Tourism Data Analysis

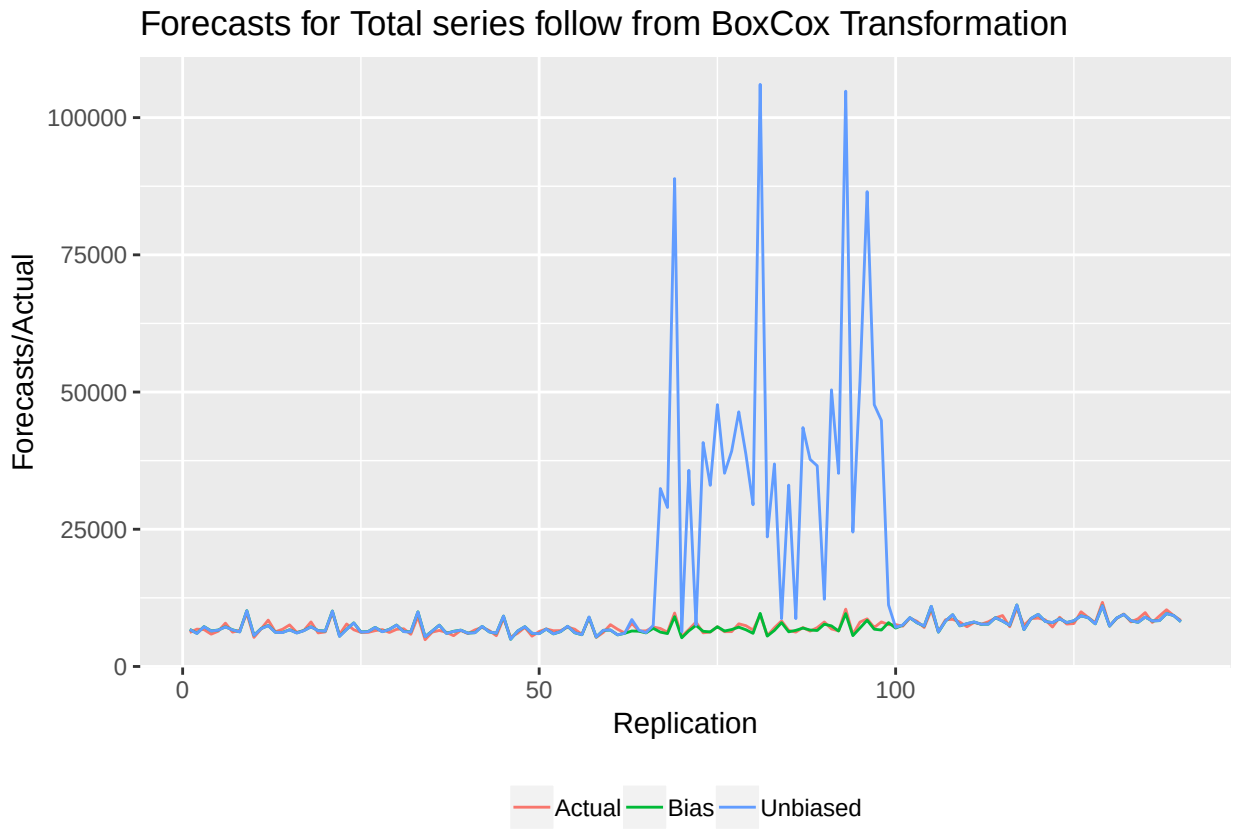
31 July 2019

R.method	Bias	Unbiased (Method 1)	Unbiased (Method 2)
Base	4.64	3521.02	6.73
Bottom-up	6.47	5.29	8.10
MinT(Shrink)	4.35	4.23	4.67
OLS	4.50	2938.01	6.38
WLS	4.90	4.40	5.55

More analysis on BoxCox transformed results

```
## # A tibble: 1 x 4
##   Series  Bias Unbiased_M1 `Bias-Unbiase`
##   <fct>  <dbl>       <dbl>          <dbl>
## 1 Total   275.     386984.       -386710.
```

Bias correction is giving wiered results for `Total` series.




```
DF_BoxCoxTrans_Tot_h1 %>%
  spread(key = `Forecasts/Actual`, value = Val) %>%
  mutate("Unb-Act" = Unbiased_M1 - Actual) %>%
  filter(`Unb-Act`>1000)
```

```
##    Replication    Actual     Bias Unbiased_M1   Unb-Act
## 1           24  6690.320 7942.644    7929.872  1239.551
## 2           67  6922.011 6256.862   32427.930 25505.920
## 3           68  6253.403 5990.885   28965.728 22712.325
## 4           69  9742.298 9074.247   88888.278 79145.980
## 5           71  6735.583 6492.927   35741.613 29006.030
## 6           73  6129.598 6408.122   40797.225 34667.627
## 7           74  6266.592 6297.023   32998.078 26731.487
## 8           75  7240.105 7236.467   47696.759 40456.654
## 9           76  6333.832 6456.888   35207.147 28873.315
## 10          77  6347.804 6727.321   39234.151 32886.347
## 11          78  7770.601 7162.167   46393.968 38623.367
## 12          79  7427.809 6689.464   38641.113 31213.304
## 13          80  6640.508 6034.327   29470.252 22829.744
## 14          81  9419.311 9672.074  106057.019 96637.708
## 15          82  5605.903 5534.823   23602.321 17996.418
## 16          83  7058.426 6578.873   36903.003 29844.577
## 17          85  6566.996 6308.722   33039.816 26472.820
## 18          86  6258.781 6578.185    8721.849  2463.068
## 19          87  7098.180 6994.092   43522.680 36424.500
## 20          88  6436.991 6633.680   37729.921 31292.930
## 21          89  7071.804 6554.580   36555.710 29483.905
## 22          90  8106.166 7703.135   12245.860  4139.695
## 23          91  6839.574 7390.927   50397.661 43558.087
## 24          92  6482.379 6460.287   35188.677 28706.298
## 25          93 10445.812 9631.840  104815.092 94369.280
## 26          94  5847.470 5615.639   24482.840 18635.370
## 27          95  8089.742 7008.365   51987.132 43897.390
## 28          96  8678.255 8423.320   86514.330 77836.075
## 29          97  7095.224 6789.855   47695.463 40600.240
## 30          98  8116.618 6635.950   44825.584 36708.966
## 31          99  7611.168 7986.305   11211.788  3600.620
```

Very large bias corrected forecasts are given for some replications. For example in the rolling window Jul-2003 to Oct-2011. Observing the model auto.arima fits:

```
## Warning: Missing column names filled in: 'X1' [1]
```

```
## Parsed with column specification:
## cols(
##   .default = col_double(),
##   `Month returned from trip` = col_character()
## )
```

```
## See spec(...) for full column specifications.
```

```
## Series: TS
## ARIMA(0,0,0)(2,1,1)[12] with drift
## Box Cox transformation: lambda= -0.9999242
##
## Coefficients:
```

```
## Warning in sqrt(diag(x$var.coef)): NaNs produced
```

```
##          sar1     sar2     sma1  drift
##       -0.1604  -0.3213  -0.7427      0
## s.e.   0.0666   0.0815   0.1256    NaN
```

```
##
## sigma^2 estimated as 1.429e-07:  log likelihood=878.67
## AIC=-1747.34   AICc=-1746.61   BIC=-1734.95
##
## Training set error measures:
##                     ME     RMSE      MAE      MPE     MAPE     MASE
## Training set 758.0031 2294.668 1101.829 9.975516 15.40116 2.203493
##                    ACF1
## Training set 0.8644324
```

The estimated drift term has a very large standard error. Further the variance of $\hat{y}_{t+1}$ is 6903171.42872441 which is very large

Log Transformation

R.method	Bias	Unbiased (Method 1)	Unbiased (Method 2)
Base	4.47	4.43	4.51
Bottom-up	6.36	5.26	8.06
MinT(Shrink)	4.32	4.16	4.61
OLS	4.34	4.31	4.36
WLS	4.82	4.38	5.43