# FAO Detrending Analysis

Rafaela Flach

February 20th, 2021

```
# --- ----- --- ----- --- ----- --- ----- --- ----- ---
# Basic code setup
# --- ----- --- ----- --- ----- --- ----- --- ----- ---

# Load necessary packages

Packages <- c("dplyr","tidyverse","FAOSTAT","smooth",
              "abind","stringr","Metrics",
              "RColorBrewer","ggthemes","ggplot2",
              "ggpubr","wesanderson","kableExtra")

lapply(Packages, library, character.only = TRUE)
```

```
# --- ----- --- ----- --- ----- --- ----- --- ----- ---
# Read and pre-process data
# --- ----- --- ----- --- ----- --- ----- --- ----- ---

inp.folder <- "Data/"

data.fao.code <- "QC" # FAO Dataset code for crop production data

countries <- c("Morocco","United States of America","Germany")
count.code <- c(143,231,79) #codes of the three countries in the FAO dataset


#-----------------------------------
# The next two lines download and save data from FAO.
# If the data are already downloaded, comment these lines

#data.fao.bulk <- get_faostat_bulk(data.fao.code,inp.folder)
#saveRDS(data.fao.bulk, paste0(inp.folder,data.fao.code,"_all_data.rds"))
#-----------------------------------

# Read data saved in folder, pre-process
production_crops <- readRDS(paste0(inp.folder,data.fao.code,"_all_data.rds")) %>%
  filter(area_code %in% count.code,
         element == "Yield",
         item == "Wheat") %>%
  dplyr::select(area,year,value) %>%
  mutate(type = "FAO data")
```

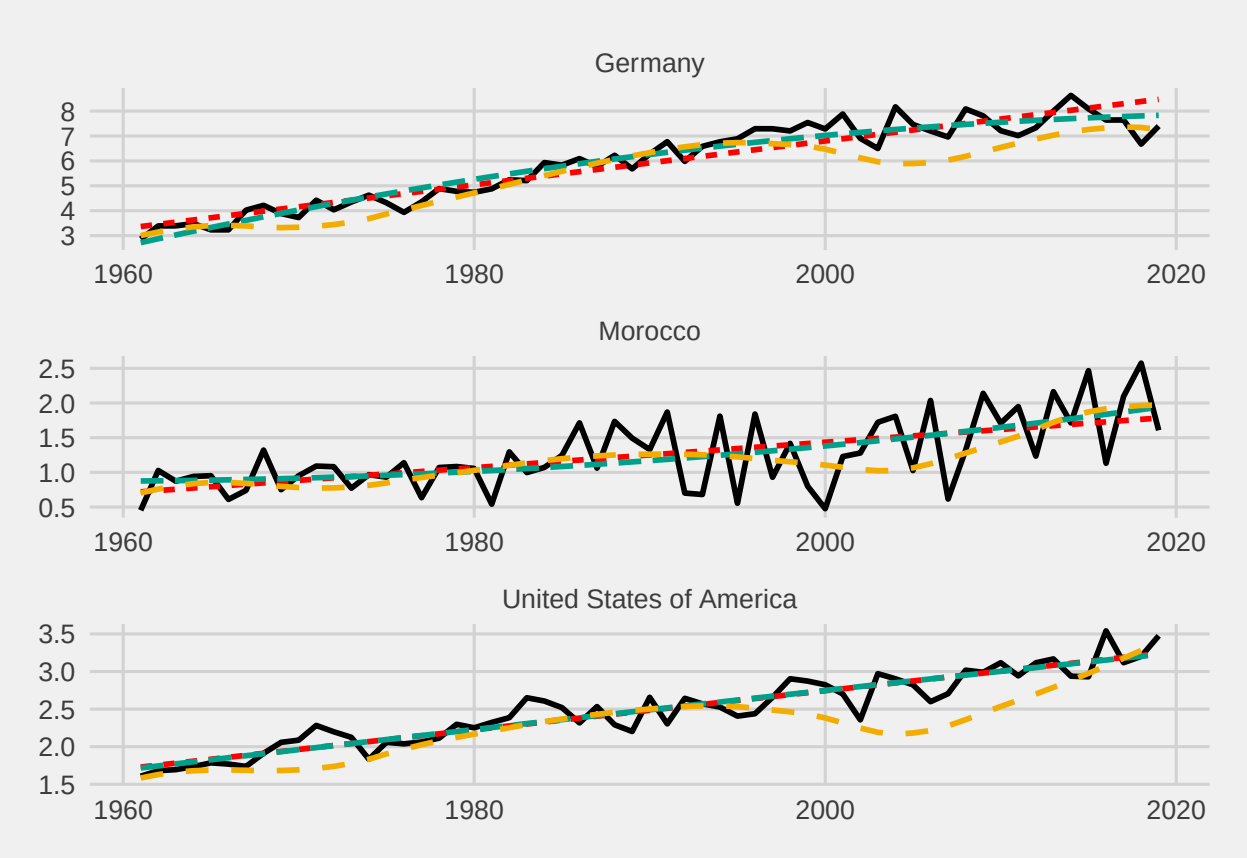
```
# ---- parameters ------
f = 0.9   # This is the loess smoothing parameter
```

```r
# ---- descriptors ------
det.methods <- c("Linear regression",
                 "2-order polynomial",
                 paste0("Loess f=",f))
dec.models <- c("additive","multiplicative")

time.frame <- data.frame(year = sort(unique(production_crops$year)))

# This function uses three different models to fit trend

gendata <- function(country)
{
  countrydata <- production_crops[which(production_crops$area == country),]

  model.lr <- lm(value ~ year, data = countrydata)
  model.pr <- lm(value ~ poly(year,2), data = countrydata)
  model.ql <- loess(value ~ poly(year,2),span = f, countrydata)

  models <- list(model.lr,model.pr,model.ql)
  resultstable <- do.call(rbind,lapply(1:length(models), function(x)
    data.frame(area = country,
               year = time.frame,
               value = models[[x]] %>% predict(time.frame) %>% as.vector,
               type = det.methods[[x]])))
  return(resultstable)
}

# This runs the function above for all countries in list, creates table
# with results

resultstable <- rbind(production_crops,
                      do.call(rbind,
                              lapply(countries, function(x) gendata(x)))) %>%
  mutate(type = factor(type,levels = c("FAO data",det.methods)),
         value = value/10000) %>% drop_na()


# This creates the RMSE for all models and countries
RMSEtable <- data.frame(do.call(rbind,lapply(countries, function(x)
  unlist(lapply(det.methods, function(y)
    rmse(resultstable %>% filter(area == x, type == y) %>% .$value,
         resultstable %>%filter(area == x, type == "FAO data") %>% .$value))))),
  row.names = countries)
colnames(RMSEtable) <- det.methods
```

```r
# This creates the de-trended lines for all models,
# for additive and multiplicative decomposition methods

detrended.tab <- resultstable %>%
  filter(type%in% c("FAO data",det.methods[2])) %>%
  pivot_wider(names_from = type,values_from=value) %>%
  mutate(Multiplicative = `2-order polynomial`*100/`FAO data`,
         Additive = `2-order polynomial`-`FAO data`) %>%
```

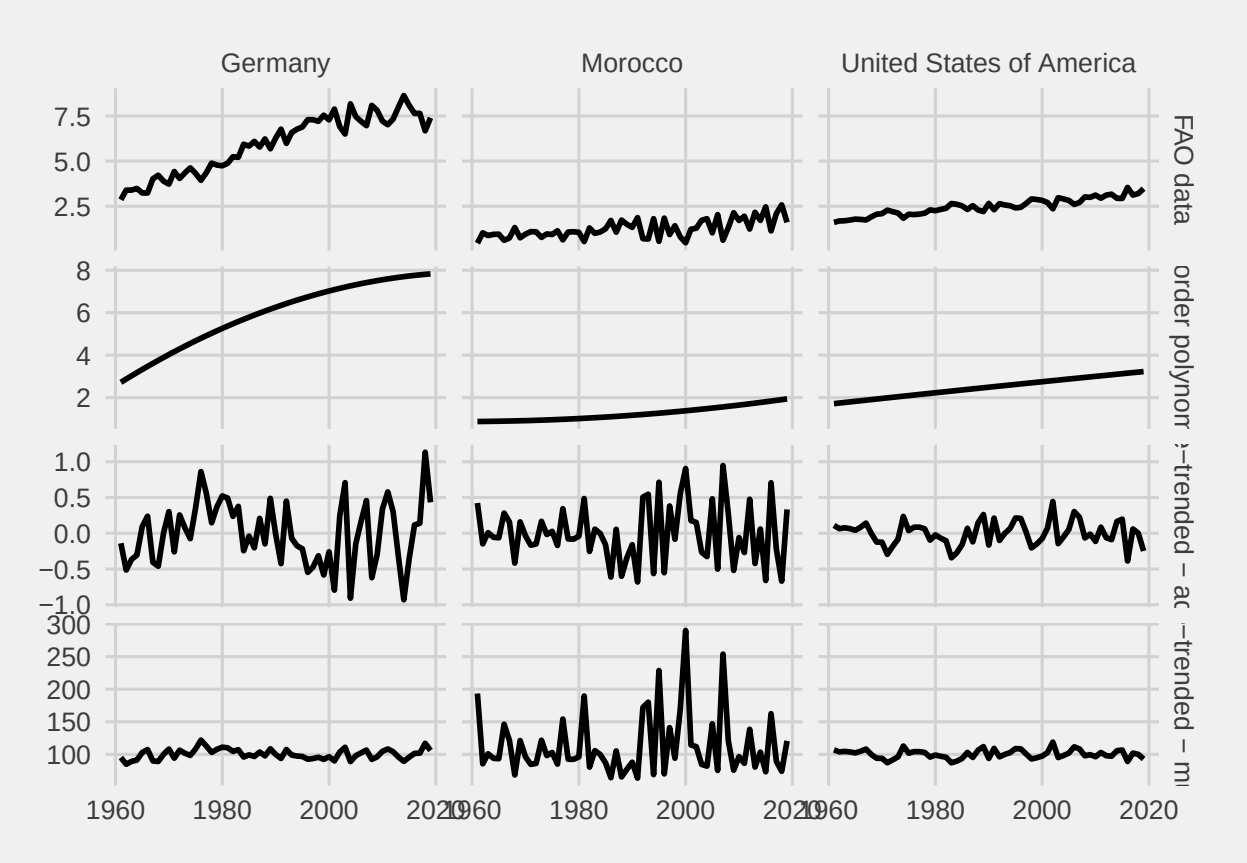
```
  pivot_longer(`FAO data`:Additive, names_to = "type",values_to="value") %>%
  mutate(type = factor(type,levels = c("FAO data", det.methods[2],
                                       "Additive","Multiplicative"),
                       labels = c("FAO data", "2-order polynomial",
                                  "De-trended - add.","De-trended - mult.")))
```

```
fig1 <- ggline(resultstable,
               x = "year",
               y = "value",
               color = "type",
               plot_type = "l",
               facet.by = "area",
               nrow=3,
               scales = "free",
               size = 1.0,
               linetype = "type",
               xlab = "Year",
               ylab = "Yield (ton/ha)") +
  theme_fivethirtyeight() +
  color_palette(c("#000000",wes_palette("Darjeeling1",3))) +
  theme(legend.title = element_blank(), legend.position = "botton")

plot(fig1)
```




```
fig2 <- ggline(detrended.tab ,
               x = "year",
```

```
                y = "value",
                plot_type = "l",
                facet.by = c("type","area"),
                scales = "free",
                size = 1.0) + theme_fivethirtyeight()

plot(fig2)
```




```
kable(RMSEtable)
```

| | Linear regression | 2-order polynomial | Loess f=0.9 |
|---|---|---|---|
| Morocco | 0.4057627 | 0.3990320 | 0.4132280 |
| United States of America | 0.1636212 | 0.1635190 | 0.3935963 |
| Germany | 0.5287394 | 0.4329664 | 0.9357898 |