

ARIMA MODELS - without log-back transformations

Victor Felix

November 11, 2025

Here we run ensembles and single automatic ARIMA models for forecasting weekly hospitalizations in 48 states on the contiguous U.S. The models fit on a rolling window of previous 104 weeks to generate forecasts. These models do not include log-back transformations. It returns some metrics that evaluate the performance of the models: target_end_date, abs_error, cases, forecast, number of models, weighted interval score (WIS) and predictive quantiles (%). The user defines if it will use an AUTO ARIMA (auto=TRUE), or ensembles of 27 (ES27=TRUE) or ensembles of 64 models (ES64=TRUE). It can also choose the number of weeks ahead for each forecast, and the size of the rolling window which is set as 104 (2 years).

```
knitr::opts_chunk$set(echo = TRUE)
```

```
!!!!!!!!!!!!!! LOADING THE PACKAGES !!!!!!!
```

```
library("tidyverse")
library("MMWRweek")
library("data.table")
library("caret")
```

```
## Loading required package: ggplot2
```

```
## Loading required package: lattice
```

```
library("purrr")
```

```
##
## Attaching package: 'purrr'
```

```
## The following object is masked from 'package:caret':
##   lift
```

```
## The following object is masked from 'package:data.table':
##   transpose
```

```
library("dplyr")
```

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

```

## The following objects are masked from 'package:data.table':
##
##     between, first, last

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

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

library("tseries")

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

library("gtools")
library("forecast")
library("scoringutils")

## Note: scoringutils is currently undergoing major development changes (with an update planned for the

library("covidHubUtils")
library("parallel")
library("future")#https://cran.r-project.org/web/packages/future/vignettes/future-4-issues.html

##
## Attaching package: 'future'

## The following object is masked from 'package:tseries':
##
##     value

## The following object is masked from 'package:caret':
##
##     cluster

library("listenv")

##
## Attaching package: 'listenv'

## The following object is masked from 'package:purrr':
##
##     map

```

```
library("epitoools")
```

!!!!!!!!!!!!!! LOADING DATASET AND FUNCTIONS !!!!!!!

```
#####
#      LOADING AND CLEANING THE DATASET #
#####

# Loads the ADJACENT states models
source("ES_ARIMA_nolog.R", local = TRUE, chdir = TRUE)

# Loads the ILI dataset
my_data = read.csv("treated_influenza_hosp_dataframe_v2.csv")
my_data$target_end_date<-as.Date(my_data$target_end_date) # set the dates as dates

list_of_states <- split(my_data, my_data$state_name)
```

AUTO ARIMA WEEK1

```
start_time <- Sys.time()

AUTO_ARIMA_WEEK1_list <- mclapply(list_of_states, ES_ARIMA, auto=TRUE, n_weeks_ahead=1, mc.cores=2)%>%
  setNames(names(list_of_states))

end_time <- Sys.time()
run_time <- end_time - start_time

print(run_time)

## Time difference of 11.18519 mins

# Combine the list of data frames into a single data frame with names as a column
AUTO_ARIMA_WEEK1 <- bind_rows(AUTO_ARIMA_WEEK1_list, .id = "State")
```

AUTO ARIMA WEEK2

```
start_time <- Sys.time()

AUTO_ARIMA_WEEK2_list <- mclapply(list_of_states , ES_ARIMA, auto=TRUE, n_weeks_ahead=2, mc.cores=2)%>%
  setNames(names(list_of_states))

end_time <- Sys.time()
run_time <- end_time - start_time

print(run_time)

## Time difference of 10.47833 mins

# Combine the list of data frames into a single data frame with names as a column
AUTO_ARIMA_WEEK2 <- bind_rows(AUTO_ARIMA_WEEK2_list, .id = "State")
```

AUTO ARIMA WEEK3

```
start_time <- Sys.time()

AUTO_ARIMA_WEEK3_list <- mclapply(list_of_states , ES_ARIMA, auto=TRUE, n_weeks_ahead=3, mc.cores=2)%>%
  setNames(names(list_of_states))

end_time <- Sys.time()
run_time <- end_time - start_time

print(run_time)

## Time difference of 10.23431 mins

# Combine the list of data frames into a single data frame with names as a column
AUTO_ARIMA_WEEK3 <- bind_rows(AUTO_ARIMA_WEEK3_list, .id = "State")
```

AUTO ARIMA WEEK4

```
start_time <- Sys.time()

AUTO_ARIMA_WEEK4_list <- mclapply(list_of_states, ES_ARIMA, auto=TRUE, n_weeks_ahead=4 ,mc.cores=2)%>%
  setNames(names(list_of_states))

end_time <- Sys.time()
run_time <- end_time - start_time

print(run_time)

## Time difference of 10.18757 mins

# Combine the list of data frames into a single data frame with names as a column
AUTO_ARIMA_WEEK4 <- bind_rows(AUTO_ARIMA_WEEK4_list, .id = "State")

save.image("ARIMA_MODELS_influenza_hospitalization_nolog.Rdata")
```

ES27 ARIMA WEEK1

```
start_time <- Sys.time()

ES27_ARIMA_WEEK1_list <- mclapply(list_of_states, ES_ARIMA, ES27=TRUE, n_weeks_ahead=1, mc.cores=2)%>%
  setNames(names(list_of_states))

end_time <- Sys.time()
run_time <- end_time - start_time

print(run_time)

## Time difference of 14.31929 mins
```

```
# Combine the list of data frames into a single data frame with names as a column  
ES27_ARIMA_WEEK1 <- bind_rows(ES27_ARIMA_WEEK1_list, .id = "State")
```

ES27 ARIMA WEEK2

```
start_time <- Sys.time()  
  
ES27_ARIMA_WEEK2_list <- mclapply(list_of_states, ES_ARIMA, ES27=TRUE, n_weeks_ahead=2, mc.cores=2)%>%  
  setNames(names(list_of_states))  
  
end_time <- Sys.time()  
run_time <- end_time - start_time  
  
print(run_time)
```

Time difference of 14.40703 mins

```
# Combine the list of data frames into a single data frame with names as a column  
ES27_ARIMA_WEEK2 <- bind_rows(ES27_ARIMA_WEEK2_list, .id = "State")
```

ES27 ARIMA WEEK3

```
start_time <- Sys.time()  
  
ES27_ARIMA_WEEK3_list <- mclapply(list_of_states, ES_ARIMA, ES27=TRUE, n_weeks_ahead=3, mc.cores=2)%>%  
  setNames(names(list_of_states))  
  
end_time <- Sys.time()  
run_time <- end_time - start_time  
  
print(run_time)
```

Time difference of 14.61848 mins

```
# Combine the list of data frames into a single data frame with names as a column  
ES27_ARIMA_WEEK3 <- bind_rows(ES27_ARIMA_WEEK3_list, .id = "State")
```

ES27 ARIMA WEEK4

```
start_time <- Sys.time()  
  
ES27_ARIMA_WEEK4_list <- mclapply(list_of_states, ES_ARIMA, ES27=TRUE, n_weeks_ahead=4 ,mc.cores=2)%>%  
  setNames(names(list_of_states))  
  
end_time <- Sys.time()  
run_time <- end_time - start_time  
  
print(run_time)
```

Time difference of 14.41285 mins

```
# Combine the list of data frames into a single data frame with names as a column  
ES27_ARIMA_WEEK4 <- bind_rows(ES27_ARIMA_WEEK4_list, .id = "State")
```

```
save.image("ARIMA_MODELS_influenza_hospitalization_nolog.Rdata")
```

ES64 ARIMA WEEK1

```
start_time <- Sys.time()  
  
ES64_ARIMA_WEEK1_list <- mclapply(list_of_states, ES_ARIMA, ES64=TRUE, n_weeks_ahead=1, mc.cores=2)%>%  
  setNames(names(list_of_states))  
  
end_time <- Sys.time()  
run_time <- end_time - start_time  
  
print(run_time)
```

Time difference of 38.32509 mins

```
# Combine the list of data frames into a single data frame with names as a column  
ES64_ARIMA_WEEK1 <- bind_rows(ES64_ARIMA_WEEK1_list, .id = "State")
```

ES64 ARIMA WEEK2

```
start_time <- Sys.time()  
  
ES64_ARIMA_WEEK2_list <- mclapply(list_of_states, ES_ARIMA, ES64=TRUE, n_weeks_ahead=2, mc.cores=2)%>%  
  setNames(names(list_of_states))  
  
end_time <- Sys.time()  
run_time <- end_time - start_time  
  
print(run_time)
```

Time difference of 38.08917 mins

```
# Combine the list of data frames into a single data frame with names as a column  
ES64_ARIMA_WEEK2 <- bind_rows(ES64_ARIMA_WEEK2_list, .id = "State")
```

ES64 ARIMA WEEK3

```
start_time <- Sys.time()  
  
ES64_ARIMA_WEEK3_list <- mclapply(list_of_states, ES_ARIMA, ES64=TRUE, n_weeks_ahead=3, mc.cores=2)%>%  
  setNames(names(list_of_states))  
  
end_time <- Sys.time()  
run_time <- end_time - start_time  
  
print(run_time)
```

```
## Time difference of 38.16521 mins

# Combine the list of data frames into a single data frame with names as a column
ES64_ARIMA_WEEK3 <- bind_rows(ES64_ARIMA_WEEK3_list, .id = "State")
```

ES64 ARIMA WEEK4

```
start_time <- Sys.time()

ES64_ARIMA_WEEK4_list <- mclapply(list_of_states, ES_ARIMA, ES64=TRUE, n_weeks_ahead=4 ,mc.cores=2)%>%
  setNames(names(list_of_states))

end_time <- Sys.time()
run_time <- end_time - start_time

print(run_time)
```

```
## Time difference of 37.82942 mins
```

```
# Combine the list of data frames into a single data frame with names as a column
ES64_ARIMA_WEEK4 <- bind_rows(ES64_ARIMA_WEEK4_list, .id = "State")
```

```
save.image("ARIMA_MODELS_influenza_hospitalization_nolog.Rdata")
```