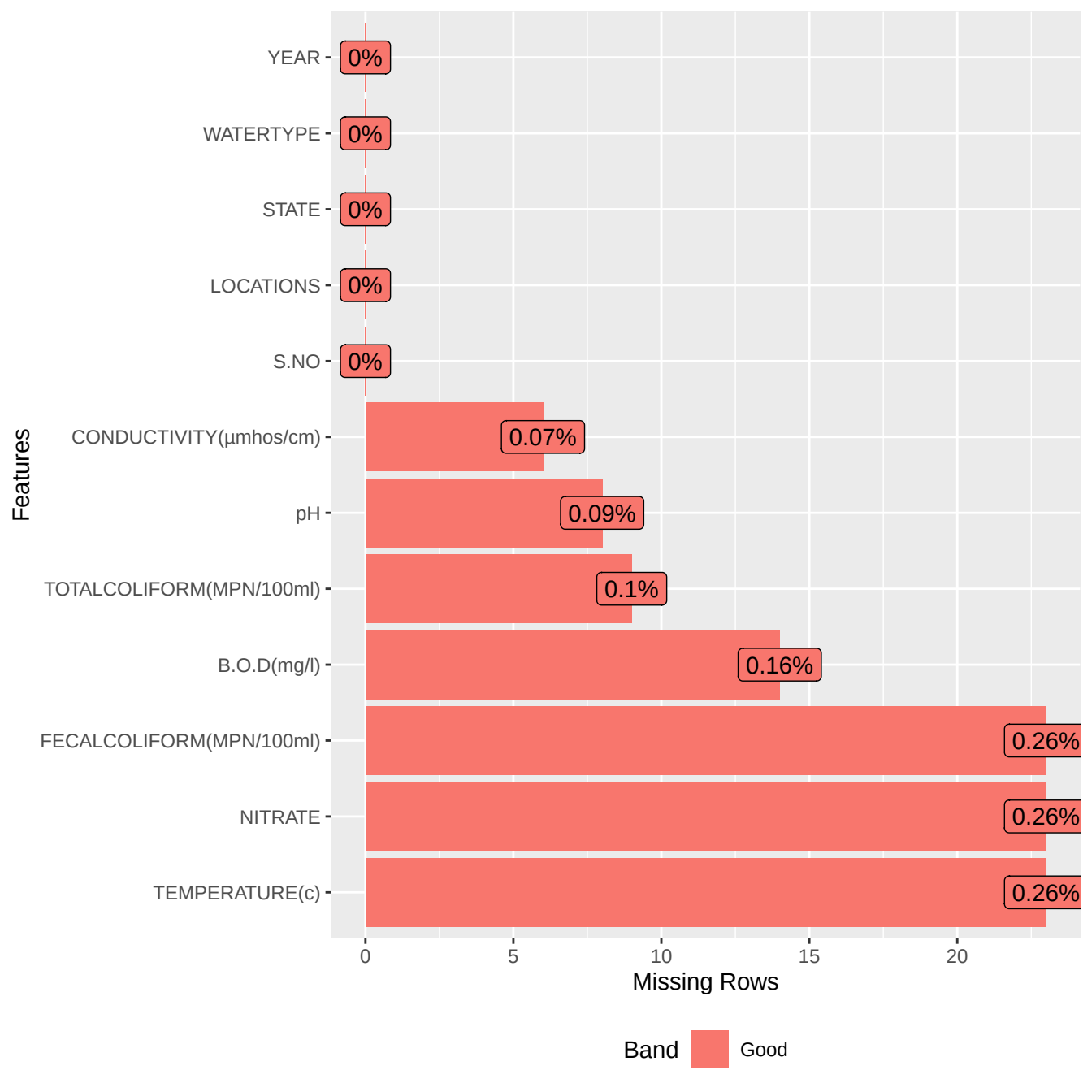
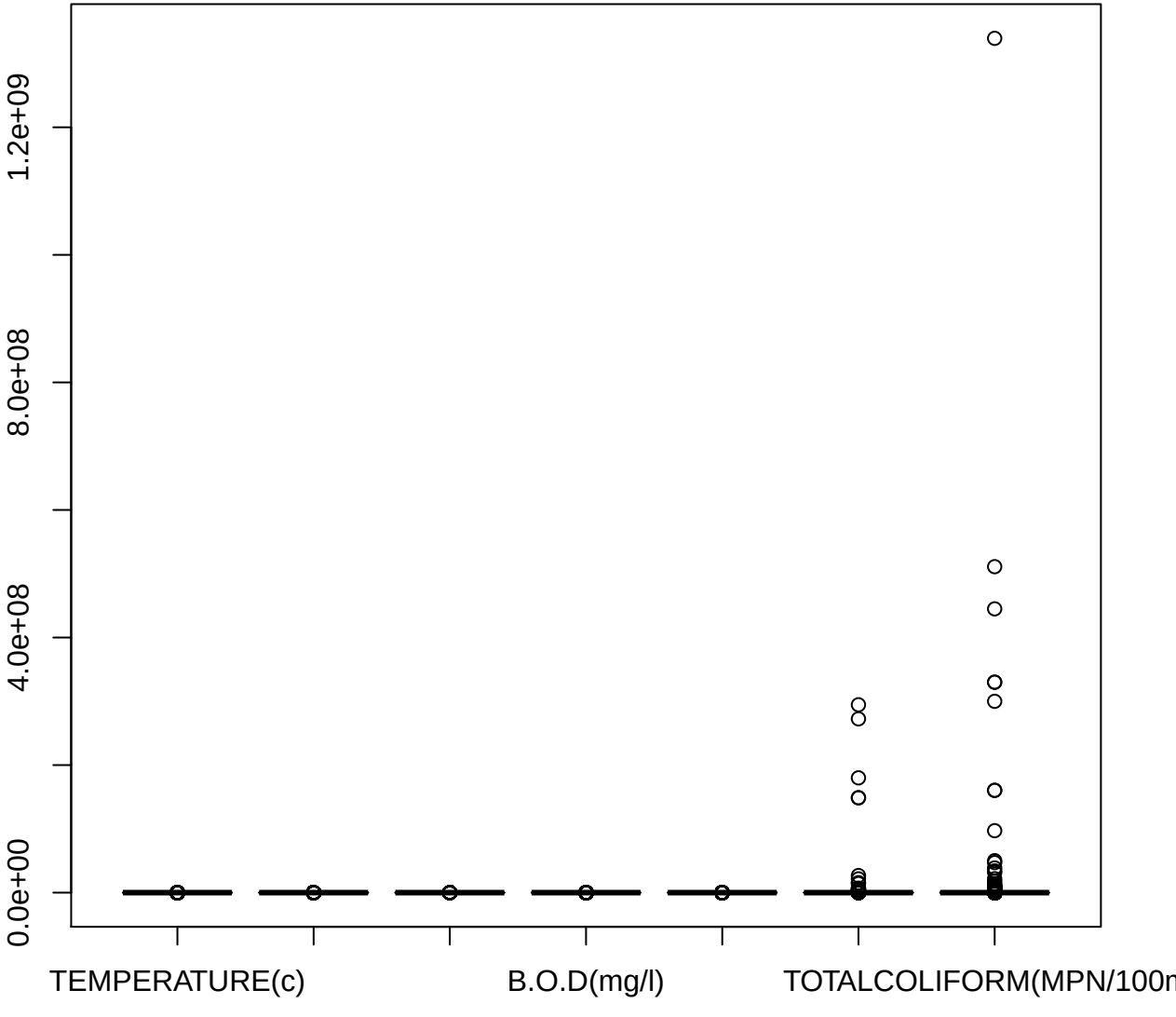
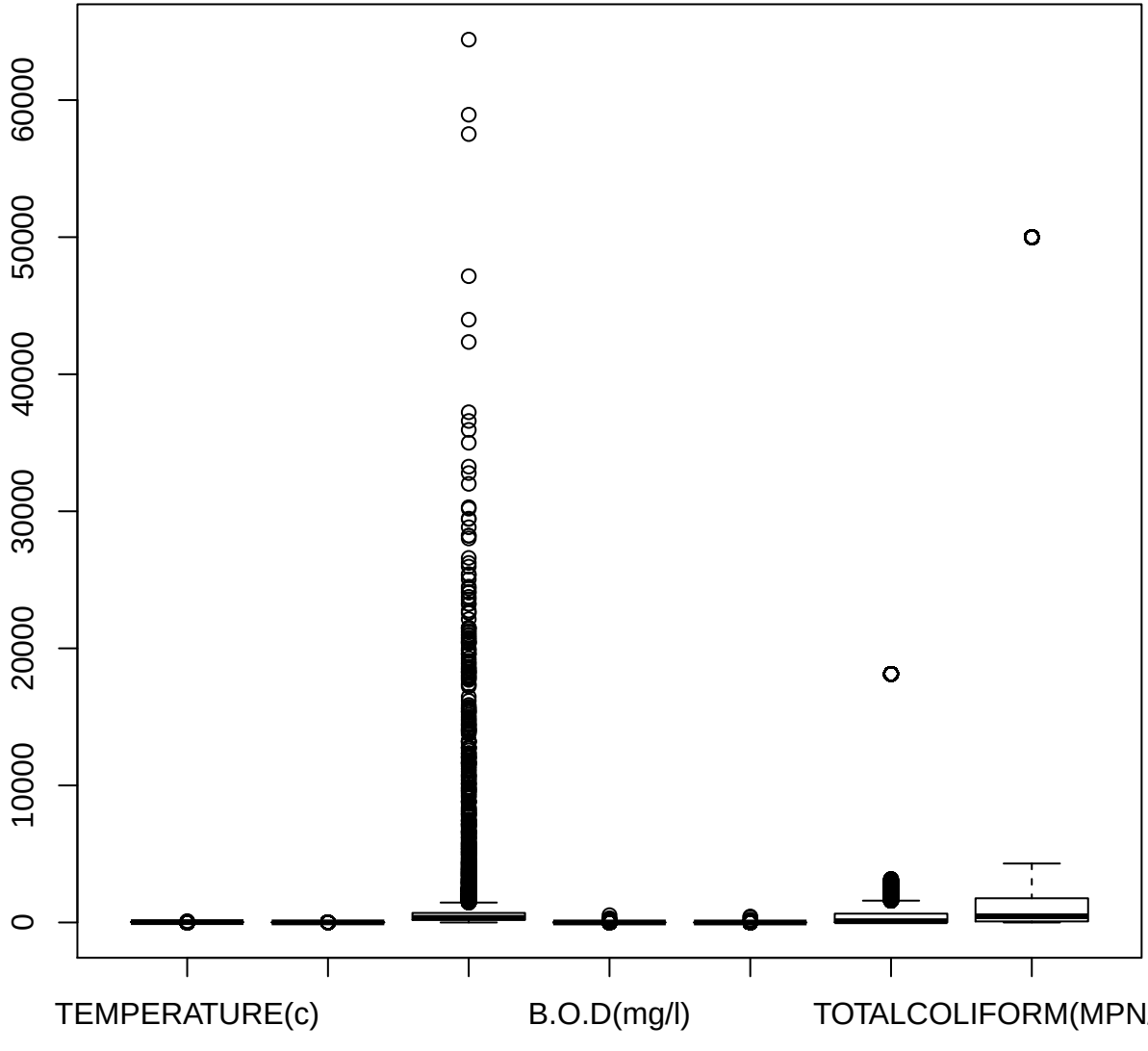
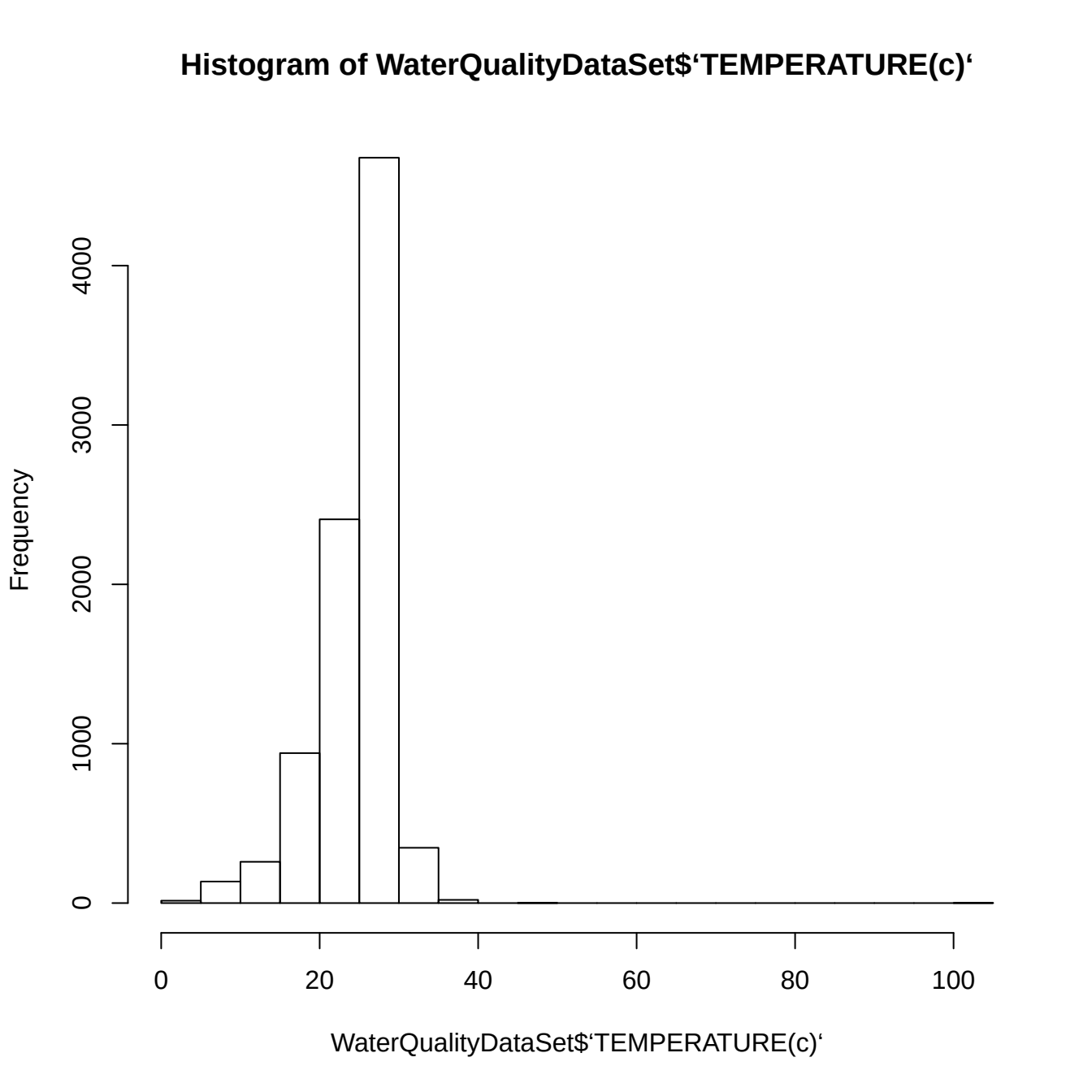
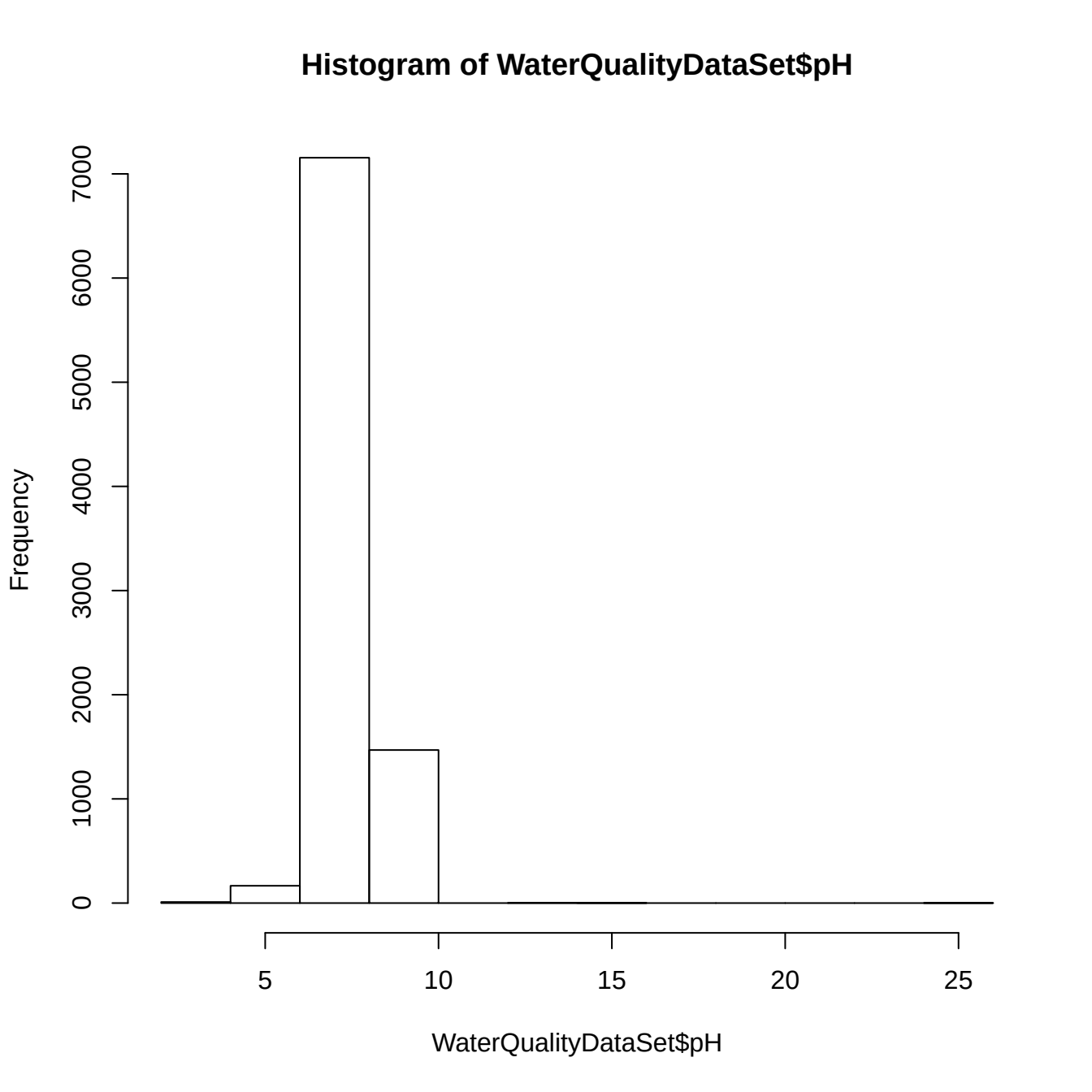
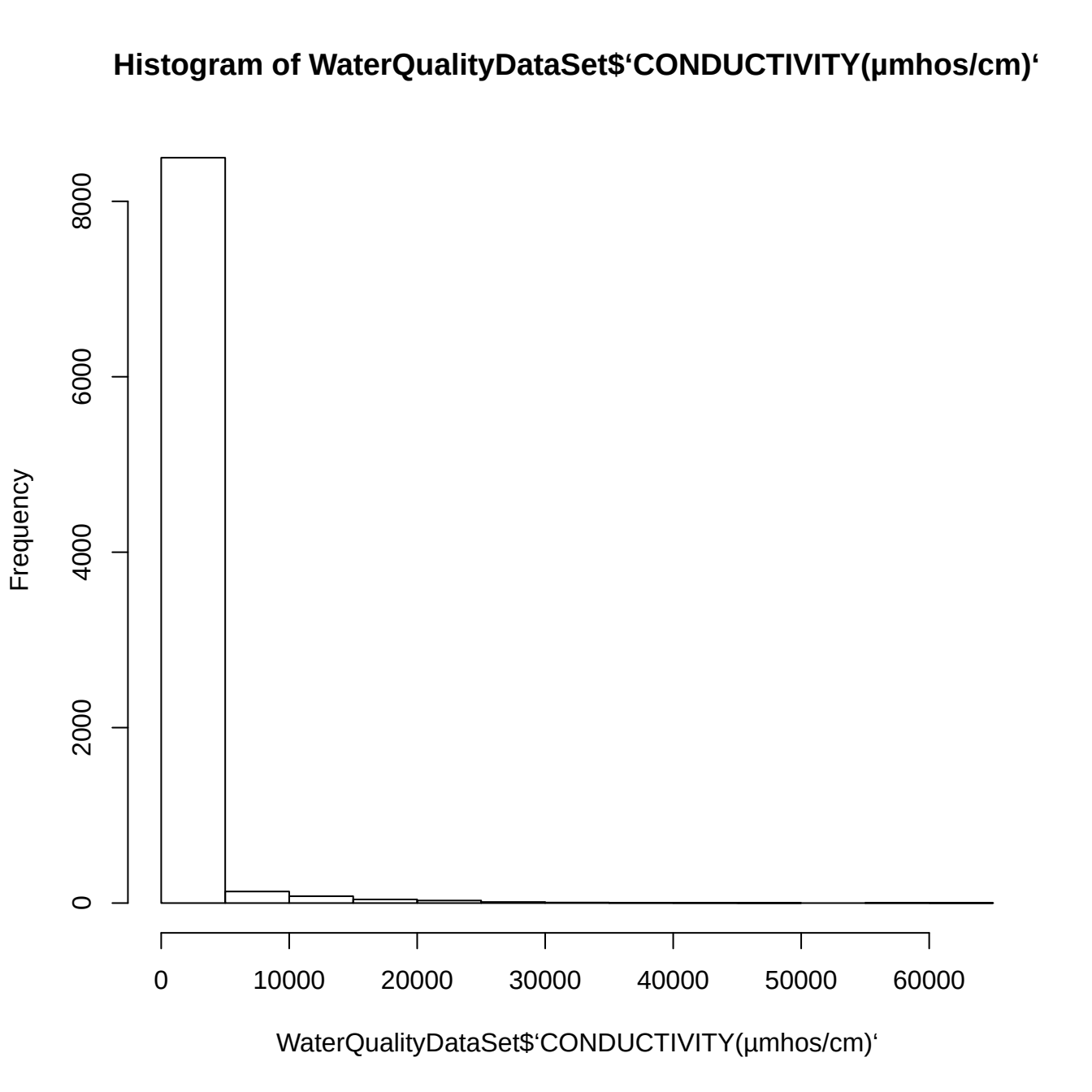

YEAR
WATERTYPE
STATE
LOCATIONS
S.NO
CONDUCTIVITY(µmhos/cm)
pH
TOTALCOLIFORM(MPN/100ml)
B.O.D(mg/l)
FECALCOLIFORM(MPN/100ml)
NITRATE
TEMPERATURE(c)
0%
0%
0%
0%
0%
0.07%
0.09%
0.1%
0.16%
0.26%
0.26%
0.26%
Features
0
5
10
15
20
Missing Rows
Band
Good

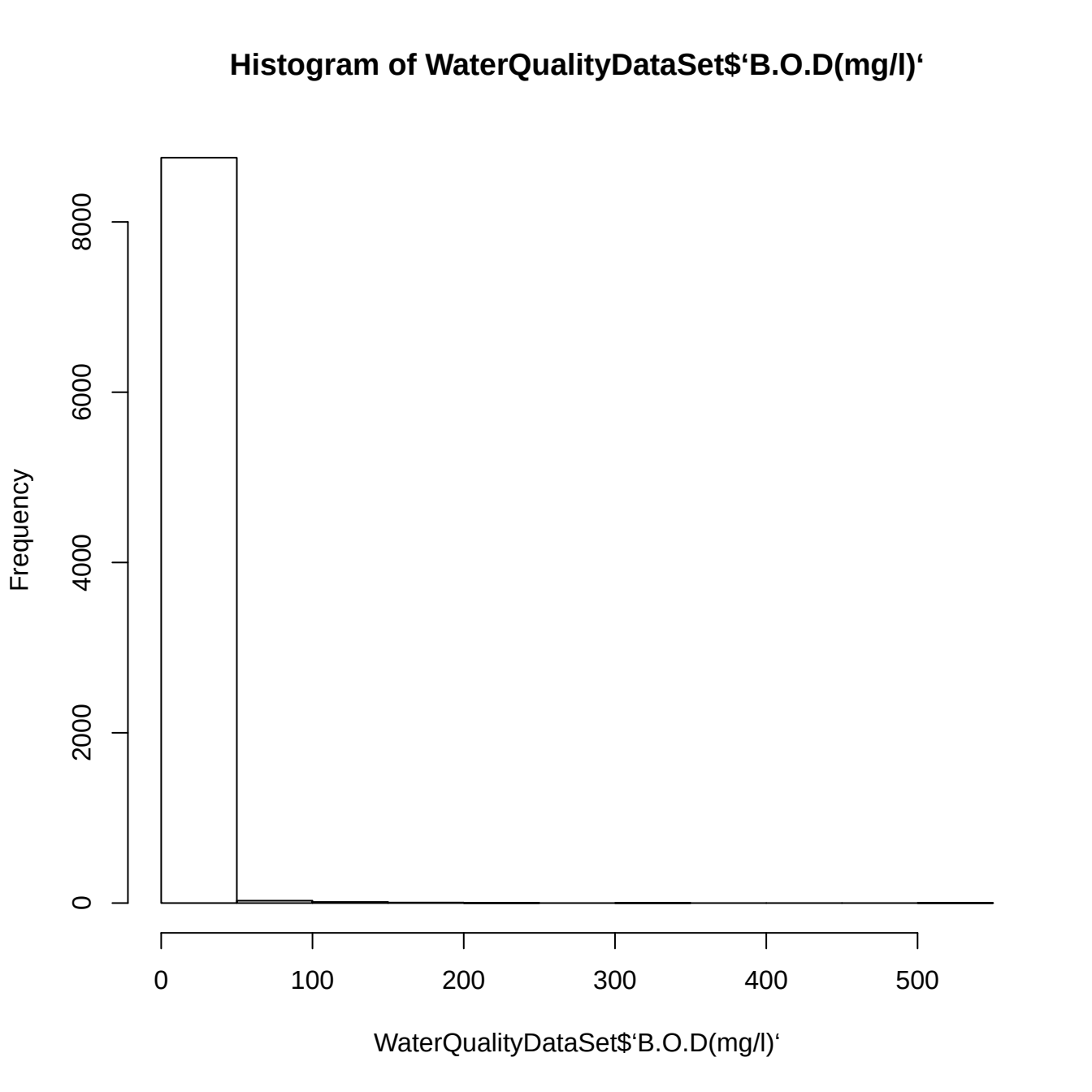

0.0e+00
4.0e+08
8.0e+08
1.2e+09
TEMPERATURE(c)
B.O.D(mg/l)
TOTALCOLIFORM(MPN/100m

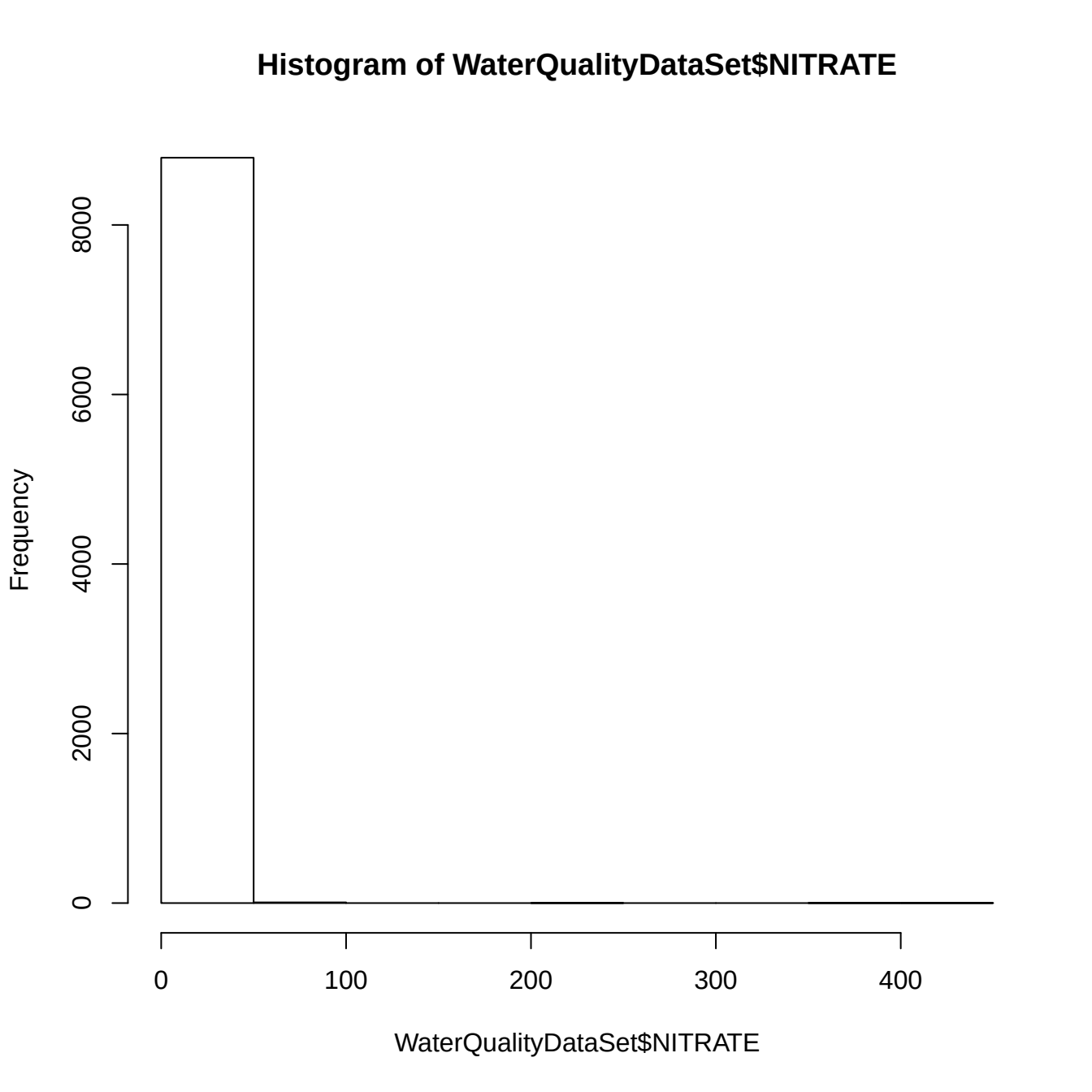

0
10000
20000
30000
40000
50000
60000
TEMPERATURE(c)
B.O.D(mg/l)
TOTALCOLIFORM(MPN/100m

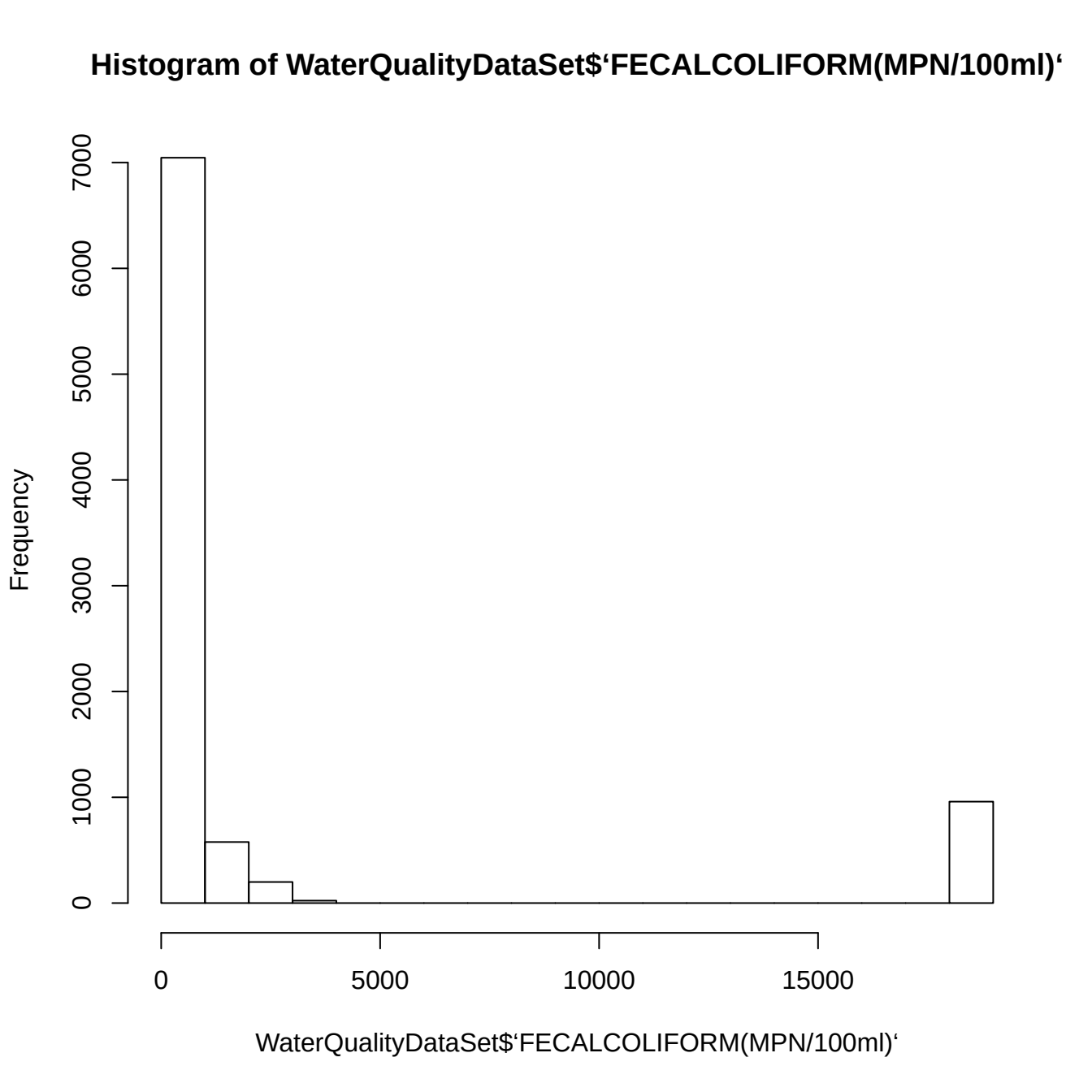

Histogram of WaterQualityDataSet$‘TEMPERATURE(c)‘
Frequency
0
1000
2000
3000
4000
0
20
40
60
80
100
WaterQualityDataSet$‘TEMPERATURE(c)‘

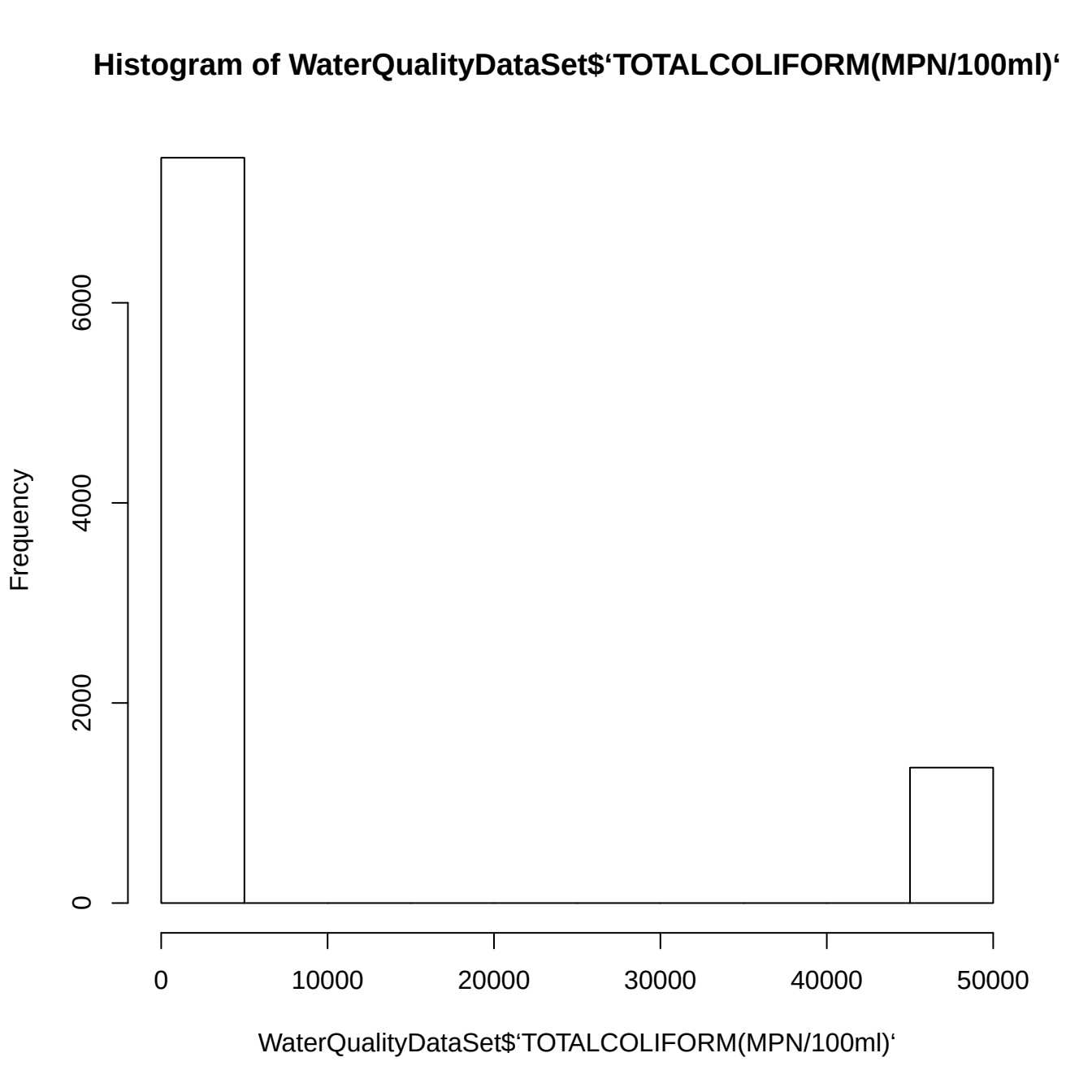

Histogram of WaterQualityDataSet$pH
Frequency
0
1000
2000
3000
4000
5000
6000
7000
5
10
15
20
25
WaterQualityDataSet$pH

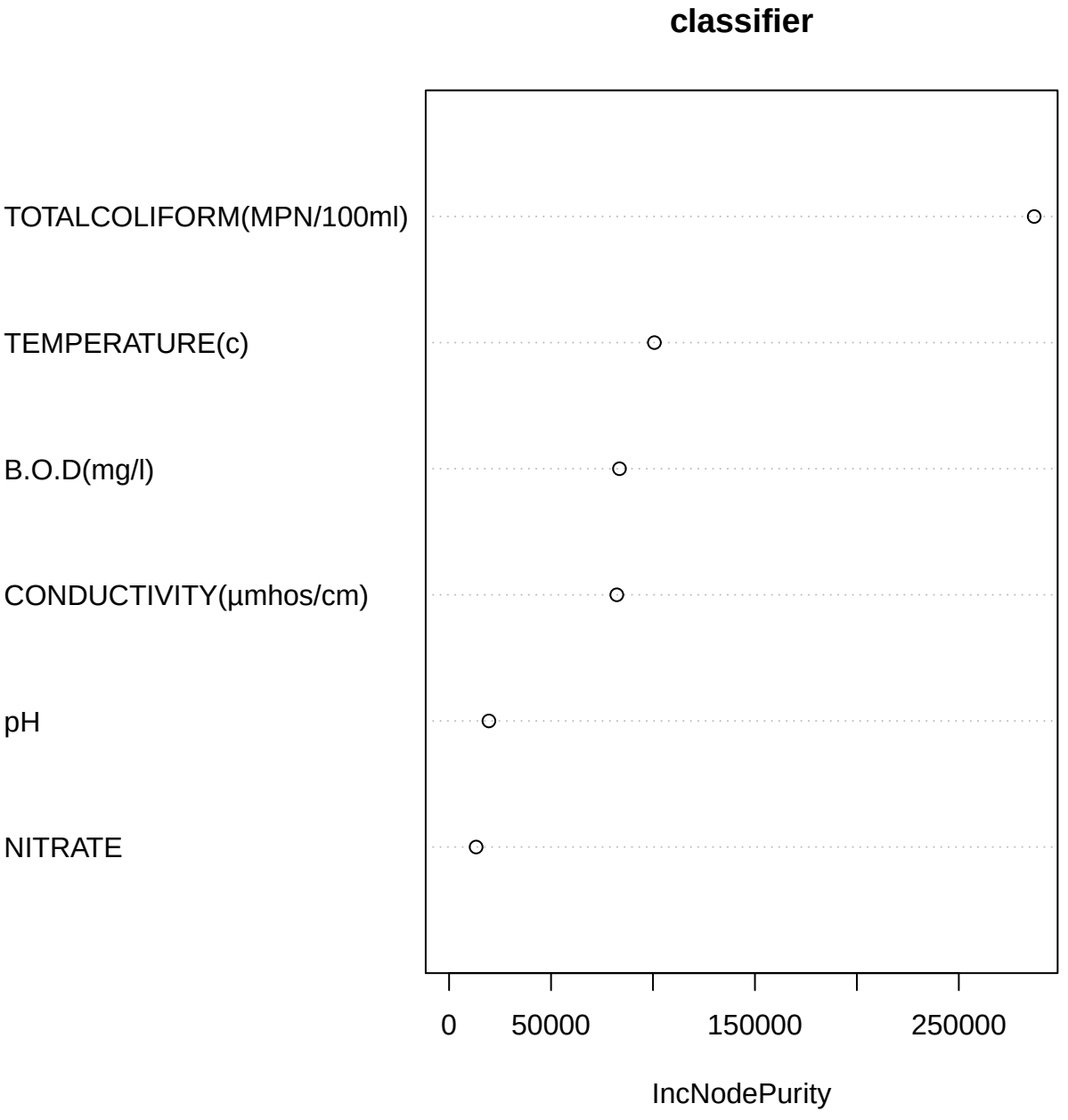

Histogram of WaterQualityDataSet$‘CONDUCTIVITY(µmhos/cm)‘
Frequency
0
2000
4000
6000
8000
0
10000
20000
30000
40000
50000
60000
WaterQualityDataSet$‘CONDUCTIVITY(µmhos/cm)‘

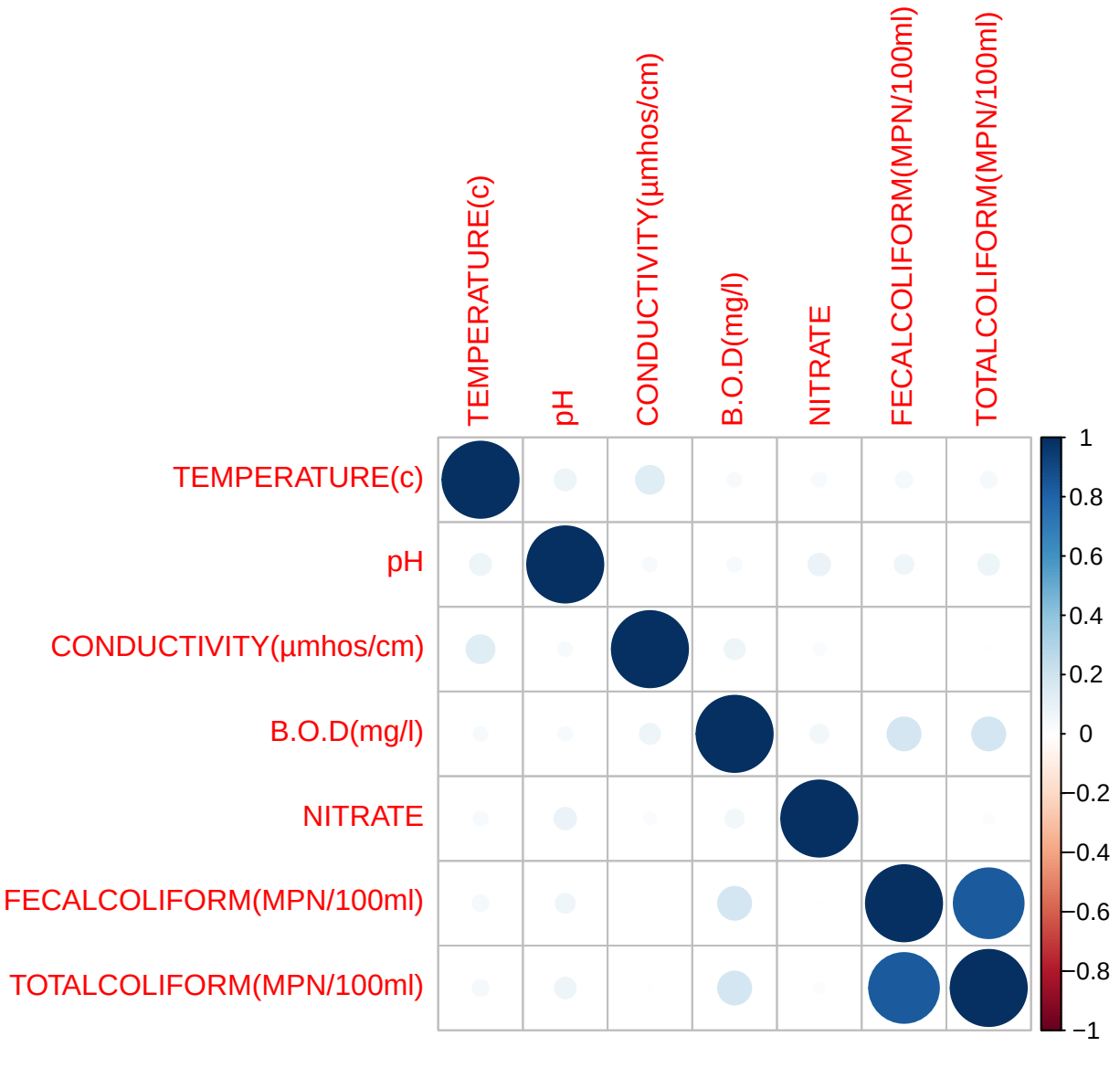

Histogram of WaterQualityDataSet$‘B.O.D(mg/l)‘
Frequency
0
2000
4000
6000
8000
0
100
200
300
400
500
WaterQualityDataSet$‘B.O.D(mg/l)‘

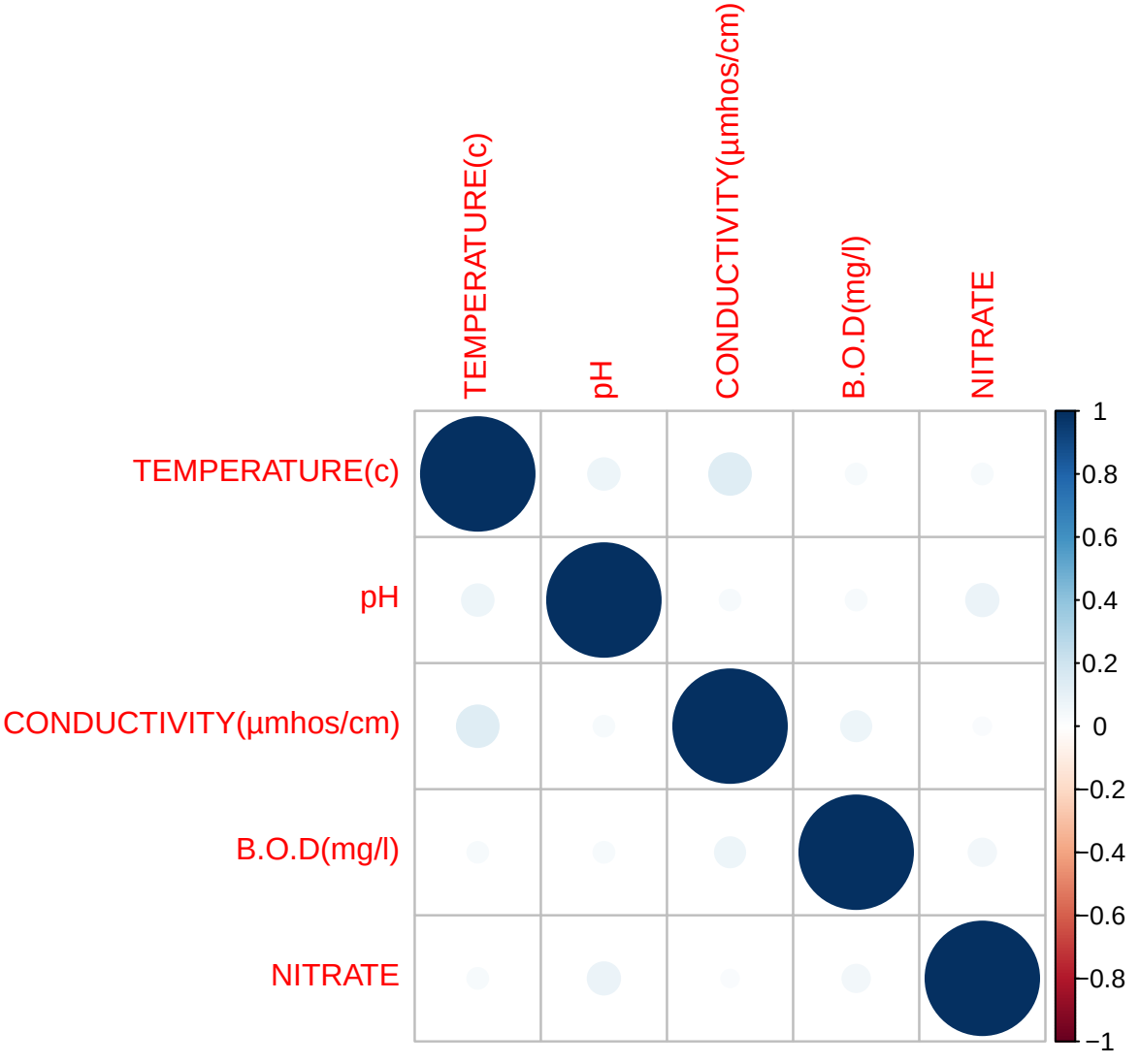

Histogram of WaterQualityDataSet$NITRATE
Frequency
0
2000
4000
6000
8000
0
100
200
300
400
WaterQualityDataSet$NITRATE

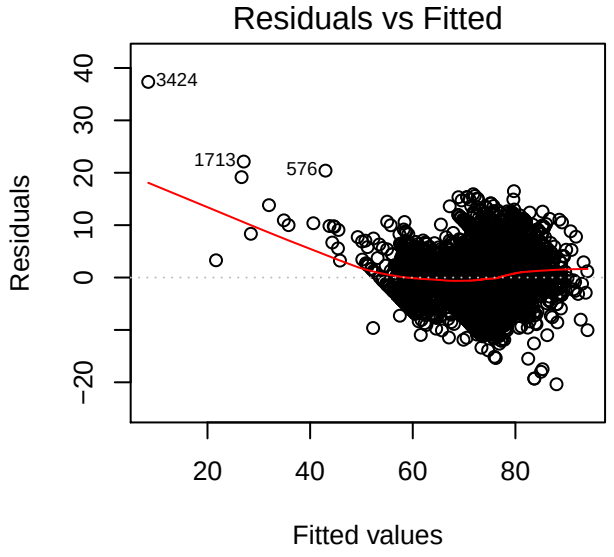

Histogram of WaterQualityDataSet$‘FECALCOLIFORM(MPN/100ml)‘
Frequency
0
1000
2000
3000
4000
5000
6000
7000
0
5000
10000
15000
WaterQualityDataSet$‘FECALCOLIFORM(MPN/100ml)‘

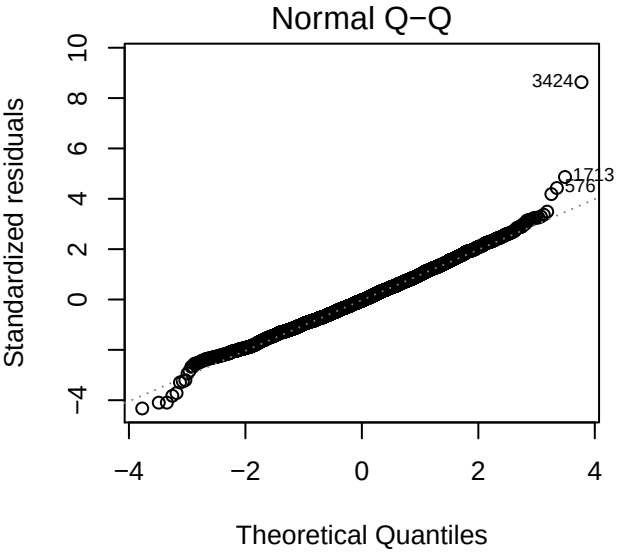

Histogram of WaterQualityDataSet$‘TOTALCOLIFORM(MPN/100ml)‘
Frequency
0
2000
4000
6000
0
10000
20000
30000
40000
50000
WaterQualityDataSet$‘TOTALCOLIFORM(MPN/100ml)‘

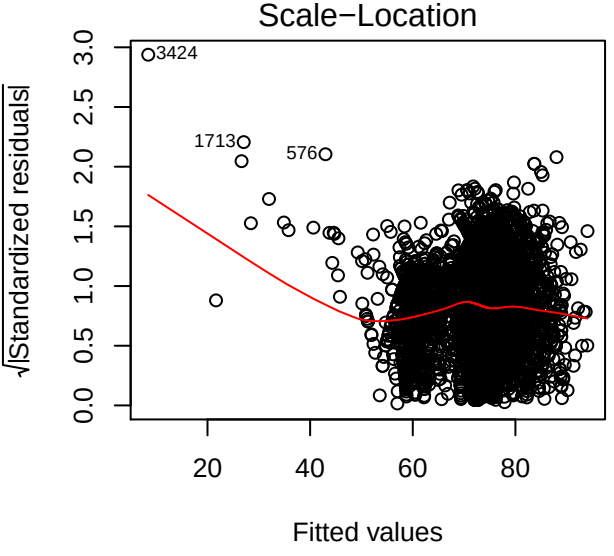

classifier
TOTALCOLIFORM(MPN/100ml)
TEMPERATURE(c)
B.O.D(mg/l)
CONDUCTIVITY(µmhos/cm)
pH
NITRATE
0
50000
150000
250000
IncNodePurity

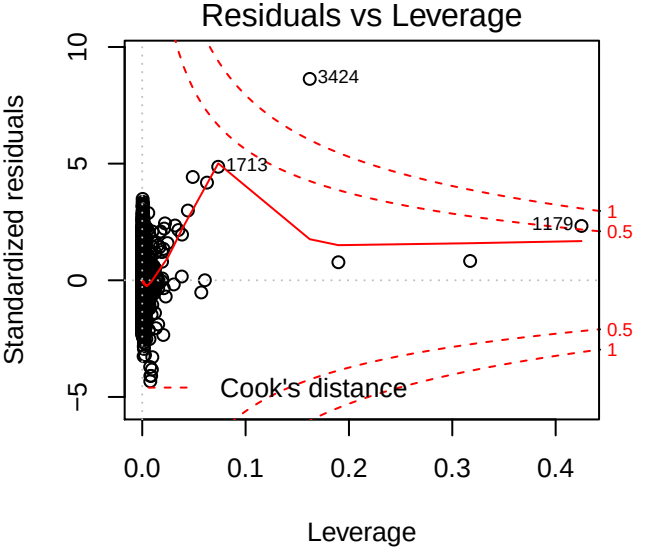

TEMPERATURE(c)
pH
CONDUCTIVITY(µmhos/cm)
B.O.D(mg/l)
NITRATE
FECALCOLIFORM(MPN/100ml)
TOTALCOLIFORM(MPN/100ml)
1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−1

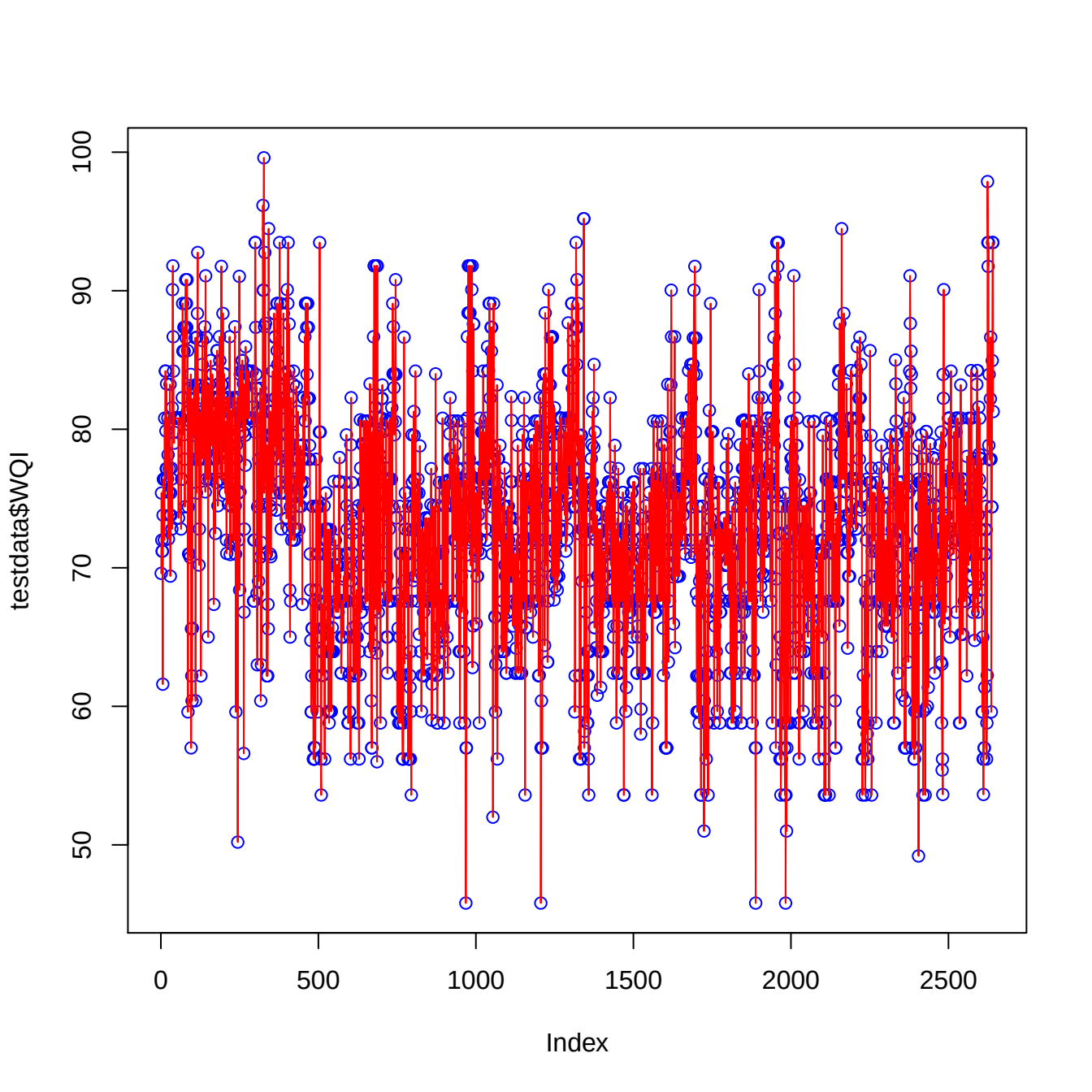

TEMPERATURE(c)
pH
CONDUCTIVITY(µmhos/cm)
B.O.D(mg/l)
NITRATE
TEMPERATURE(c)
pH
CONDUCTIVITY(µmhos/cm)
B.O.D(mg/l)
NITRATE
1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−1

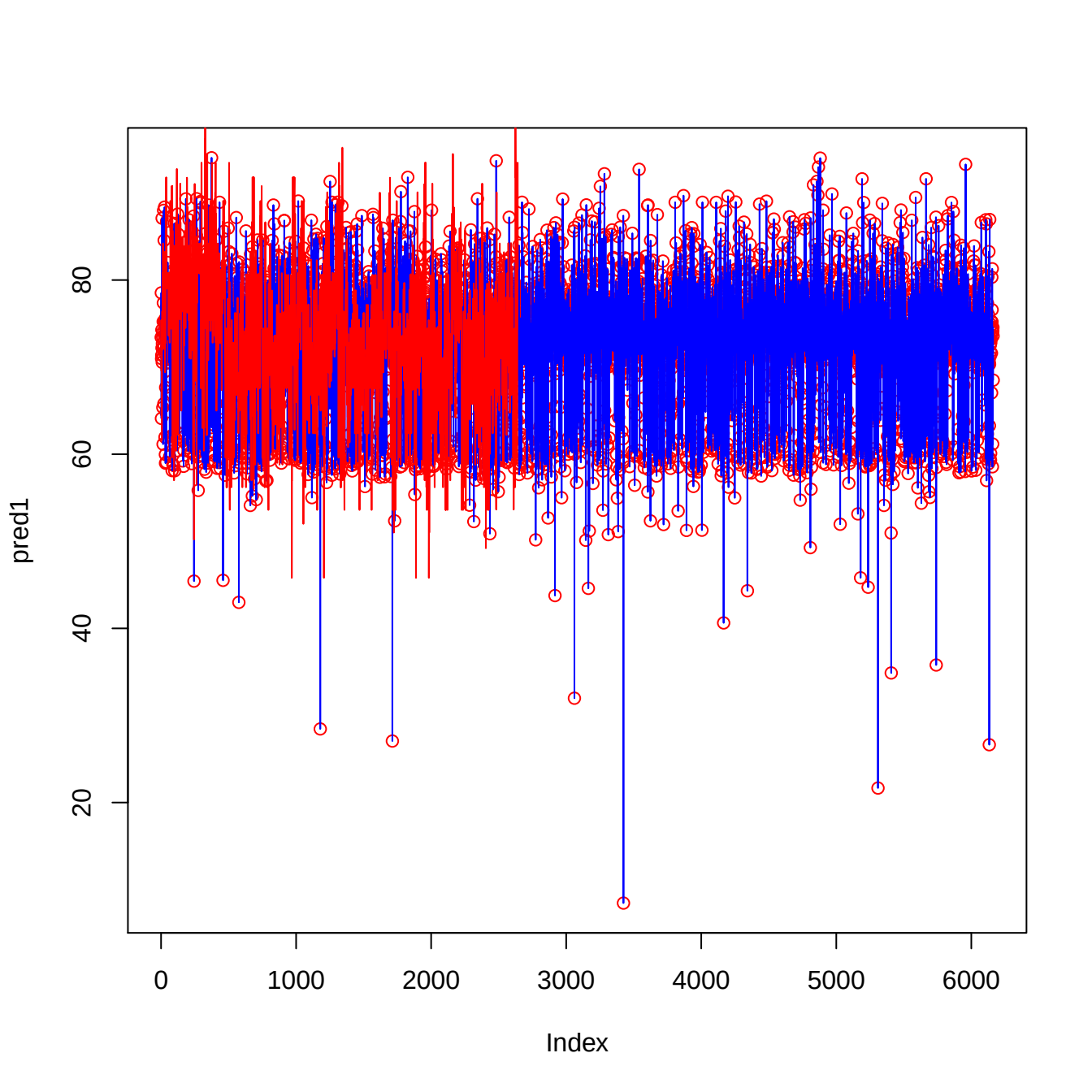

Residuals vs Fitted
Residuals
Fitted values
3424
1713
576
40
30
20
10
0
−20
20
40
60
80

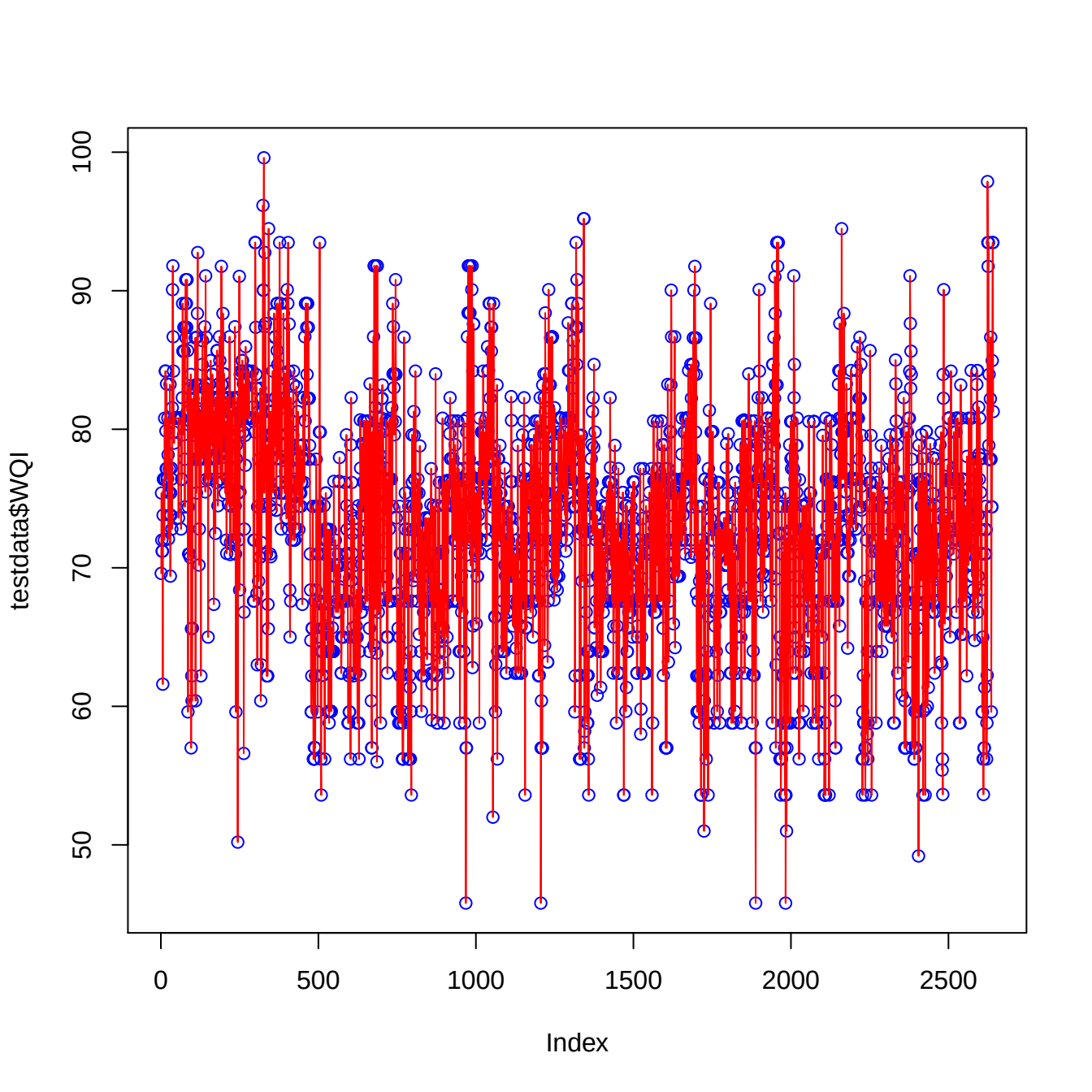

Normal Q−Q
Standardized residuals
Theoretical Quantiles
3424
1713
576
10
8
6
4
2
0
−4
−4
−2
0
2
4

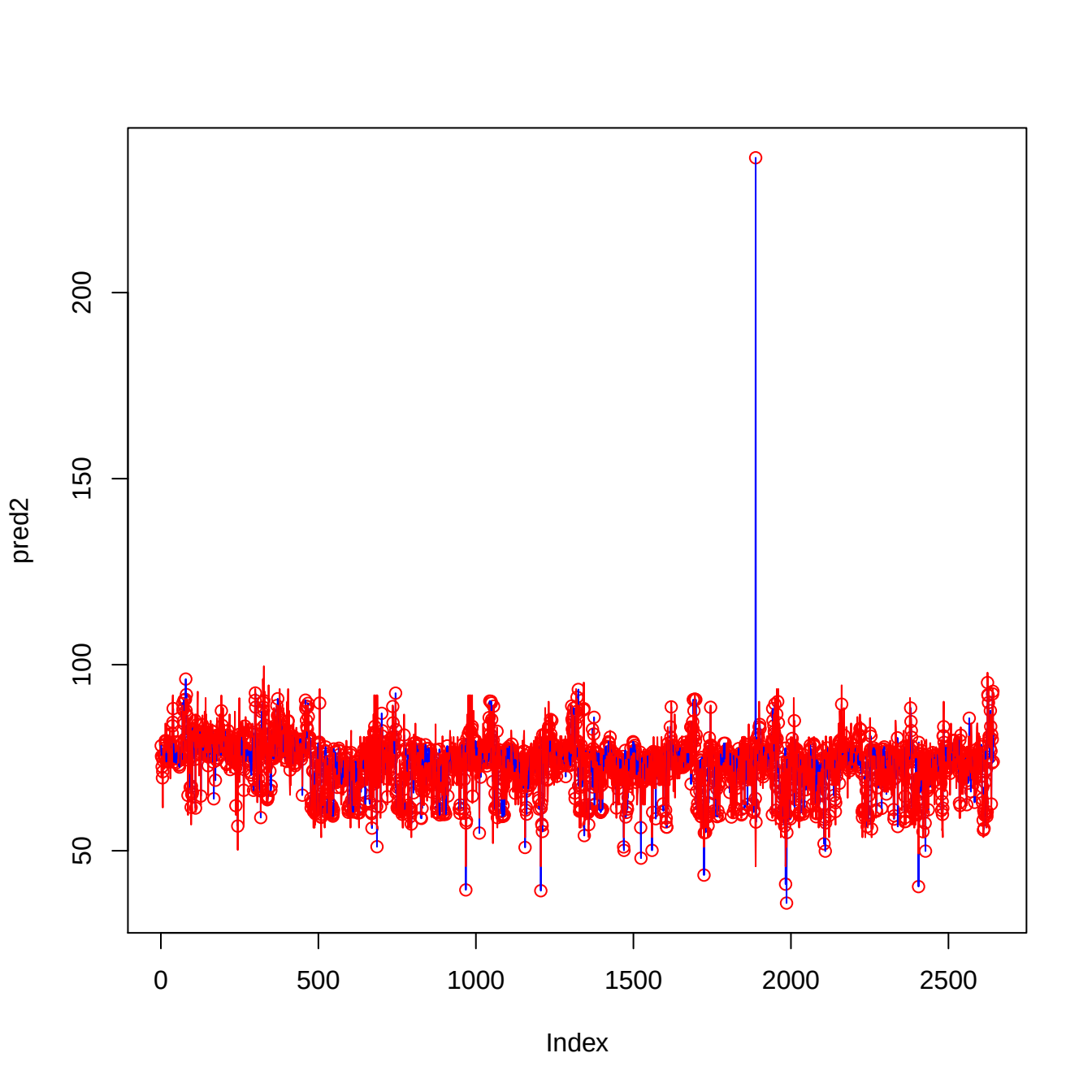

Scale−Location
√|Standardized residuals|
Fitted values
3424
1713
576
3.0
2.5
2.0
1.5
1.0
0.5
0.0
20
40
60
80

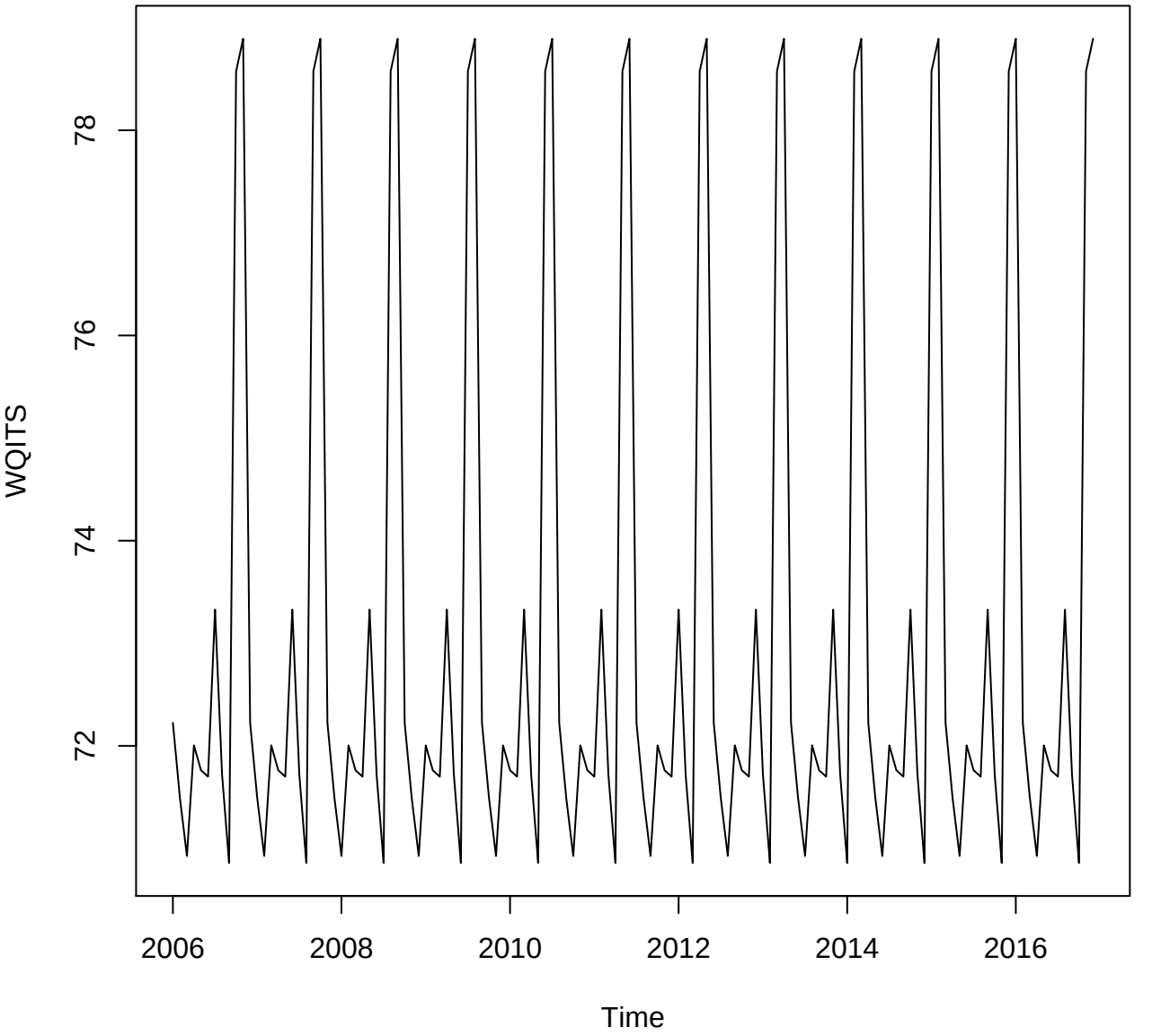

Residuals vs Leverage
Standardized residuals
Leverage
Cook's distance
3424
1713
1179
1
0.5
0.5
1
10
5
0
−5
0.0
0.1
0.2
0.3
0.4

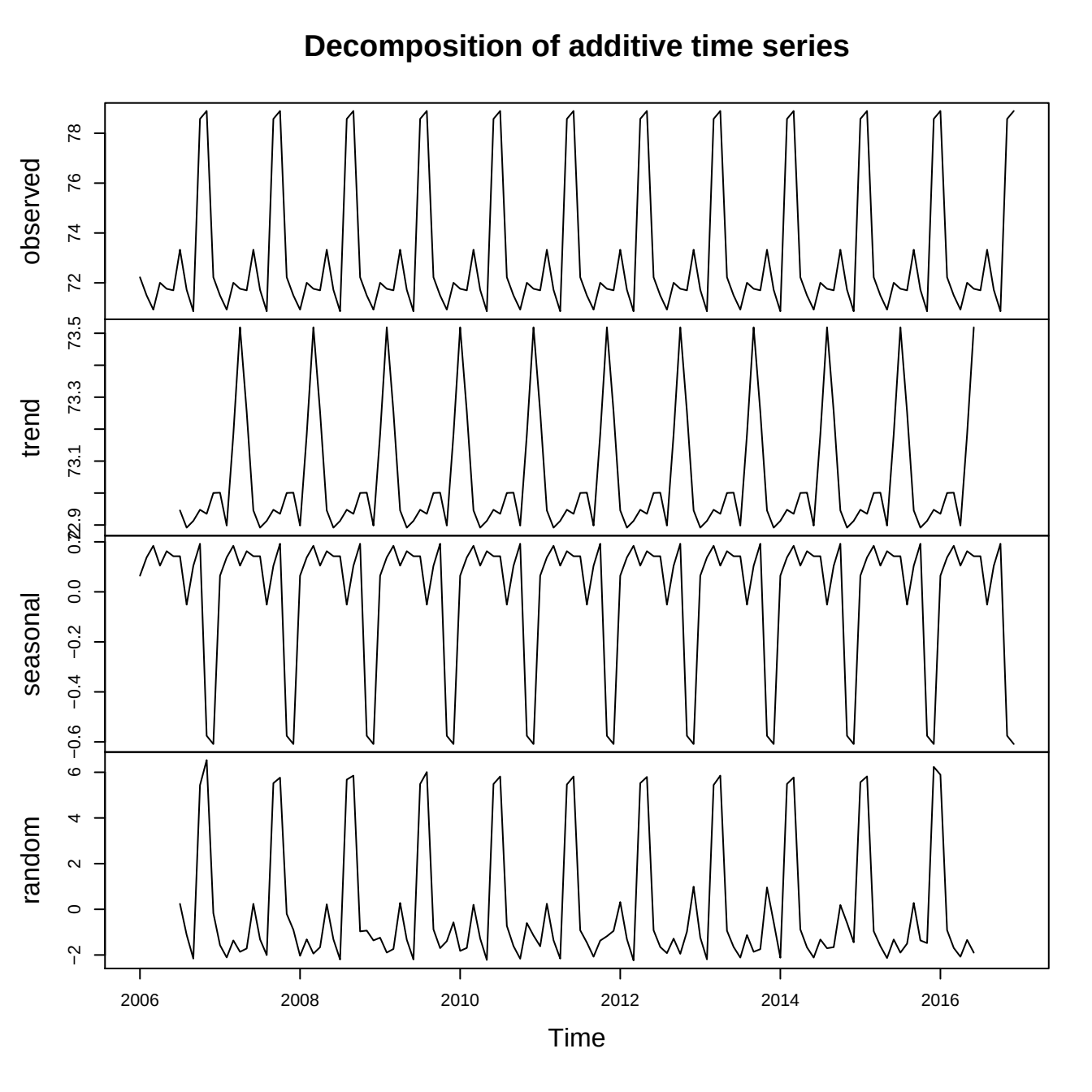

testdata$WQI
Index
50
60
70
80
90
100
0
500
1000
1500
2000
2500

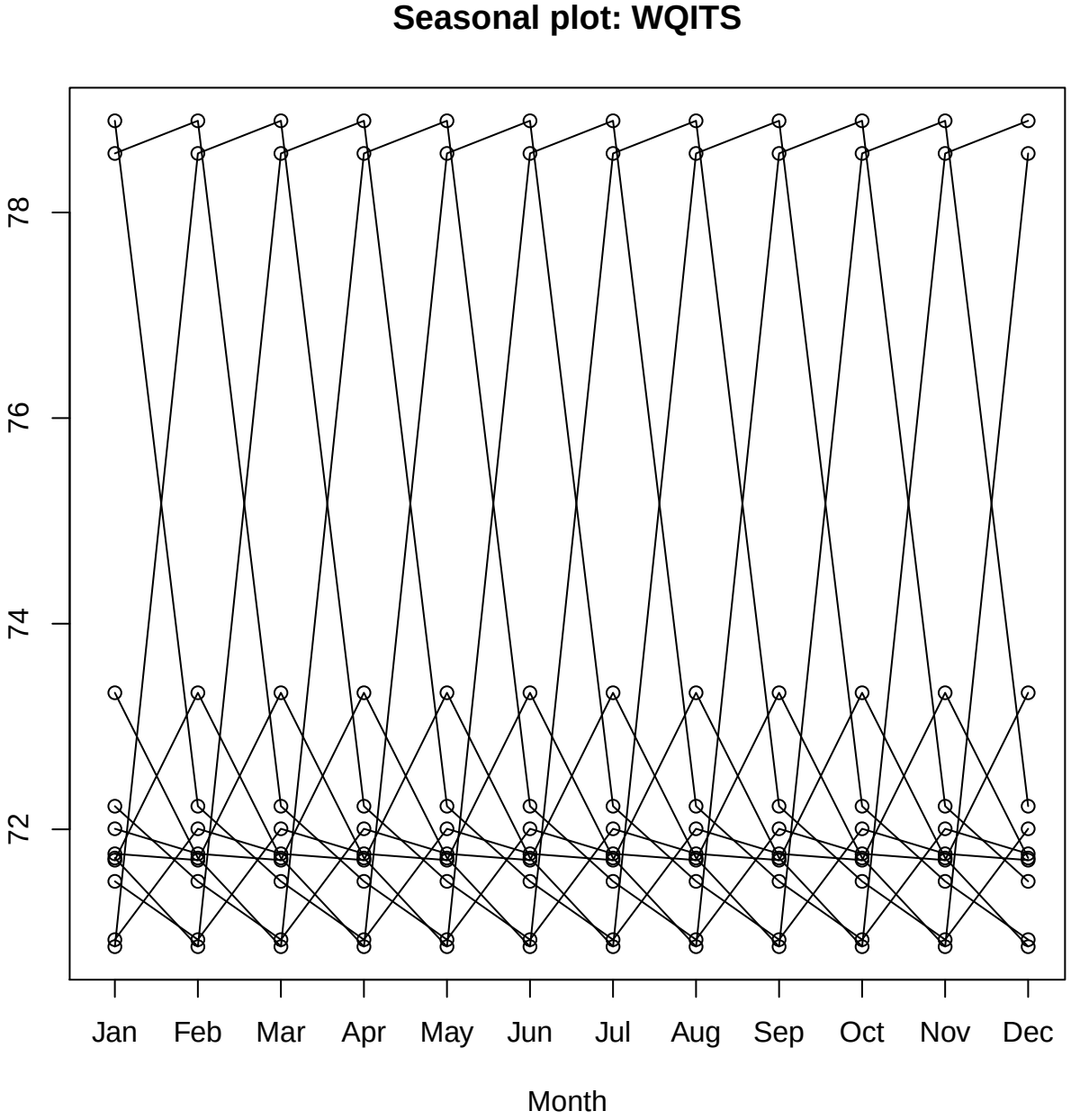

pred1
20
40
60
80
0
1000
2000
3000
4000
5000
6000
Index

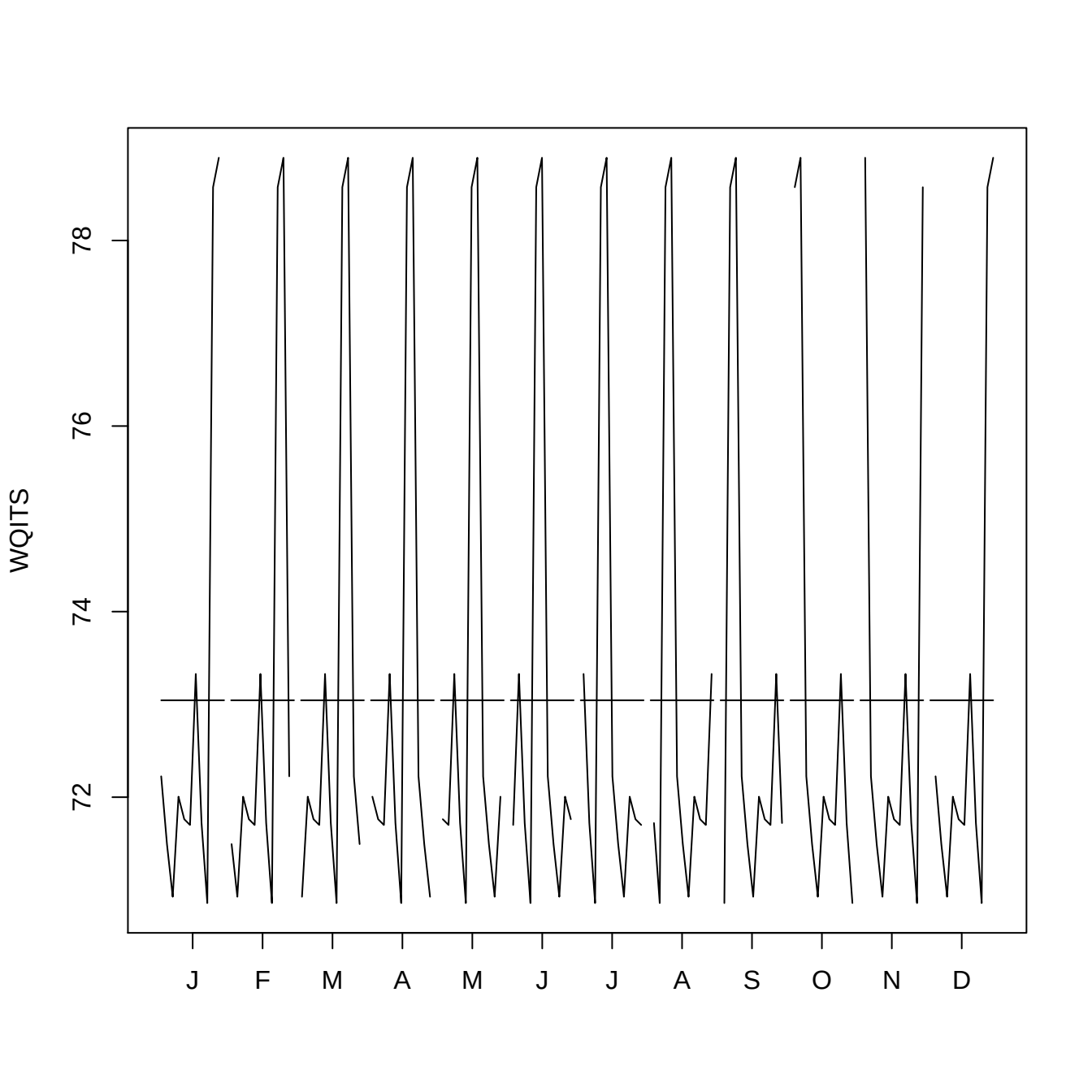

testdata$WQI
Index
50
60
70
80
90
100
0
500
1000
1500
2000
2500

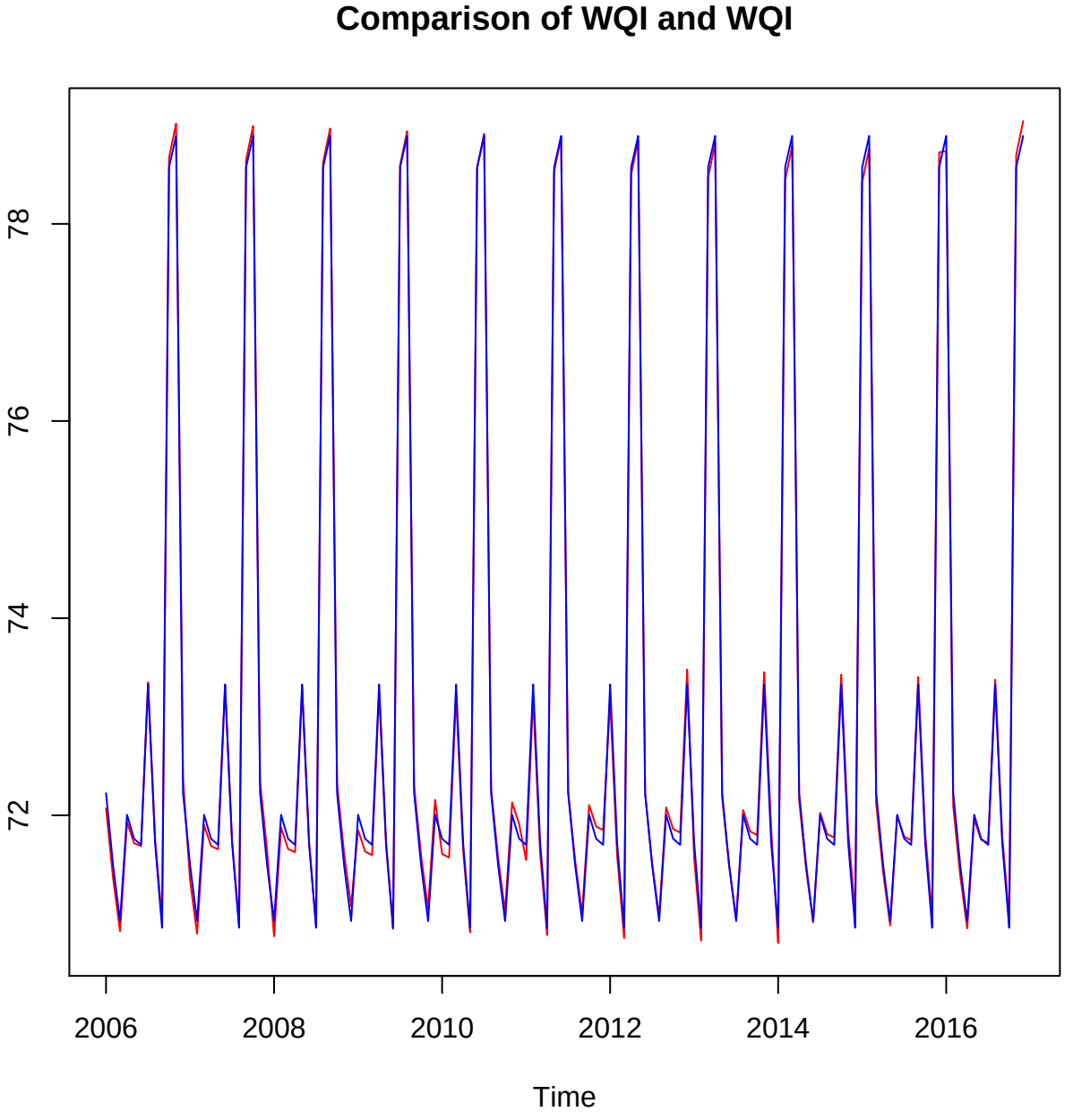

pred2
50
100
150
200
0
500
1000
1500
2000
2500
Index

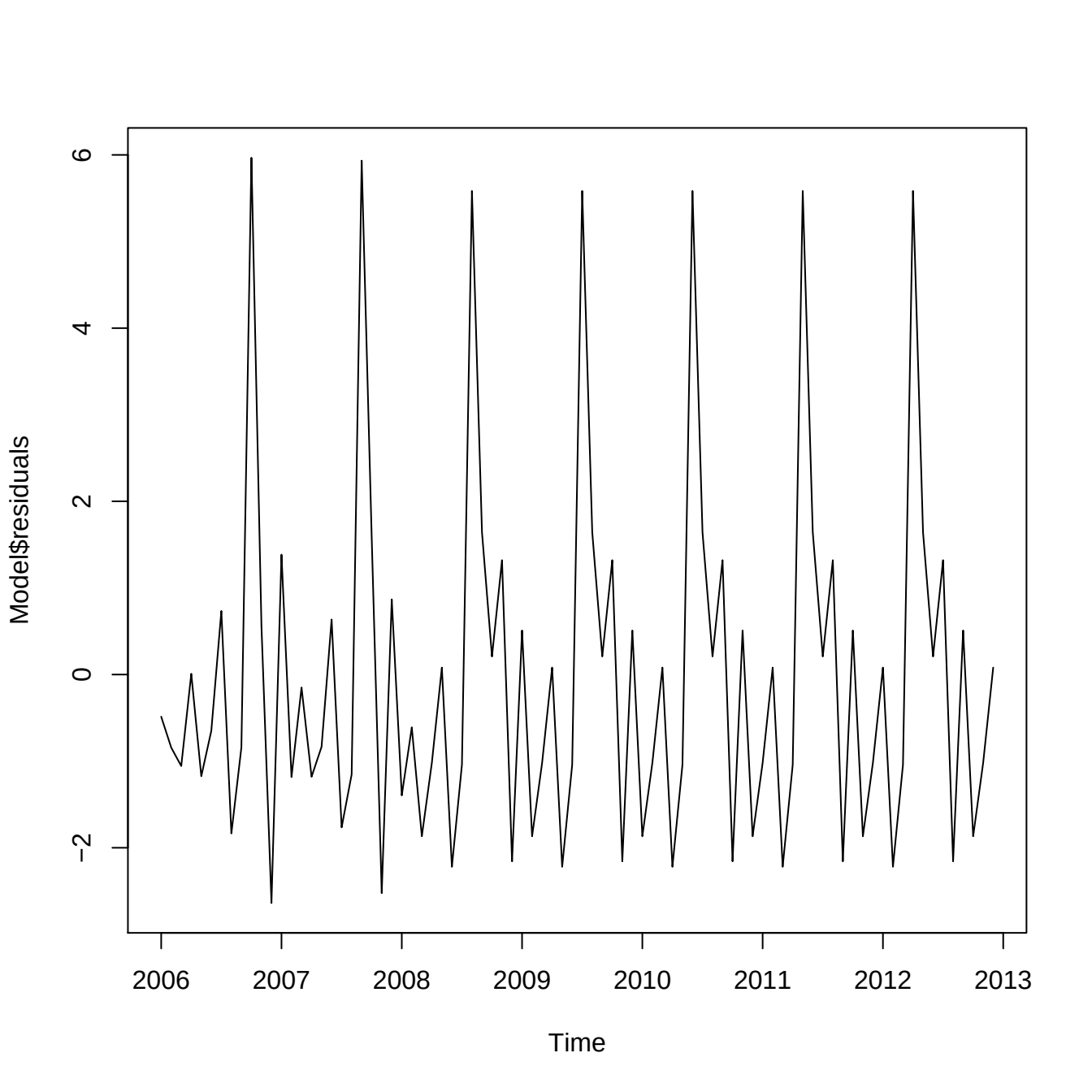

WQITS
72
74
76
78
2006
2008
2010
2012
2014
2016
Time

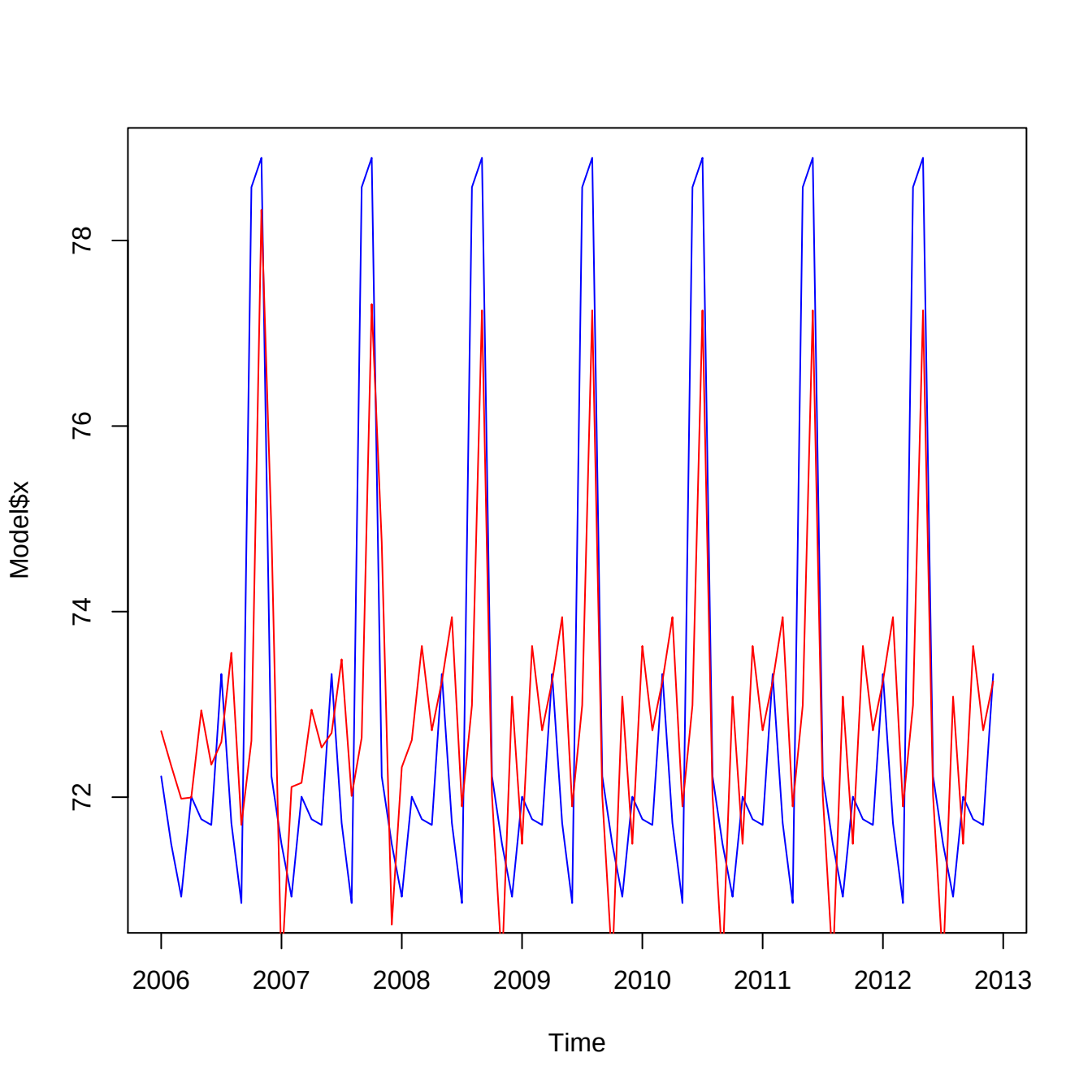

Decomposition of additive time series
observed
trend
seasonal
random
Time
2006
2008
2010
2012
2014
2016

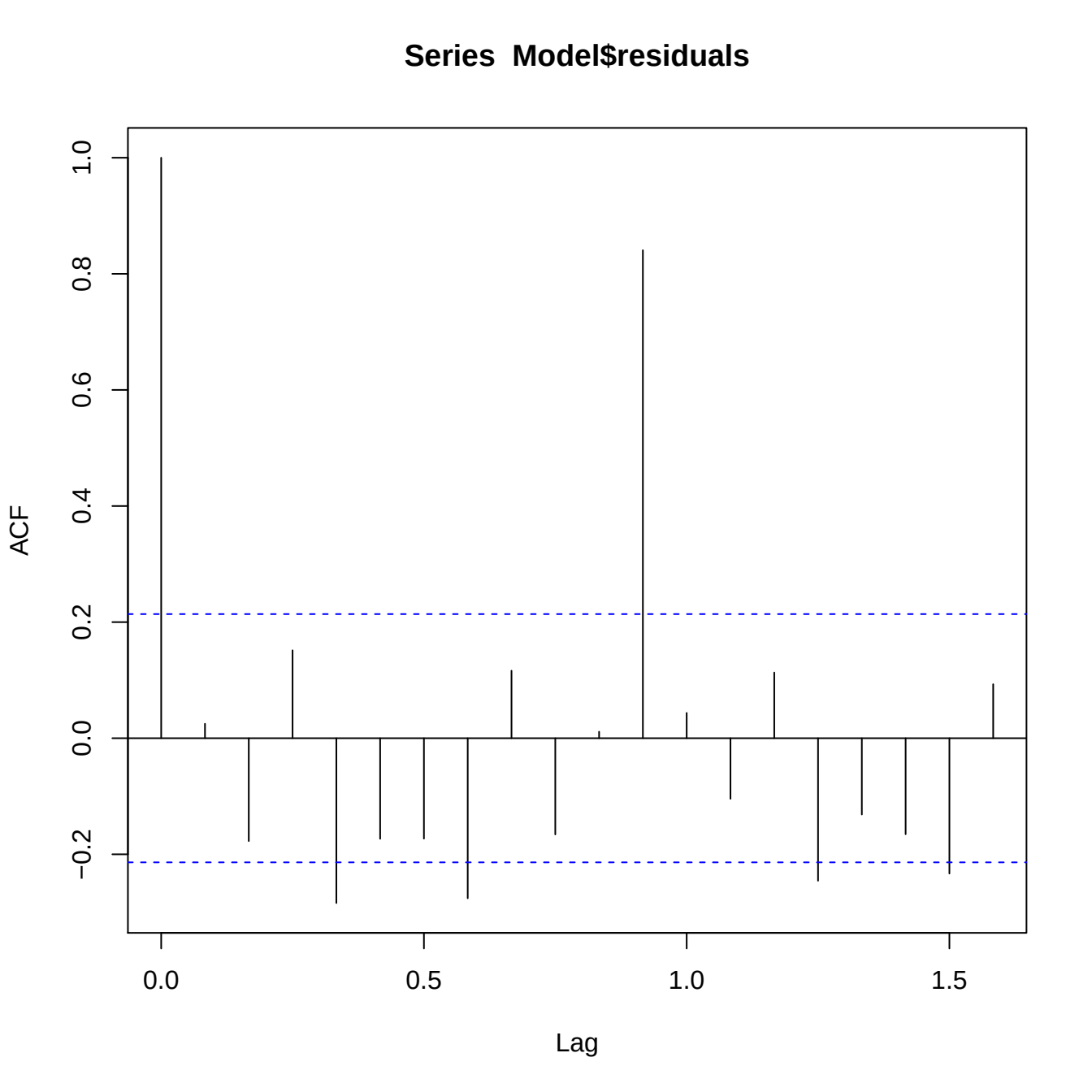

Seasonal plot: WQITS
78
76
74
72
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Month

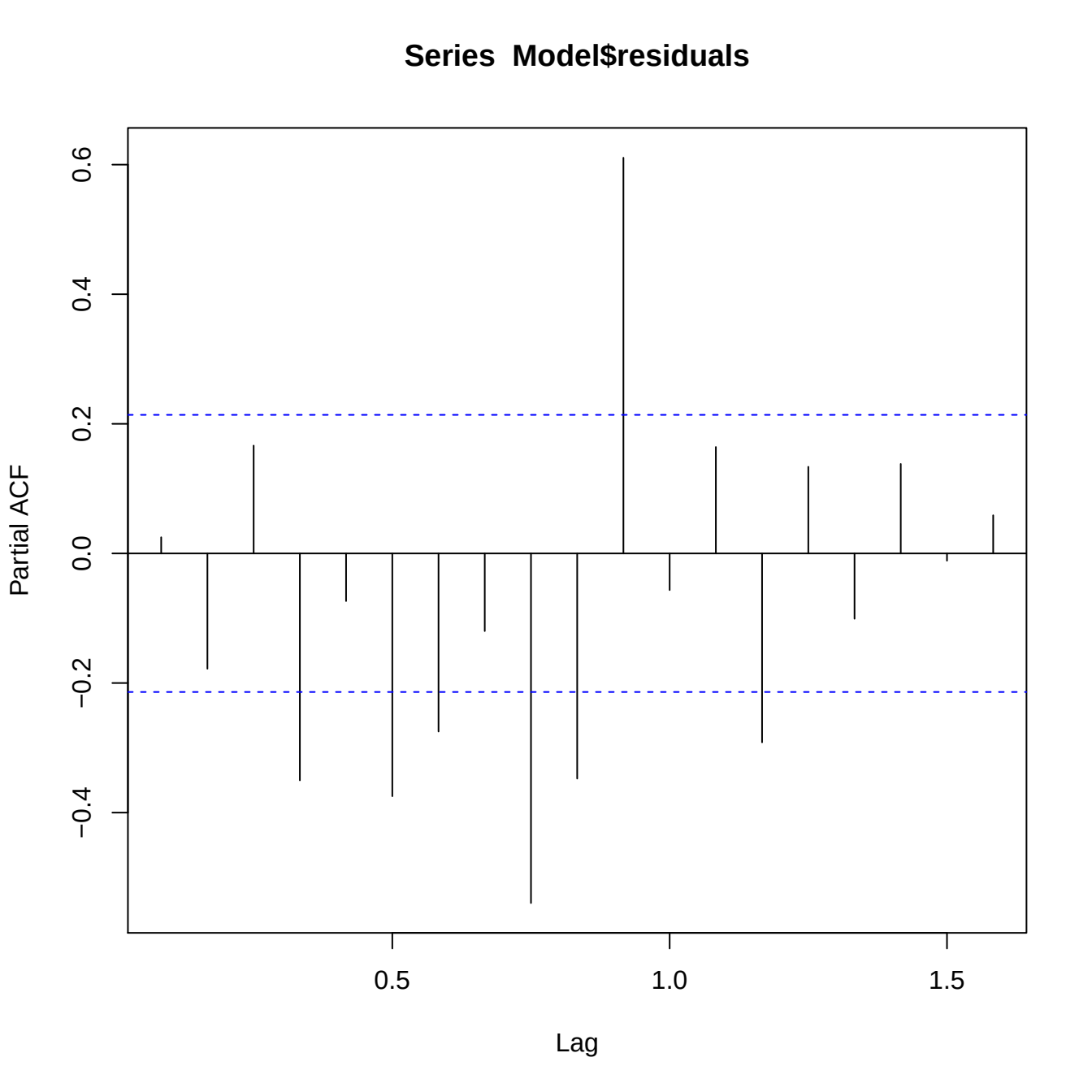

WQITS
72
74
76
78
J
F
M
A
M
J
J
A
S
O
N
D

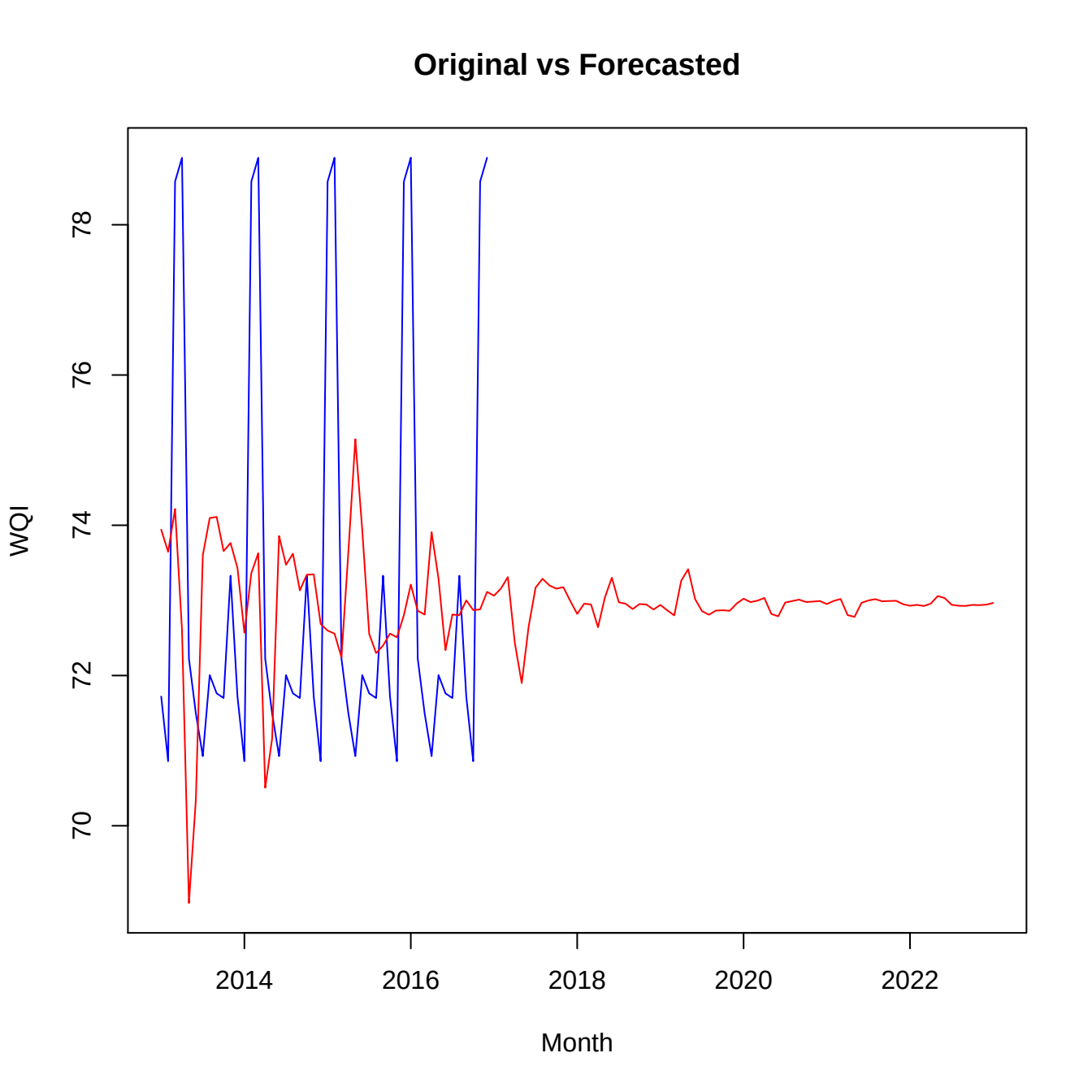

Comparison of WQI and WQI
78
76
74
72
2006
2008
2010
2012
2014
2016
Time


Model$residuals
6
4
2
0
−2
2006
2007
2008
2009
2010
2011
2012
2013
Time


Model$x
78
76
74
72
2006
2007
2008
2009
2010
2011
2012
2013
Time


Series Model$residuals
ACF
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
0.0
0.5
1.0
1.5
Lag


Series Model$residuals
Partial ACF
0.6
0.4
0.2
0.0
−0.2
−0.4
0.5
1.0
1.5
Lag


Original vs Forecasted
WQI
70
72
74
76
78
2014
2016
2018
2020
2022
Month