

The model that is chosen by the default settings in stepwise contains: Intercept, TW-Ratio, Temp, Solar.R.

1.3 Question 3

The CV MSPE for each fold, as well as the average across all folds, is as follows for the seed 2928893.

##	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8
##	328.0475	403.9859	577.2344	106.9023	178.0752	369.4887	233.2169	634.9388
##	Fold 9	Fold 10	Average					
##	1074.3133	786.7373	469.2940					

Rerunning this 1000 times and summarizing the results, MSPEs answers for other random cv splits should fall within this range almost all the time:

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
##	44.5	225.2	380.5	474.0	580.3	4962.2
##	2.5%	5%	10%	90%	95%	97.5%
##	117.0946	136.2184	161.7774	912.3060	1098.3036	1239.4809

And the average across the 10 folds should fall within this range:

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
##	385.5	441.5	454.5	474.0	467.0	871.5

##	2.5%	5%	10%	90%	95%	97.5%
##	407.8276	411.8332	419.8524	481.5239	766.0876	796.7516