STL for Great Lakes Temperature Data

Song Qian

8/1/2019

Introduction

This document includes R code for analyzing GL temperature data using STL. Because the current implementation of STL in R does not allow missing values, we used the median polishing algorithm to impute missings before using STL. The function `median.polish.ts()` replaces missing monthly temperature with the estimated annual and monthly median. It also produces two figures to show the seasonal and de-seasoanlized trends. We also modified the STL function to produce desired figures using the function `stl.rfs2()`. This function is a wrapper of the R function `stl()`. It produces the figure and returns data used for the figure. The R code chunk is not included in the knitted PDF file.

Setting Up File System

The following R code chunk defines the default file systems: all input data are in the subfolder `Data`, and resulting figures are saved in subfolder `Figs`.

```
base <- getwd()  ##'~/Google Drive/UT/Research/CILER/STL'
dataDIR <- paste(base, "Data", sep = "/")
figDIR <- paste(base, "Figs", sep = "/")
dataOUTDIR <- paste(dataDIR, "output", sep = "/")

packages(tidyverse)
```

```
## Loading required package: tidyverse
```

```
## -- Attaching packages ------------------------------------------------------------------------ tidyvers
```

```
## v ggplot2 3.1.1       v purrr   0.3.2
## v tibble  2.1.1       v dplyr   0.8.0.1
## v tidyr   0.8.3       v stringr 1.4.0
## v readr   1.3.1       v forcats 0.4.0
```

```
## -- Conflicts --------------------------------------------------------------------------- tidyverse_conf
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
packages(readxl)
```

```
## Loading required package: readxl
```

```
packages(lattice)
```

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

Reading Data

The following R code chunk imports data, processes missing values, and creates time series objects.

```
stl_data <- read_excel(paste(dataDIR, "glsea_trends.xlsx", sep = "/"), col_names = F)
```

```
## New names:
## * `` -> ...1
## * `` -> ...2
## * `` -> ...3
## * `` -> ...4
## * `` -> ...5
## * ... and 8 more problems
```

```
stl_data_therm <- read_excel(paste(dataDIR, "thermistor_trends.xlsx", sep = "/"),
    col_names = F)
```

```
## New names:
## * `` -> ...1
## * `` -> ...2
## * `` -> ...3
## * `` -> ...4
## * `` -> ...5
## * ... and 8 more problems
```

```
Superior <- matrix(as.numeric(unlist(stl_data[4:27, -1])), ncol = 12)
Michigan <- matrix(as.numeric(unlist(stl_data[30:53, -1])), ncol = 12)
Huron <- matrix(as.numeric(unlist(stl_data[56:79, -1])), ncol = 12)
StClair <- matrix(as.numeric(unlist(stl_data[82:105, -1])), ncol = 12)
Erie <- matrix(as.numeric(unlist(stl_data[108:131, -1])), ncol = 12)
Ontario <- matrix(as.numeric(unlist(stl_data[134:157, -1])), ncol = 12)

therm_surf <- matrix(as.numeric(unlist(stl_data_therm[4:27, -1])), ncol = 12)
therm_25m <- matrix(as.numeric(unlist(stl_data_therm[30:58, -1])), ncol = 12)
therm_50m <- matrix(as.numeric(unlist(stl_data_therm[60:89, -1])), ncol = 12)
therm_75m <- matrix(as.numeric(unlist(stl_data_therm[92:120, -1])), ncol = 12)
therm_100m <- matrix(as.numeric(unlist(stl_data_therm[123:151, -1])), ncol = 12)
therm_125m <- matrix(as.numeric(unlist(stl_data_therm[154:182, -1])), ncol = 12)

therm_surf[therm_surf < -900] <- NA
therm_25m[therm_25m < -900] <- NA
therm_50m[therm_50m < -900] <- NA
therm_75m[therm_75m < -900] <- NA
therm_100m[therm_100m < -900] <- NA
therm_125m[therm_125m < -900] <- NA

SP_monthly <- ts(as.vector(t(Superior)), start = c(1995, 1), end = c(2018, 12),
    frequency = 12)
MI_monthly <- ts(as.vector(t(Michigan)), start = c(1995, 1), end = c(2018, 12),
    frequency = 12)
HR_monthly <- ts(as.vector(t(Huron)), start = c(1995, 1), end = c(2018, 12),
    frequency = 12)
SC_monthly <- ts(as.vector(t(StClair)), start = c(1995, 1), end = c(2018, 12),
    frequency = 12)
ER_monthly <- ts(as.vector(t(Erie)), start = c(1995, 1), end = c(2018, 12),
    frequency = 12)
ON_monthly <- ts(as.vector(t(Ontario)), start = c(1995, 1), end = c(2018, 12),
    frequency = 12)
```

```
thermSurf_monthly <- ts(as.vector(t(therm_surf)), start = c(1990, 1), frequency = 12)
therm25m_monthly <- ts(as.vector(t(therm_25m)), start = c(1990, 1), frequency = 12)
therm50m_monthly <- ts(as.vector(t(therm_50m)), start = c(1990, 1), frequency = 12)
therm75m_monthly <- ts(as.vector(t(therm_75m)), start = c(1990, 1), frequency = 12)
therm100m_monthly <- ts(as.vector(t(therm_100m)), start = c(1990, 1), frequency = 12)
therm125m_monthly <- ts(as.vector(t(therm_125m)), start = c(1990, 1), frequency = 12)
```

STL Analysis

The analysis uses the function `stl.rfs2()` to produce the STL plots for each lake and export data for reproducing the figures using other graphical systems.

```
pdf(paste(figDIR, "SuperiorSTL.pdf", sep = "/"), height = 4, width = 5)
Superior.STL <- stl.rfs2(data = SP_monthly, aspect = 2.25, ylab = "Temperature",
    ss.w = 11, ss.d = 1, fc.w = 91, fc.d = 1)
dev.off()
```

```
## pdf
##   2
```

```
write.csv(Superior.STL[[3]], paste(dataOUTDIR, "superior_monthly.csv", sep = "/"))
write.csv(Superior.STL[[1]], paste(dataOUTDIR, "superior_components.csv", sep = "/"))

pdf(paste(figDIR, "MichiganSTL.pdf", sep = "/"), height = 4, width = 5)
Michigan.STL <- stl.rfs2(data = MI_monthly, aspect = 2.25, ylab = "Temperature",
    ss.w = 15, ss.d = 1, fc.w = 91, fc.d = 1)
dev.off()
```

```
## pdf
##   2
```

```
write.csv(Michigan.STL[[3]], paste(dataOUTDIR, "michigan_monthly.csv", sep = "/"))
write.csv(Michigan.STL[[1]], paste(dataOUTDIR, "michigan_components.csv", sep = "/"))

pdf(paste(figDIR, "HuronSTL.pdf", sep = "/"), height = 4, width = 5)
Huron.STL <- stl.rfs2(data = HR_monthly, aspect = 2.25, ylab = "Temperature",
    ss.w = 15, ss.d = 1, fc.w = 91, fc.d = 1)
dev.off()
```

```
## pdf
##   2
```

```
write.csv(Huron.STL[[3]], paste(dataOUTDIR, "huron_monthly.csv", sep = "/"))
write.csv(Huron.STL[[1]], paste(dataOUTDIR, "huron_components.csv", sep = "/"))

pdf(paste(figDIR, "ErieSTL.pdf", sep = "/"), height = 4, width = 5)
Erie.STL <- stl.rfs2(data = ER_monthly, aspect = 2.25, ylab = "Temperature",
    ss.w = 15, ss.d = 1, fc.w = 91, fc.d = 1)
dev.off()
```

```
## pdf
##   2
```

```
write.csv(Erie.STL[[3]], paste(dataOUTDIR, "erie_monthly.csv", sep = "/"))
write.csv(Erie.STL[[1]], paste(dataOUTDIR, "erie_components.csv", sep = "/"))
```

```
pdf(paste(figDIR, "StClairSTL.pdf", sep = "/"), height = 4, width = 5)
StClair.STL <- stl.rfs2(data = SC_monthly, aspect = 2.25, ylab = "Temperature",
    ss.w = 15, ss.d = 1, fc.w = 91, fc.d = 1)
dev.off()
```

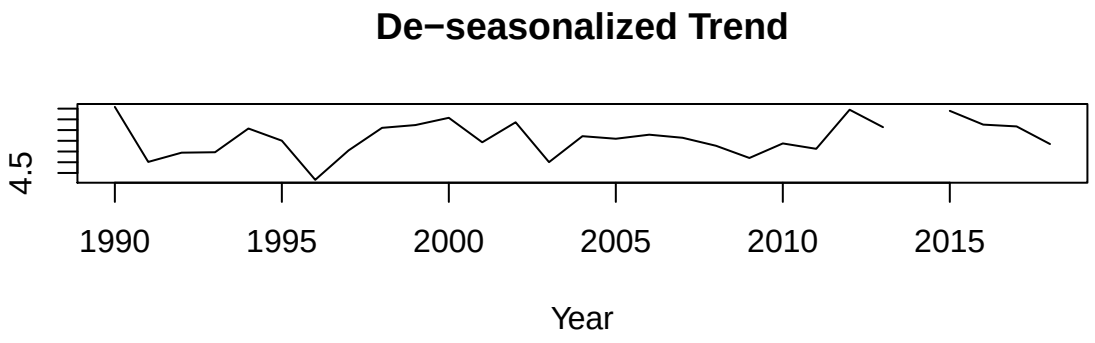
```
## pdf
##   2
```

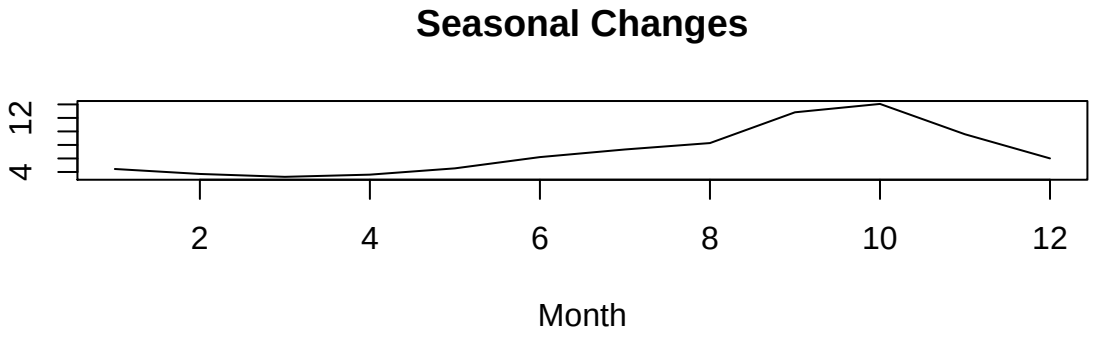
```
write.csv(StClair.STL[[3]], paste(dataOUTDIR, "stclair_monthly.csv", sep = "/"))
write.csv(StClair.STL[[1]], paste(dataOUTDIR, "stclair_components.csv", sep = "/"))

pdf(paste(figDIR, "OntarioSTL.pdf", sep = "/"), height = 4, width = 5)
Ontario.STL <- stl.rfs2(data = ON_monthly, aspect = 2.25, ylab = "Temperature",
    ss.w = 15, ss.d = 1, fc.w = 91, fc.d = 1)
dev.off()
```

```
## pdf
##   2
```

```
write.csv(Ontario.STL[[3]], paste(dataOUTDIR, "ontario_monthly.csv", sep = "/"))
write.csv(Ontario.STL[[1]], paste(dataOUTDIR, "ontario_components.csv", sep = "/"))

therm25_ts <- median.polish.ts(therm25m_monthly)
```

```
## 1: 311.4488
## 2: 231.8249
## 3: 230.8607
## Final: 230.7111
## [1] "overall"   "row"       "col"       "residuals" "name"
```



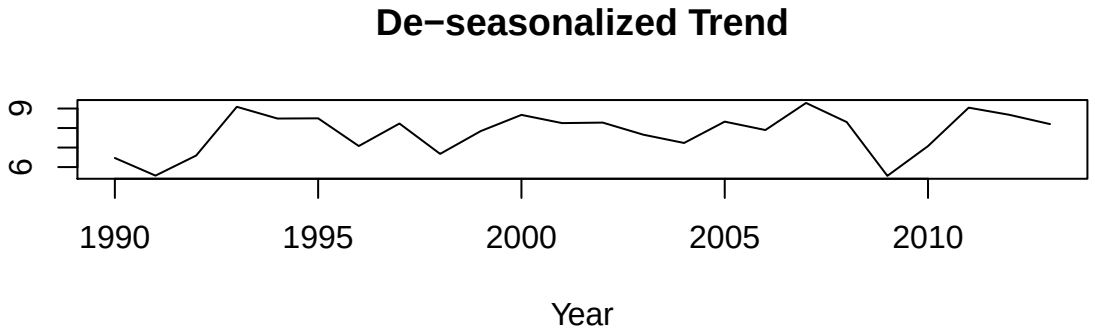


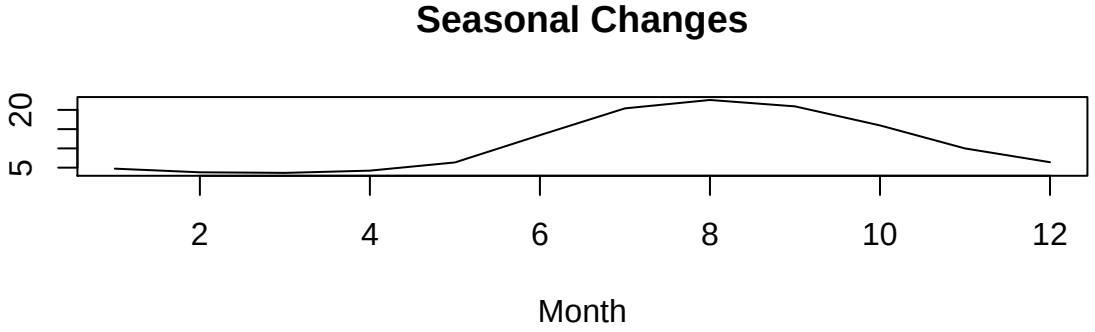
```
pdf(paste(figDIR, "Term_25mSTL.pdf", sep = "/"), height = 4.75, width = 5)
Therm_25m.STL <- stl.rfs2(data = therm25_ts, aspect = 2.25, ylab = "Temperature",
    ss.w = 15, ss.d = 1, fc.w = 91, fc.d = 1)
dev.off()
```

```
## pdf
##   2
```

```
write.csv(Therm_25m.STL[[3]], paste(dataOUTDIR, "therm25m_monthly.csv", sep = "/"))
write.csv(Therm_25m.STL[[1]], paste(dataOUTDIR, "therm25m_components.csv", sep = "/"))

thermsurf_ts <- median.polish.ts(thermSurf_monthly)
```

```
## 1: 294.3012
## 2: 240.5454
## 3: 238.8431
## Final: 238.6529
## [1] "overall"   "row"       "col"       "residuals" "name"
```





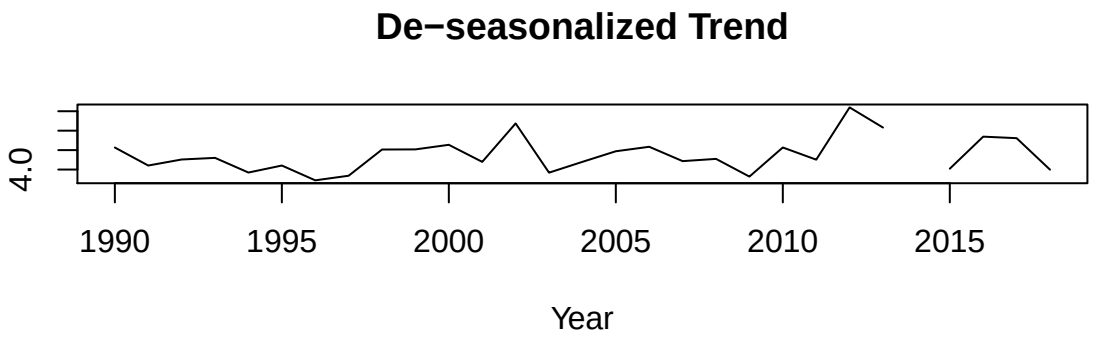

```
pdf(paste(figDIR, "Term_surfSTL.pdf", sep = "/"), height = 4.75, width = 5)
Therm_surf.STL <- stl.rfs2(data = thermsurf_ts, aspect = 2.25, ylab = "Temperature",
    ss.w = 15, ss.d = 1, fc.w = 91, fc.d = 1)
dev.off()
```

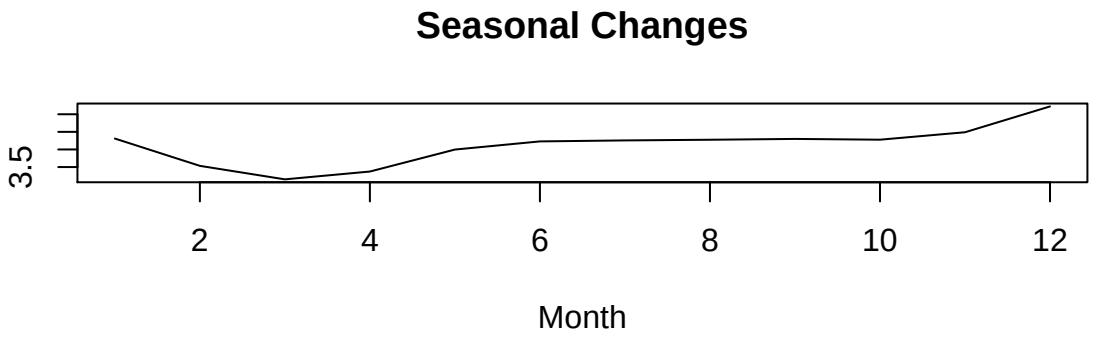
```
## pdf
##   2
```

```
write.csv(Therm_surf.STL[[3]], paste(dataOUTDIR, "thermsurf_monthly.csv", sep = "/"))
write.csv(Therm_surf.STL[[1]], paste(dataOUTDIR, "thermsurf_components.csv",
    sep = "/"))
```

```
therm75_ts <- median.polish.ts(therm75m_monthly)
```

```
## 1: 54.15748
## 2: 48.98291
## 3: 48.9082
## Final: 48.88895
## [1] "overall"   "row"       "col"       "residuals" "name"
```



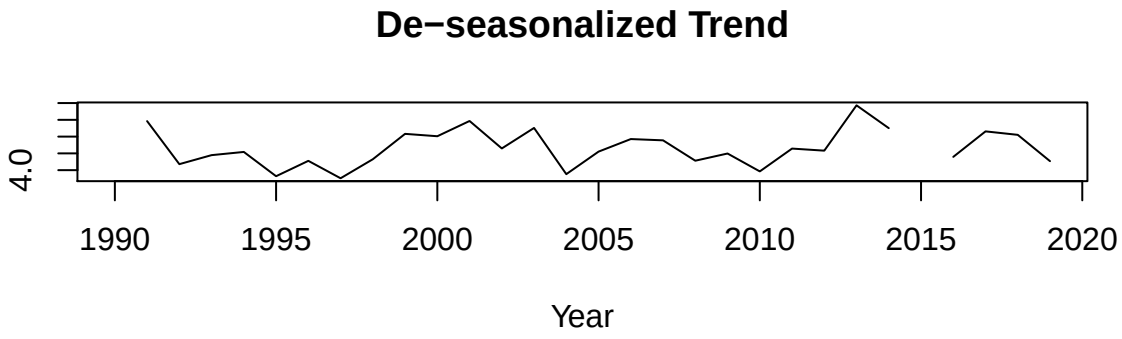


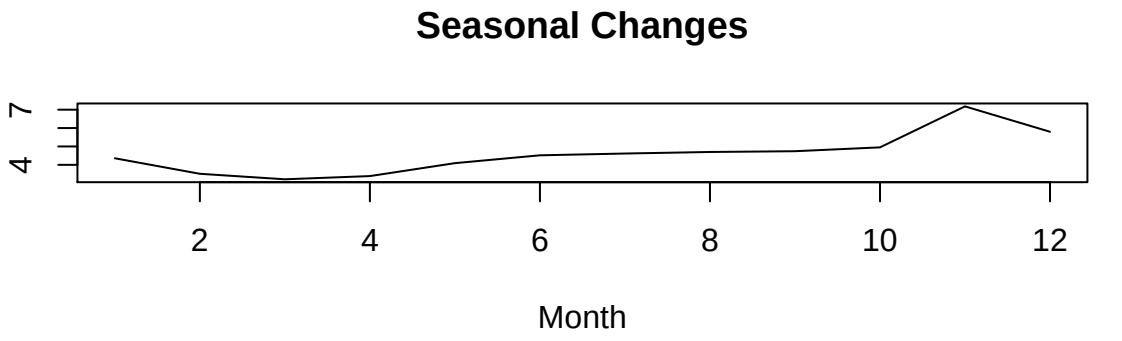

```
pdf(paste(figDIR, "Term_75mSTL.pdf", sep = "/"), height = 4.75, width = 5)
Therm_75m.STL <- stl.rfs2(data = therm75_ts, aspect = 2.25, ylab = "Temperature",
    ss.w = 15, ss.d = 1, fc.w = 91, fc.d = 1)
dev.off()
```

```
## pdf
##   2
```

```
write.csv(Therm_75m.STL[[3]], paste(dataOUTDIR, "therm75m_monthly.csv", sep = "/"))
write.csv(Therm_75m.STL[[1]], paste(dataOUTDIR, "therm75m_components.csv", sep = "/"))

therm50_ts <- median.polish.ts(therm50m_monthly)
```

```
## 1: 100.5456
## 2: 83.1835
## 3: 82.36739
## 4: 82.18339
## Final: 82.10725
## [1] "overall"   "row"       "col"       "residuals" "name"
```



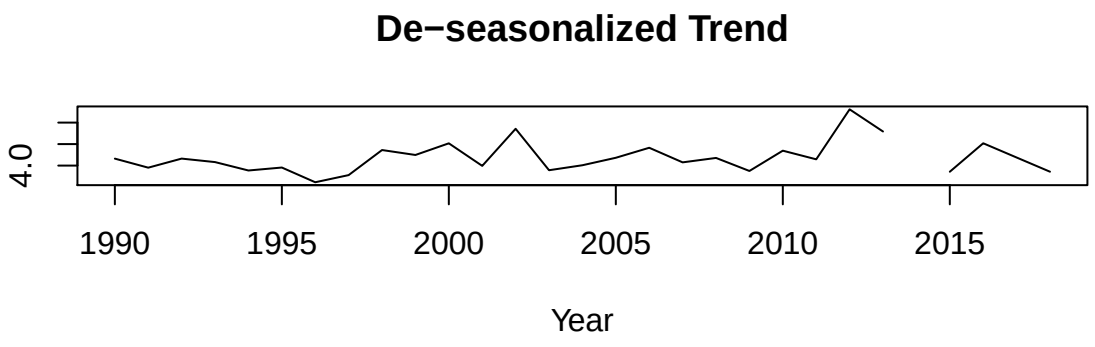


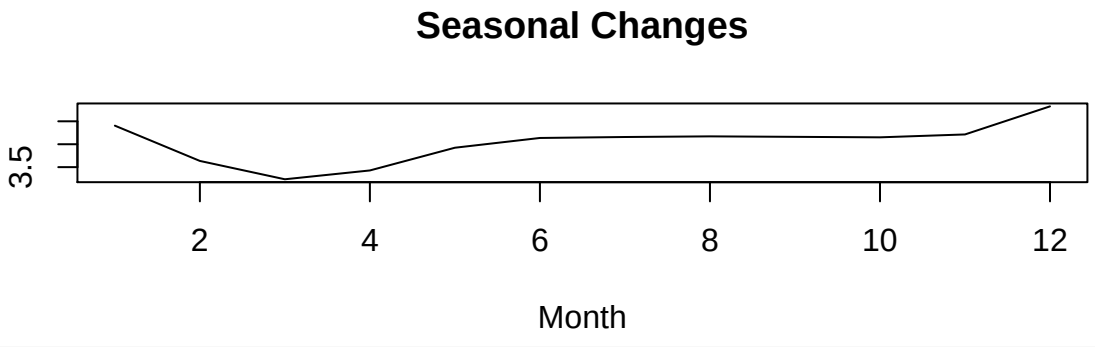

```
pdf(paste(figDIR, "Term_50mSTL.pdf", sep = "/"), height = 4.75, width = 5)
Therm_50m.STL <- stl.rfs2(data = therm50_ts, aspect = 2.25, ylab = "Temperature",
    ss.w = 15, ss.d = 1, fc.w = 91, fc.d = 1)
dev.off()
```

```
## pdf
##   2
```

```
write.csv(Therm_50m.STL[[3]], paste(dataOUTDIR, "therm50m_monthly.csv", sep = "/"))
write.csv(Therm_50m.STL[[1]], paste(dataOUTDIR, "therm50m_components.csv", sep = "/"))

therm100_ts <- median.polish.ts(therm100m_monthly)
```

```
## 1: 47.09763
## 2: 43.53106
## 3: 43.40232
## Final: 43.37905
## [1] "overall"   "row"       "col"       "residuals" "name"
```



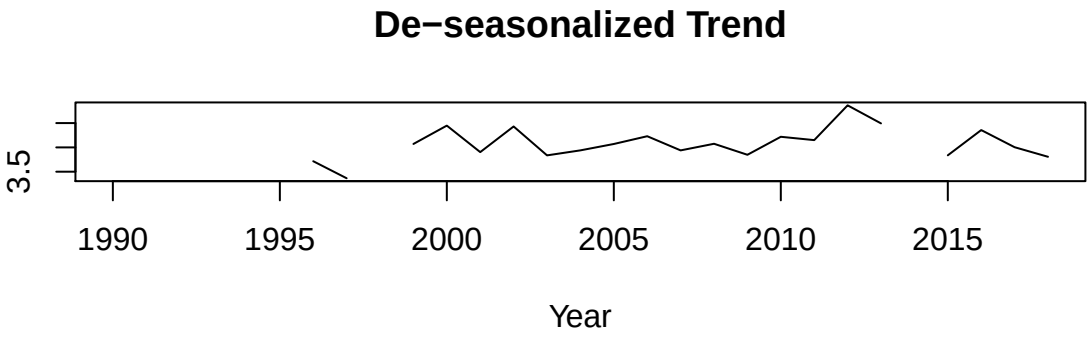


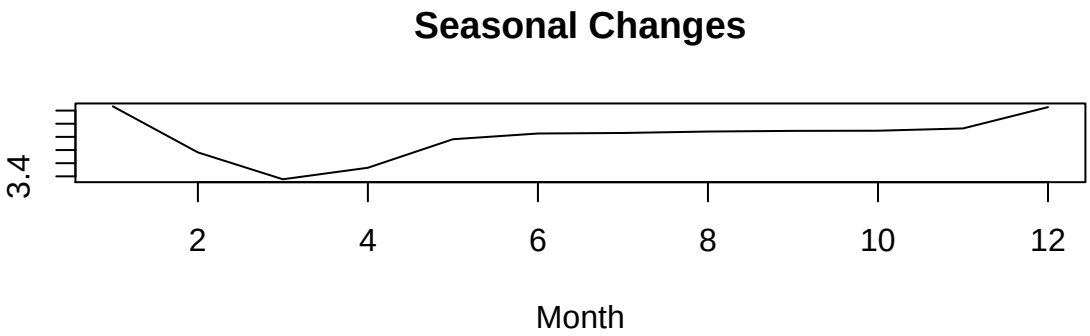

```
pdf(paste(figDIR, "Term_100mSTL.pdf", sep = "/"), height = 4.75, width = 5)
Therm_100m.STL <- stl.rfs2(data = therm100_ts, aspect = 2.25, ylab = "Temperature",
    ss.w = 15, ss.d = 1, fc.w = 91, fc.d = 1)
dev.off()
```

```
## pdf
##   2
```

```
write.csv(Therm_100m.STL[[3]], paste(dataOUTDIR, "therm100m_monthly.csv", sep = "/"))
write.csv(Therm_100m.STL[[1]], paste(dataOUTDIR, "therm100m_components.csv",
    sep = "/"))

therm125_ts <- median.polish.ts(therm125m_monthly)
```

```
## 1: 29.94145
## 2: 26.3536
## 3: 26.22647
## 4: 26.19686
## Final: 26.18505
## [1] "overall"   "row"       "col"       "residuals" "name"
```






```
pdf(paste(figDIR, "Term_125mSTL.pdf", sep = "/"), height = 4.75, width = 5)
Therm_125m.STL <- stl.rfs2(data = therm125_ts, aspect = 2.25, ylab = "Temperature",
    ss.w = 15, ss.d = 1, fc.w = 91, fc.d = 1)
dev.off()
```

```
## pdf
##   2
```

```
write.csv(Therm_125m.STL[[3]], paste(dataOUTDIR, "therm125m_monthly.csv", sep = "/"))
write.csv(Therm_125m.STL[[1]], paste(dataOUTDIR, "therm125m_components.csv",
    sep = "/"))
```

New data

This R code chunk was written to process the updated data.

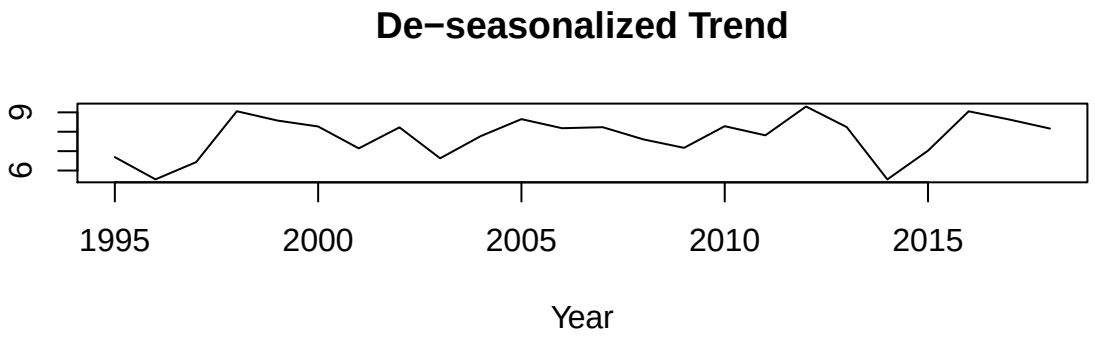
```
stl_data2 <- read_excel(paste(dataDIR, "updated_temperature_depth_data.xlsx",
    sep = "/"), col_names = F)
```

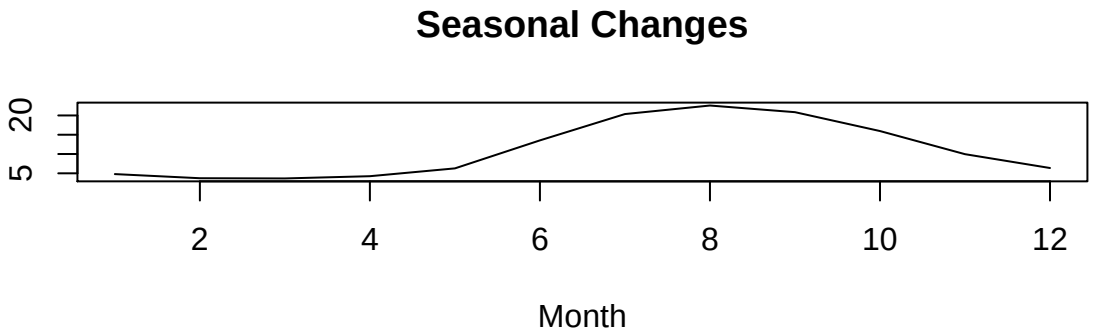
```
## New names:
## * `` -> ...1
## * `` -> ...2
## * `` -> ...3
## * `` -> ...4
## * `` -> ...5
## * ... and 9 more problems
```

```
Surface <- matrix(as.numeric(unlist(stl_data2[3:26, -c(1, 2)])), ncol = 12)
D30m <- matrix(as.numeric(unlist(stl_data2[30:58, -c(1, 2)])), ncol = 12)
D60m <- matrix(as.numeric(unlist(stl_data2[62:90, -c(1, 2)])), ncol = 12)
D75m <- matrix(as.numeric(unlist(stl_data2[94:122, -c(1, 2)])), ncol = 12)
```

```
D100 <- matrix(as.numeric(unlist(stl_data2[126:154, -c(1, 2)])), ncol = 12)
D110 <- matrix(as.numeric(unlist(stl_data2[158:186, -c(1, 2)])), ncol = 12)
D120 <- matrix(as.numeric(unlist(stl_data2[190:218, -c(1, 2)])), ncol = 12)
D140 <- matrix(as.numeric(unlist(stl_data2[222:250, -c(1, 2)])), ncol = 12)

Surface[Surface < -900] <- NA
D30m[D30m < -900] <- NA
D60m[D60m < -900] <- NA
D75m[D75m < -900] <- NA
D100[D100 < -900] <- NA
D110[D110 < -900] <- NA
D120[D120 < -900] <- NA
D140[D140 < -900] <- NA

SF_monthly <- median.polish.ts(ts(as.vector(t(Surface)), start = c(1995, 1),
    end = c(2018, 12), frequency = 12))
```

```
## 1: 292.151
## 2: 237.5483
## 3: 235.6352
## 4: 235.3183
## Final: 235.2718
## [1] "overall"   "row"       "col"       "residuals" "name"
```

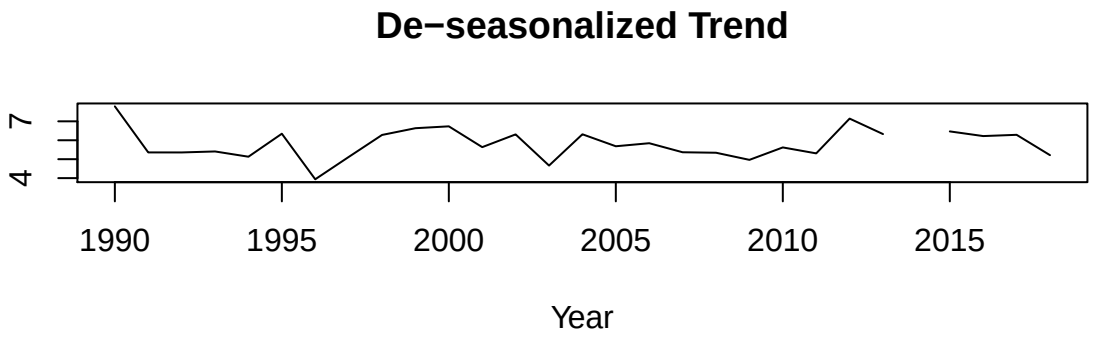


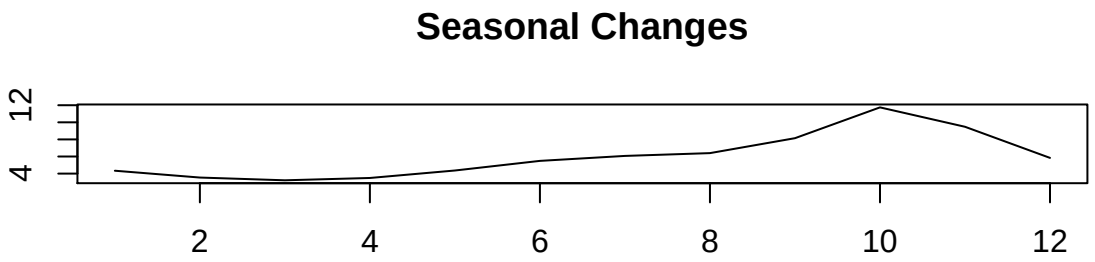



```
D30m_monthly <- median.polish.ts(ts(as.vector(t(D30m)), start = c(1990, 1),
    end = c(2018, 12), frequency = 12))
```

```
## 1: 218.8891
## 2: 183.5962
## 3: 181.8342
```

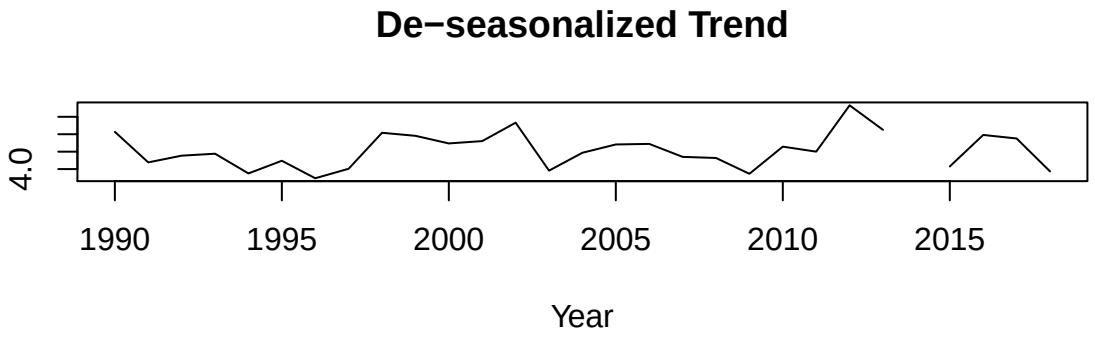
```
## 4: 181.4274
## 5: 181.2117
## Final: 181.0977
## [1] "overall"   "row"       "col"       "residuals" "name"
```

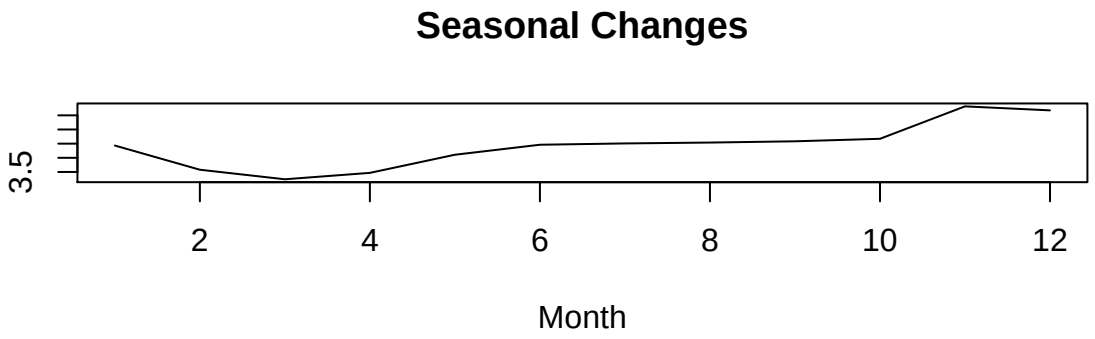





```
D60m_monthly <- median.polish.ts(ts(as.vector(t(D60m)), start = c(1990, 1),
    end = c(2018, 12), frequency = 12))
```

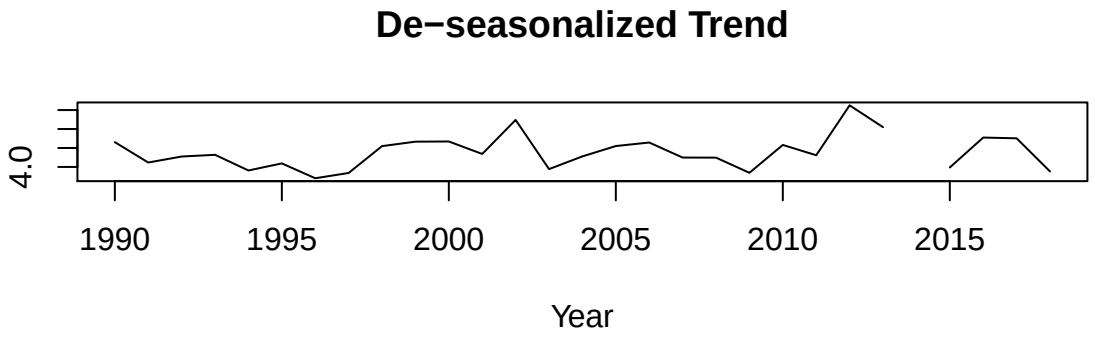
```
## 1: 78.9129
## 2: 71.62421
## 3: 71.29642
## 4: 71.21897
## Final: 71.18447
## [1] "overall"   "row"       "col"       "residuals" "name"
```

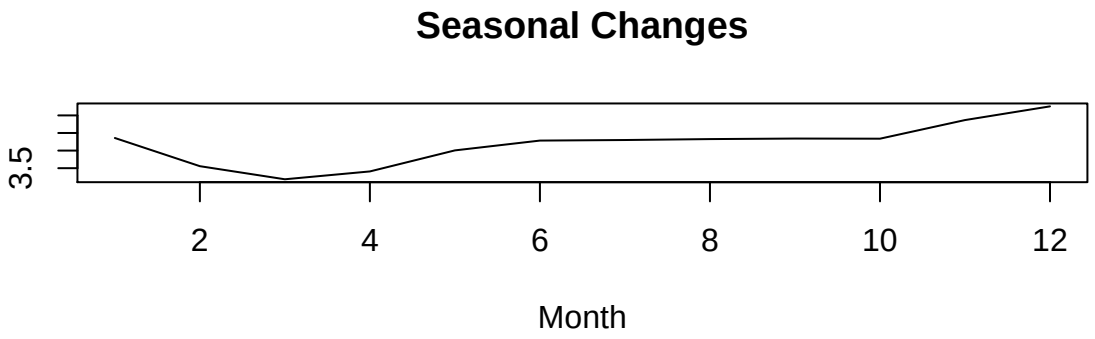





```
D75m_monthly <- median.polish.ts(ts(as.vector(t(D75m)), start = c(1990, 1),
    end = c(2018, 12), frequency = 12))
```

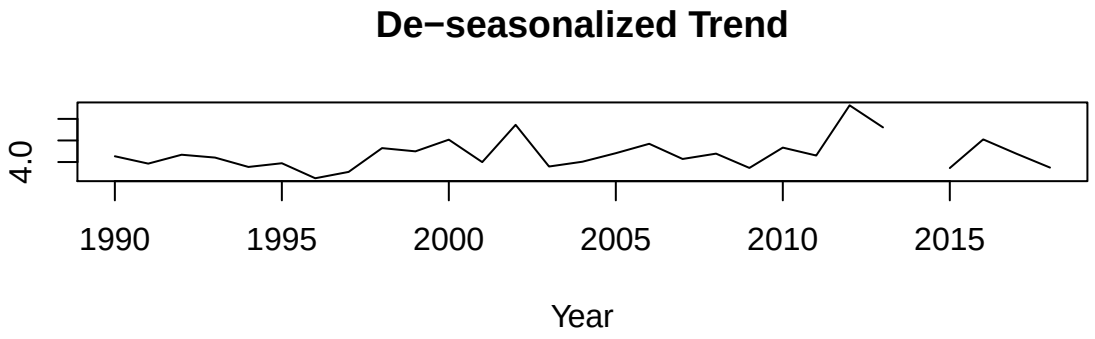
```
## 1: 55.43029
## 2: 51.43485
## 3: 51.31347
## Final: 51.303
## [1] "overall"   "row"       "col"       "residuals" "name"
```

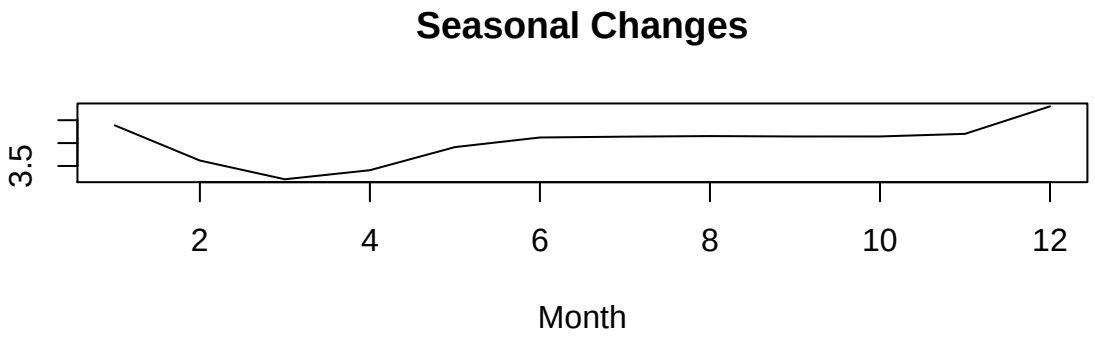





```
D100_monthly <- median.polish.ts(ts(as.vector(t(D100)), start = c(1990, 1),
    end = c(2018, 12), frequency = 12))
```

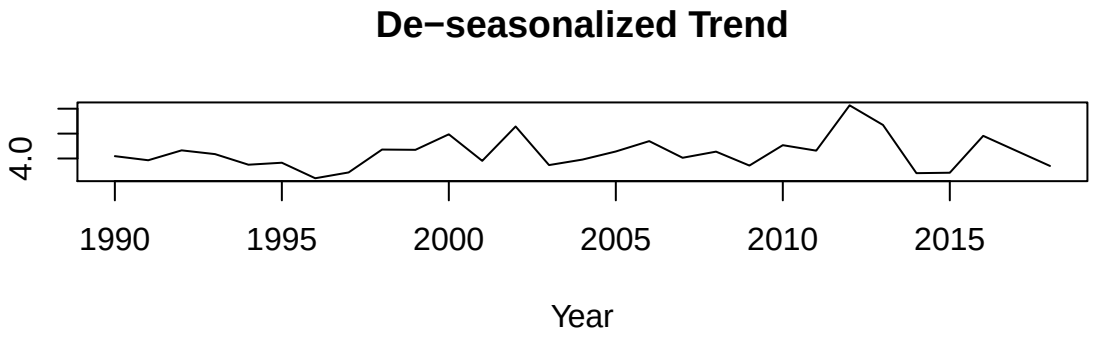
```
## 1: 45.10383
## 2: 43.04707
## 3: 42.97044
## Final: 42.95399
## [1] "overall"   "row"       "col"       "residuals" "name"
```

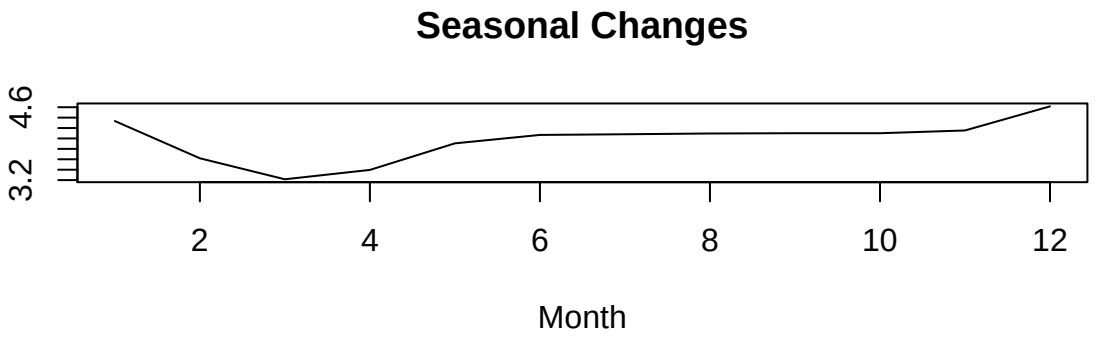





```
D110_monthly <- median.polish.ts(ts(as.vector(t(D110)), start = c(1990, 1),
    end = c(2018, 12), frequency = 12))
```

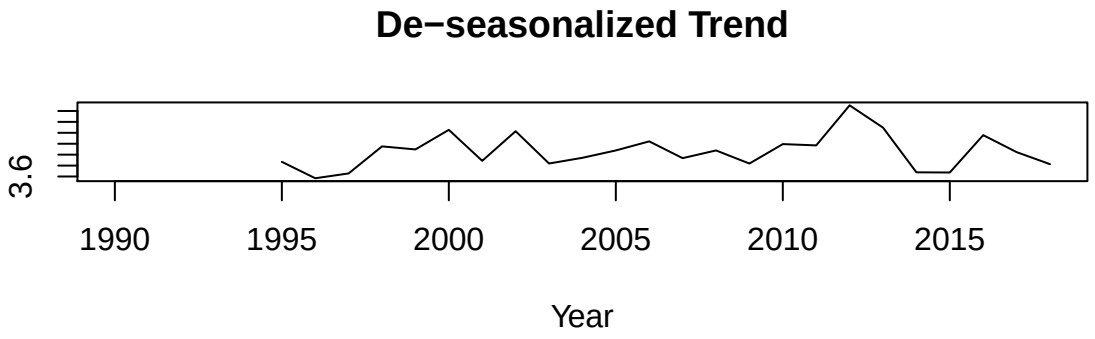
```
## 1: 51.46227
## 2: 50.17832
## 3: 50.10195
## Final: 50.0668
## [1] "overall"   "row"       "col"       "residuals" "name"
```

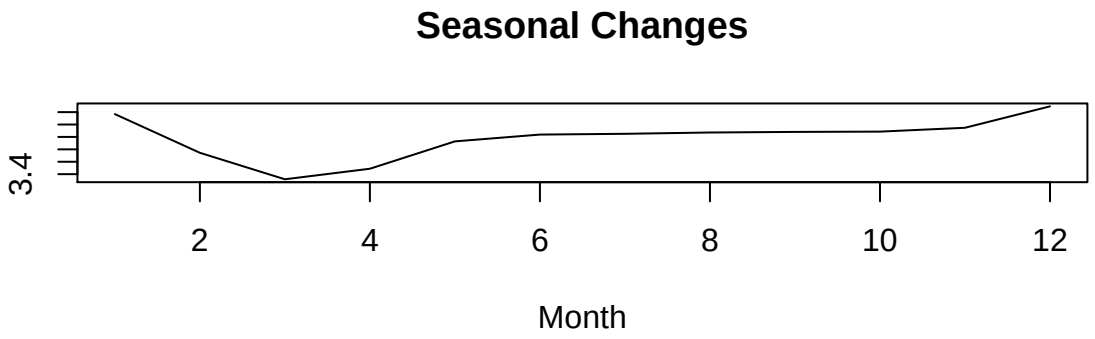





```
D120_monthly <- median.polish.ts(ts(as.vector(t(D120)), start = c(1990, 1),
    end = c(2018, 12), frequency = 12))
```

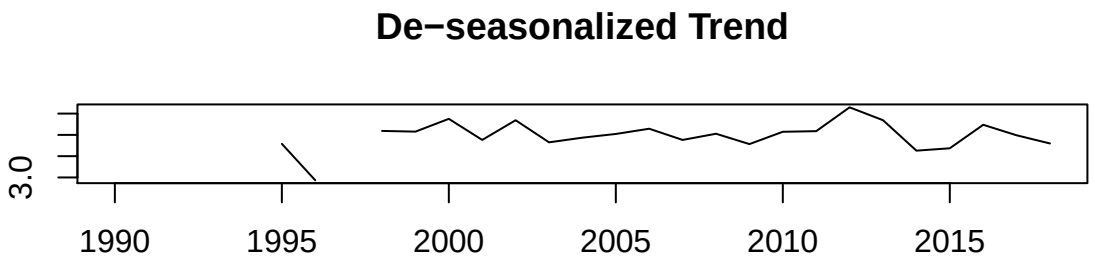
```
## 1: 43.45816
## 2: 41.96002
## 3: 41.89538
## Final: 41.88494
## [1] "overall"   "row"       "col"       "residuals" "name"
```

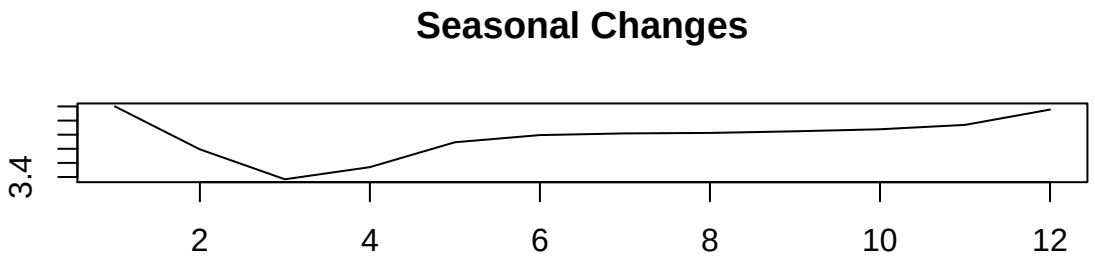





```
D140_monthly <- median.polish.ts(ts(as.vector(t(D140)), start = c(1990, 1),
    end = c(2018, 12), frequency = 12))
```

```
## 1: 34.26858
## 2: 32.2521
## 3: 32.21851
## Final: 32.21484
## [1] "overall"   "row"       "col"       "residuals" "name"
```






```
pdf(paste(figDIR, "SurfaceSTL.pdf", sep = "/"), height = 4, width = 5)
Surface.STL <- stl.rfs2(data = SF_monthly, aspect = 2.25, ylab = "Temperature",
    ss.w = 11, ss.d = 1, fc.w = 91, fc.d = 1)
dev.off()
```

```
## pdf
##   2
```

```
write.csv(Surface.STL[[3]], paste(dataOUTDIR, "surface_monthly.csv", sep = "/"))
write.csv(Surface.STL[[1]], paste(dataOUTDIR, "surface_components.csv", sep = "/"))

pdf(paste(figDIR, "D30mSTL.pdf", sep = "/"), height = 4, width = 5)
D30m.STL <- stl.rfs2(data = D30m_monthly, aspect = 2.25, ylab = "Temperature",
    ss.w = 11, ss.d = 1, fc.w = 91, fc.d = 1)
dev.off()
```

```
## pdf
##   2
```

```
write.csv(D30m.STL[[3]], paste(dataOUTDIR, "D30m_monthly.csv", sep = "/"))
write.csv(D30m.STL[[1]], paste(dataOUTDIR, "D30m_components.csv", sep = "/"))

pdf(paste(figDIR, "D60mSTL.pdf", sep = "/"), height = 4, width = 5)
D60m.STL <- stl.rfs2(data = D60m_monthly, aspect = 2.25, ylab = "Temperature",
    ss.w = 11, ss.d = 1, fc.w = 91, fc.d = 1)
dev.off()
```

```
## pdf
##   2
```

```
write.csv(D60m.STL[[3]], paste(dataOUTDIR, "D60m_monthly.csv", sep = "/"))
write.csv(D60m.STL[[1]], paste(dataOUTDIR, "D60m_components.csv", sep = "/"))
```

```
pdf(paste(figDIR, "D75mSTL.pdf", sep = "/"), height = 4, width = 5)
D75m.STL <- stl.rfs2(data = D75m_monthly, aspect = 2.25, ylab = "Temperature",
    ss.w = 11, ss.d = 1, fc.w = 91, fc.d = 1)
dev.off()
```

```
## pdf
##   2
```

```
write.csv(D75m.STL[[3]], paste(dataOUTDIR, "D75m_monthly.csv", sep = "/"))
write.csv(D75m.STL[[1]], paste(dataOUTDIR, "D75m_components.csv", sep = "/"))

pdf(paste(figDIR, "D100STL.pdf", sep = "/"), height = 4, width = 5)
D100.STL <- stl.rfs2(data = D100_monthly, aspect = 2.25, ylab = "Temperature",
    ss.w = 11, ss.d = 1, fc.w = 91, fc.d = 1)
dev.off()
```

```
## pdf
##   2
```

```
write.csv(D100.STL[[3]], paste(dataOUTDIR, "D100_monthly.csv", sep = "/"))
write.csv(D100.STL[[1]], paste(dataOUTDIR, "D100_components.csv", sep = "/"))

pdf(paste(figDIR, "D110STL.pdf", sep = "/"), height = 4, width = 5)
D110.STL <- stl.rfs2(data = D110_monthly, aspect = 2.25, ylab = "Temperature",
    ss.w = 11, ss.d = 1, fc.w = 91, fc.d = 1)
dev.off()
```

```
## pdf
##   2
```

```
write.csv(D110.STL[[3]], paste(dataOUTDIR, "D110_monthly.csv", sep = "/"))
write.csv(D110.STL[[1]], paste(dataOUTDIR, "D110_components.csv", sep = "/"))

pdf(paste(figDIR, "D120STL.pdf", sep = "/"), height = 4, width = 5)
D120.STL <- stl.rfs2(data = D120_monthly, aspect = 2.25, ylab = "Temperature",
    ss.w = 11, ss.d = 1, fc.w = 91, fc.d = 1)
dev.off()
```

```
## pdf
##   2
```

```
write.csv(D120.STL[[3]], paste(dataOUTDIR, "D120_monthly.csv", sep = "/"))
write.csv(D120.STL[[1]], paste(dataOUTDIR, "D120_components.csv", sep = "/"))

pdf(paste(figDIR, "D140STL.pdf", sep = "/"), height = 4, width = 5)
D140.STL <- stl.rfs2(data = D140_monthly, aspect = 2.25, ylab = "Temperature",
    ss.w = 11, ss.d = 1, fc.w = 91, fc.d = 1)
dev.off()
```

```
## pdf
##   2
```

```
write.csv(D140.STL[[3]], paste(dataOUTDIR, "D140_monthly.csv", sep = "/"))
write.csv(D140.STL[[1]], paste(dataOUTDIR, "D140_components.csv", sep = "/"))
```