STAT TEST

Ben Tutka

2025-10-20

```
library(readxl)
 library(forecast)
```

```
## Registered S3 method overwritten by 'quantmod':
##   method            from
##   as.zoo.data.frame zoo
```

```
 library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
 ####
 #### Load Data for Maximum Temperatures (1940-1959)
 ####
 # Load and prepare the data
 data <-read_excel("3_LRArea.xlsx")
```

```
## Warning: Expecting numeric in B1915 / R1915C2: got 'M'
```

```
## Warning: Expecting numeric in C1915 / R1915C3: got 'M'
```

```
## Warning: Expecting numeric in B13636 / R13636C2: got 'M'
```

```
 data$Date <-as.Date(data$Date, format = "%Y-%m-%d")
 data$MaxTemperature <-as.numeric(data$MaxTemperature)
 # Filter data for the period between 1940 and 1959, excluding February 29
 filtered_data <-data %>%
 filter(Date >= as.Date("1940-01-01") & Date <= as.Date("1959-12-31")) %>% filter(!(format(Date, "%m-%d
 ######
```

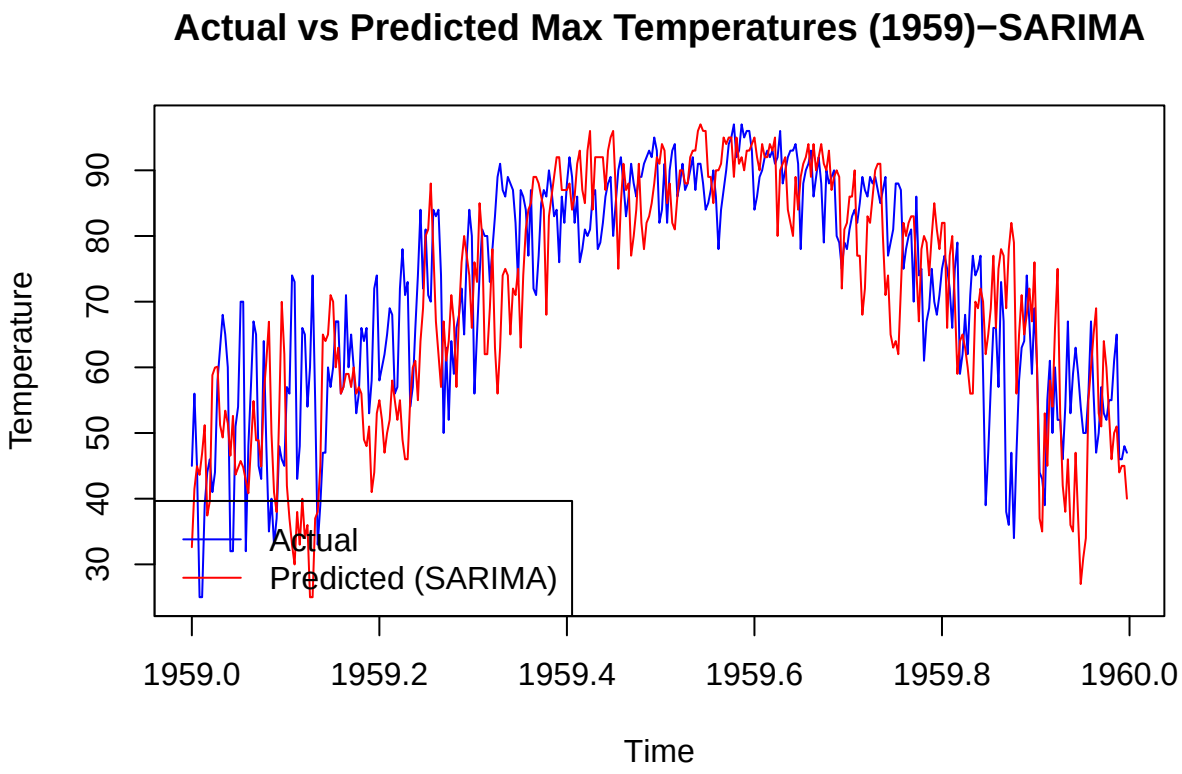
```
#### Prepare Training and Test Sets
######
# Convert to time series
full_ts <-ts(filtered_data$MaxTemperature, frequency = 365, start =
c(1940, 1))
# Split data into training (1940-1958) and test (1959) sets
train_data <-filtered_data %>%
filter(Date < as.Date("1959-01-01"))
test_data <-filtered_data %>%
filter(Date >= as.Date("1959-01-01"))
train_ts <-ts(train_data$MaxTemperature, frequency = 365, start = c(1940,
1))
test_ts <-ts(test_data$MaxTemperature, frequency = 365, start = c(1959,
1))
######
#### SARIMA Model
######
# Fit the SARIMA model
sarima_model <-auto.arima(
train_ts,
seasonal = TRUE,
stepwise = FALSE,
approximation = TRUE
)
# Forecast for 1959
forecast_1959_sarima <-forecast(sarima_model, h = length(test_ts))
# Compute error metrics for 1959
accuracy_metrics_sarima <-accuracy(forecast_1959_sarima, test_ts)
print("Error metrics for 1959 (SARIMA):")
```

```
## [1] "Error metrics for 1959 (SARIMA):"
```

```
print(accuracy_metrics_sarima)
```

```
##                       ME      RMSE      MAE        MPE     MAPE      MASE
## Training set 0.01739816  9.525205 6.966744 -1.3536957 11.26700 0.7217432
## Test set     1.58448616 12.966001 9.825880  0.4234328 15.71075 1.0179450
##                       ACF1 Theil's U
## Training set -0.0004868518        NA
## Test set      0.6317793903  1.565125
```

```
# Plot actual vs predicted for SARIMA (1959)
plot(
test_ts, main = "Actual vs Predicted Max Temperatures (1959)-SARIMA",
col = "blue",
xlab = "Time", ylab = "Temperature", ylim = range(c(test_ts,
forecast_1959_sarima$mean))
)
lines(forecast_1959_sarima$mean, col = "red")
legend("bottomleft", legend = c("Actual", "Predicted (SARIMA)"), col =
c("blue", "red"), lty = 1)
```




```
######
#### Holt-Winters Model
######
# Fit the Holt-Winters model
holt_winters_model <-HoltWinters(train_ts)
# Forecast for 1959
forecast_1959_hw <-forecast(holt_winters_model, h = length(test_ts))
# Compute error metrics for 1959
accuracy_metrics_hw <-accuracy(forecast_1959_hw, test_ts)
print("Error metrics for 1959 (Holt-Winters):")
```

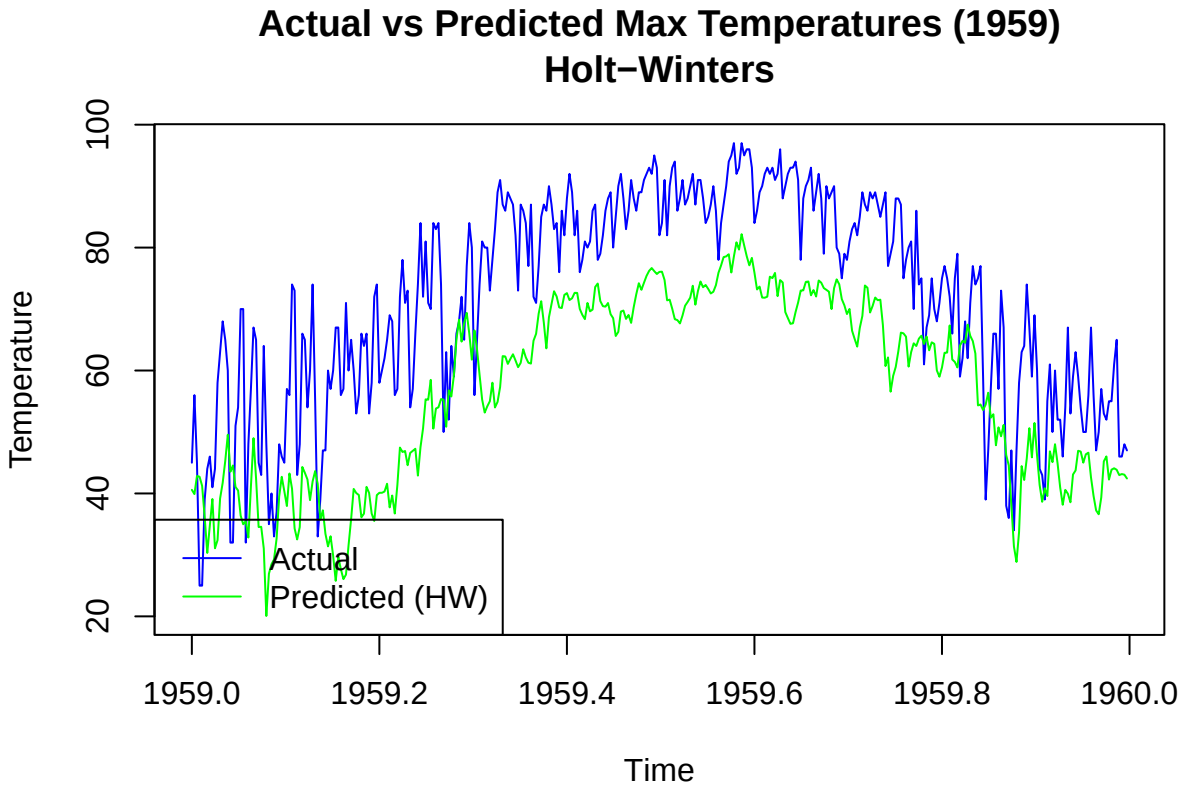
```
## [1] "Error metrics for 1959 (Holt-Winters):"
```

```
print(accuracy_metrics_hw)
```

```
##                       ME      RMSE       MAE       MPE     MAPE      MASE
## Training set -0.009569505  8.736454  6.484791 -1.161649 10.46569 0.6718136
## Test set     14.991006474 17.842199 15.834514 19.992687 22.34000 1.6404295
##                   ACF1 Theil's U
## Training set 0.1872105        NA
## Test set     0.5862409   1.85813
```

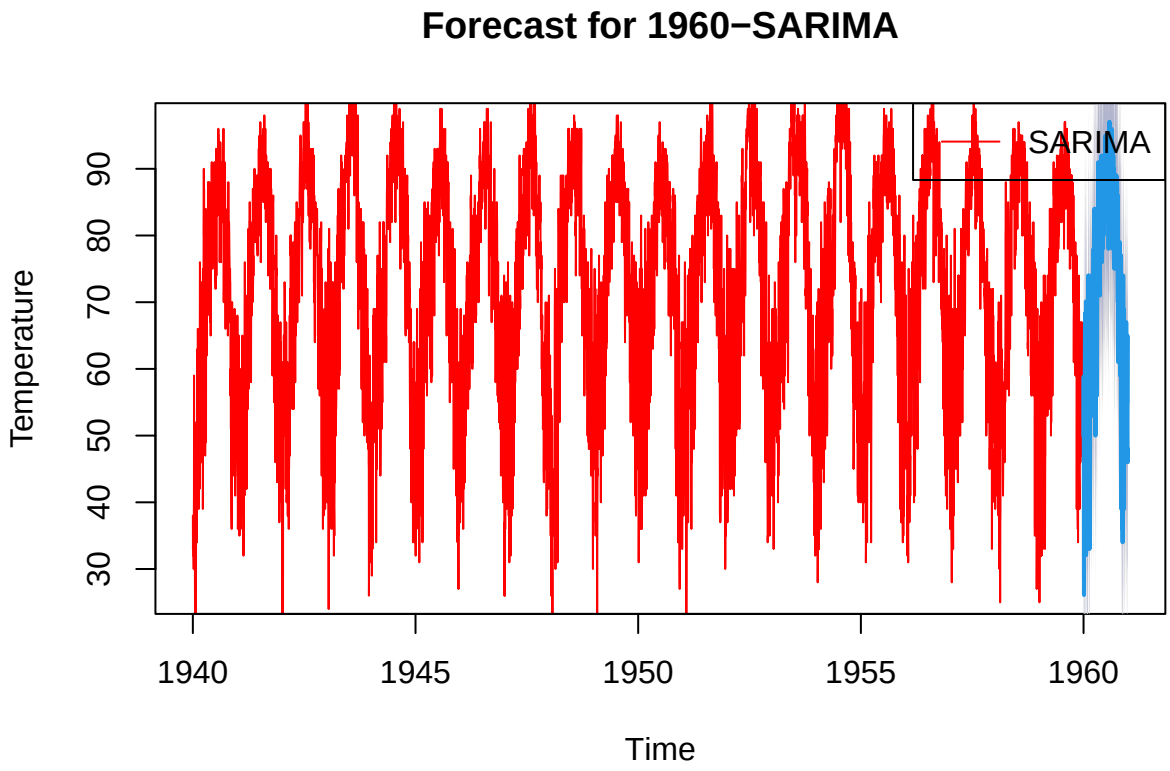
```
# Plot actual vs predicted for Holt-Winters (1959)
plot(
```

```
 test_ts, main = "Actual vs Predicted Max Temperatures (1959)
Holt-Winters", col = "blue",
 xlab = "Time", ylab = "Temperature", ylim = range(c(test_ts,
 forecast_1959_hw$mean))
 )
 lines(forecast_1959_hw$mean, col = "green")
 legend(
 "bottomleft",
 legend = c("Actual", "Predicted (HW)"),
 col = c("blue", "green"),
 lty = 1
 )
```

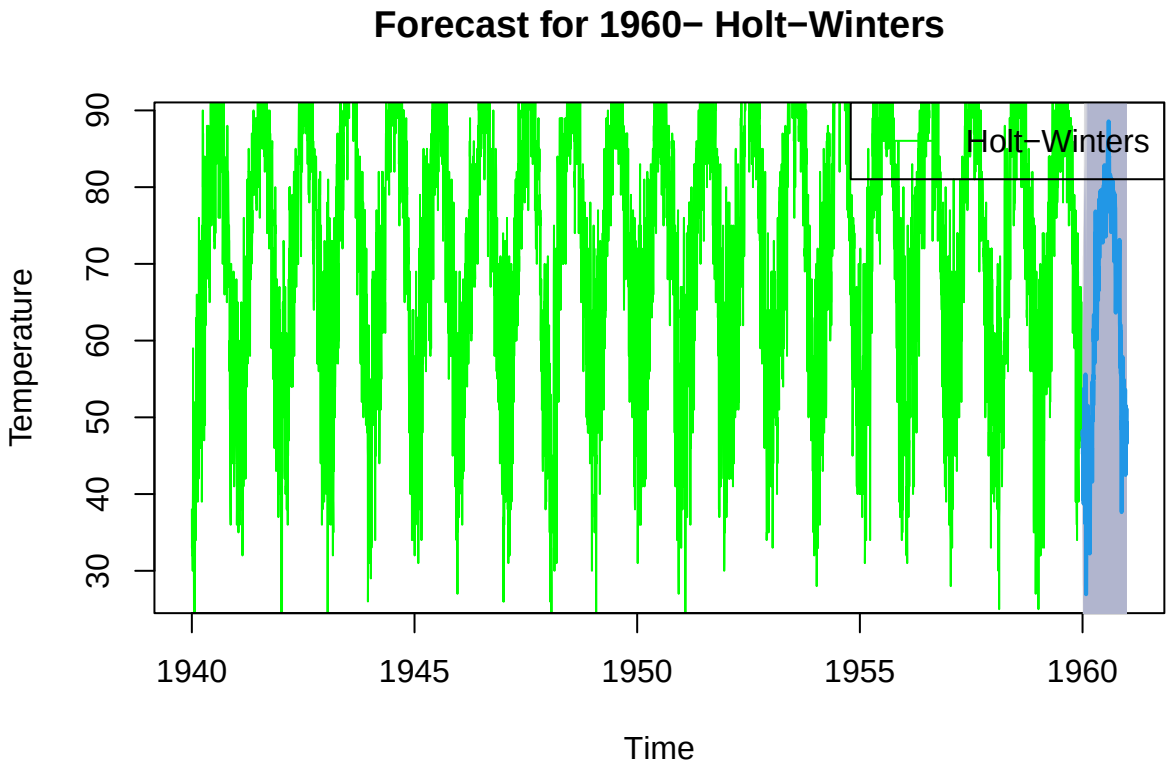



```
######
#### Forecast for 1960
######
# Refit SARIMA on the full dataset and forecast for 1960
final_sarima_model <-auto.arima(full_ts, seasonal = TRUE, stepwise =
FALSE, approximation = TRUE)
forecast_1960_sarima <-forecast(final_sarima_model, h = 365)
# Refit Holt-Winters on the full dataset and forecast for 1960
final_hw_model <-HoltWinters(full_ts)
forecast_1960_hw <-forecast(final_hw_model, h = 365)
# Plot SARIMA forecast for 1960
plot(
```

```
forecast_1960_sarima, main = "Forecast for 1960-SARIMA", xlab =
"Time", ylab = "Temperature",
col = "red", ylim = range(forecast_1960_sarima$mean)
)
legend("topright", legend = c("SARIMA"), col = c("red"), lty = 1)
```




```
 # Plot Holt-Winters forecast for 1960
plot(
forecast_1960_hw, main = "Forecast for 1960- Holt-Winters", xlab =
"Time", ylab = "Temperature",
col = "green", ylim = range(forecast_1960_hw$mean)
)
legend("topright", legend = c("Holt-Winters"), col = c("green"), lty = 1)
```


Forecast for 1960− Holt−Winters
Temperature
30
40
50
60
70
80
90
1940
1945
1950
1955
1960
Time
Holt−Winters