FDC comparison

Willem Vervoort, Michaela Dolk & Floris van Ogtrop

2020-04-15

```
# root dir
knitr::opts_knit$set(root.dir = "D:/Cloudstor/Virtual Experiments/VirtExp")
#knitr::opts_knit$set(root.dir ="C:/Users/rver4657/ownCloud/Virtual Experiments/VirtExp")
knitr::opts_chunk$set(echo = TRUE)
# LOAD REQUIRED PACKAGES # #####
library(pander)
library(tidyverse)
library(xts)
library(zoo)
library(ggplot2)
library(reshape2)
```

This rmarkdown document and the resulting pdf are stored on github. All directories (apart from the root working directory) refer to the directories in this repository

Introduction

This document is related to the manuscript "Disentangling climate change trends in Australian streamflow" (vervoort et al.), submitted . This is the fourth document as part of the response to reviewers of earlier versions of the manuscript outlining in detail the generation of the flow duration curves. These figures have now been removed from the manuscript for a later submission.

The key idea in this analysis is that by comparing the flow duration curves by decade, we might be able to pick up where exactly the rainfall and the flow distribution have changed. An alternative analysis of the rainfall using double mass curves is also presented.

Load data and find decades

```
load("data/ClimCh_project_MD.Rdata")
# define decade start and end
decade_start <- c(as.Date("1/1/1970", format="%d/%m/%Y"),
                  as.Date("1/1/1980", format="%d/%m/%Y"),
                  as.Date("1/1/1990", format="%d/%m/%Y"),
                  as.Date("1/1/2000", format="%d/%m/%Y"))
decade_end <- c(as.Date("31/12/1979", format="%d/%m/%Y"),
                as.Date("31/12/1989", format="%d/%m/%Y"),
                as.Date("31/12/1999", format="%d/%m/%Y"),
                as.Date("31/12/2009", format="%d/%m/%Y"))
```

The next step is to assign the decades to the daily data as the weekly data already have associated decades. In this document we will plot both the weekly and the daily flow duration curves (FDC) to check if aggregating to weekly makes a difference.

```r
flow_decade <- vector("character", length=nrow(flow_zoo))
decades <- c("70-80","80-90", "90-00", "00-10")
for(i in 1:4) {
  flow_decade[time(flow_zoo) >= decade_start[i] &
                  time(flow_zoo) <= decade_end[i]] <- decades[i]
}
```

Flow duration curves for the rainfall and the flow data

Flow durations curves are essentially a frequency curve and therefor we can apply the same analysis to the rainfall and the flow data. The best would be to use all the data and rank this to generate the FDC, but this is not possible as there are different numbers of NA values. So here we use the function `quantile()` to generate the distributions. For the daily data we generate a series of 1000 samples, while for the weekly data, we generate 100 samples.

flow data

```r
# combine flow_zoo with flow_decade in a data.frame
flow_df <- data.frame(decade=flow_decade,
                        as.data.frame(flow_zoo))
# now calculate the flow duration curve for each catchment by decade
# function to generate FDC, use quantile decades have different data lenghts (leap years)
FDC_gen <- function(DATA) {
   FDC <- data.frame(probs = seq(0,1,length=1000)*100,
                     flow = quantile(DATA,probs=seq(0,1,length=1000),
                                     na.rm=T))
    return(FDC)
}

# tapply this across the columns and produce a list
# do this for each decade
# an empty list for the decades
FDCs <- list()
for (i in 1:4) {
  FDCs[[i]] <- apply(subset(flow_df,
                as.character(flow_df$decade)==
                  as.character(decades[i]))[,2:14],
                      2,FDC_gen)
}

# DIFFERENCE BETWEEN daily flow FDCs DIVIDED BY 1970-1979 FDCs -> not log, ylim=c(-1,1) to prevent issu

periods <- c("((1970-1979) - (1980-1989))/(1970-1979)",
              "((1970-1979) - (1990-1999))/(1970-1979)",
              "((1970-1979) - (2000-2009))/(1970-1979)")

plot.list <- vector("list", length=13)
temp <- list()
for (j in 1:length(periods)) {
  for (i in seq_along(Stations[,1])) {
```

```
    temp[[i]] <- (FDCs[[1]][[i]]$flow -
                        FDCs[[j+1]][[i]]$flow)#/mean(FDCs[[1]][[i]]$flow,na.rm=T)
  }
  fdc <- do.call(cbind,temp)
  colnames(fdc) <- Stations[,1]
  plot.list[[j]] <- melt(data.frame(fdc = fdc,
                                    period = rep(periods[j],nrow(fdc))),
                                    id.vars="period")
}
# bring back together to a dataframe
plot.df <- do.call(rbind,plot.list)
# add the probabilities
plot.df <- cbind(plot.df,probs=rep(FDCs[[1]][[1]]$probs,13*3))
colnames(plot.df)[2] <- "station"
levels(plot.df$station) <- Stations[,1]

# make a plot
p <- ggplot(plot.df, aes(x = probs, y = value)) +
  geom_line(aes(linetype=period, colour=period),size=1.2) +
  facet_wrap(~ station,ncol=5) + ylim(c(-2,2)) +
  theme(legend.position="bottom") +
  guides(col = guide_legend(nrow = 3))
p <- p + ggtitle("Streamflow") +
  theme(plot.title = element_text(lineheight=.8, face="bold")) +
  xlab("Probability") +
  theme(axis.title.x = element_text(face="bold",  size=16),
        axis.text.x  = element_text(size=12)) +
  ylab("Scaled Difference") +
  theme(axis.title.y = element_text(face="bold",  size=16),
        axis.text.y  = element_text(size=12)) +
 scale_colour_manual(values=c("black", "gray33", "gray66")) +
  theme(legend.text = element_text( size = 12))+
  theme(legend.title = element_text(size=14, face="bold")) +
  theme(strip.text.x = element_text(size=16))
print(p)
```

```
## Warning: Removed 13 rows containing missing values (geom_path).
```

```
# # publication quality
# tiff("../manuscript/Figure6_StreamflowFDCDifference.tif",
#      width=16*480, height=12*480,
#      res=600, compression="lzw")
# print(p)
# dev.off()
```

daily rainfall data

```
# combine flow_zoo with flow_decade in a data.frame
rain_df <- data.frame(decade=flow_decade,
                          as.data.frame(rain_zoo))
```

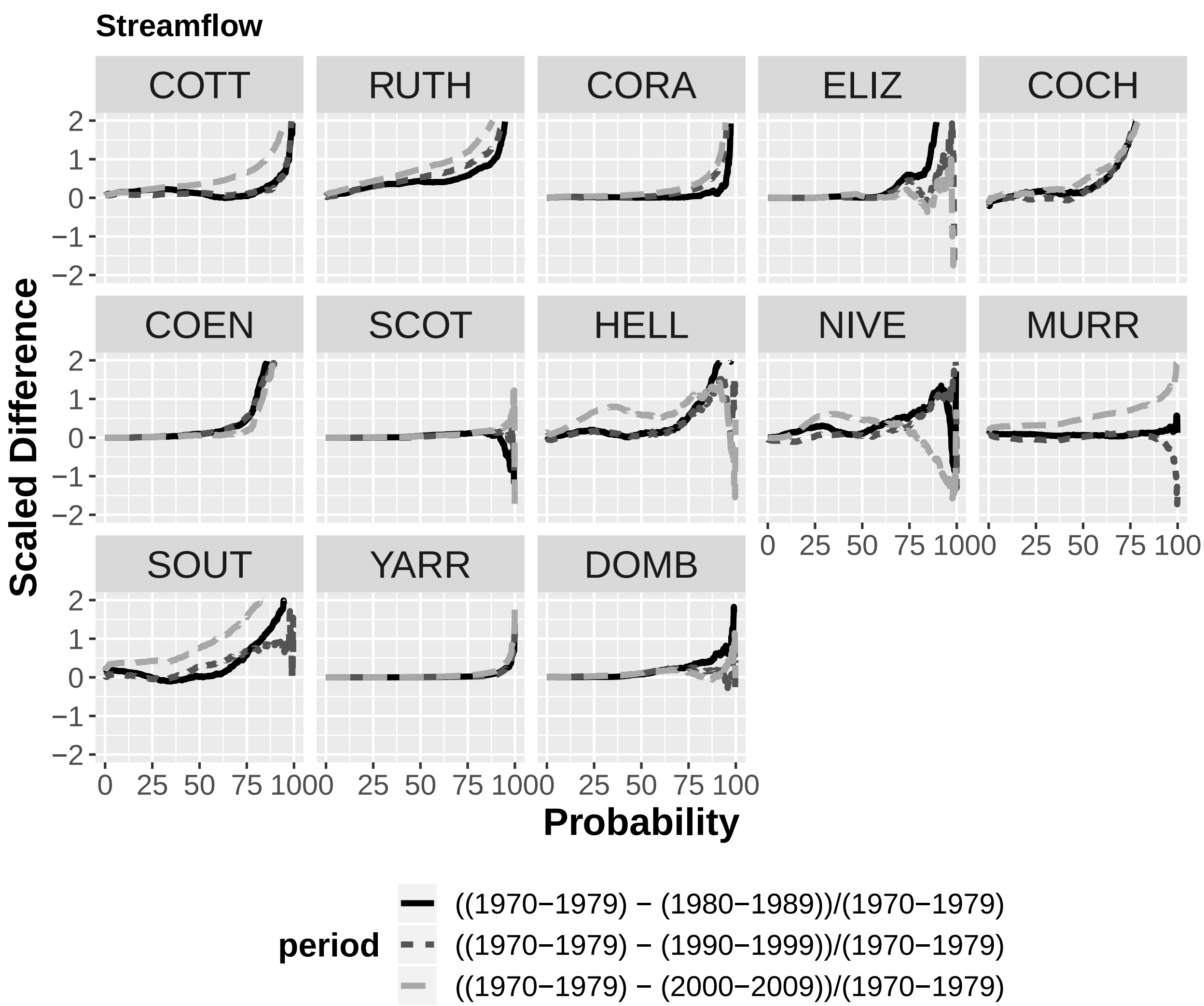

Figure 1: Difference in flow duration curves based on daily data split by decade

```r
# now calculate the rainfall duration curve for each catchment by decade
# tapply this across the columns and produce a list
# do this for each decade
# an empty list for the decades
FDCs <- list()
for (i in 1:4) {
  FDCs[[i]] <- apply(subset(rain_df,
      as.character(rain_df$decade)==as.character(decades[i]))[,2:14],
                        2,FDC_gen)
}

# DIFFERENCE BETWEEN rainfall FDCs DIVIDED BY 1970-1979 FDCs -> not log, ylim=c(-1,1) to prevent issues
plot.list <- vector("list", length=13)
temp <- list()
for (j in 1:length(periods)) {
  for (i in seq_along(Stations[,1])) {
    temp[[i]] <- (FDCs[[1]][[i]]$flow -
                    FDCs[[j+1]][[i]]$flow)#/mean(FDCs[[1]][[i]]$flow,
                                                # na.rm=T)
  }
  fdc <- do.call(cbind,temp)
  colnames(fdc) <- Stations[,1]
  plot.list[[j]] <- melt(data.frame(fdc = fdc,
                                  period = rep(periods[j],nrow(fdc))),
                                  id.vars="period")
}
# bring back together to a dataframe
plot_df_r <- do.call(rbind,plot.list)
# add the probabilities
plot_df_r <- cbind(plot_df_r,probs=rep(FDCs[[1]][[1]]$probs,13*3))
colnames(plot_df_r)[2] <- "station"
levels(plot_df_r$station) <- Stations[,1]

# make a plot
p <- ggplot(plot_df_r, aes(x = probs, y = value)) +
  geom_line(aes(linetype=period, colour=period),size=1.2) +
  facet_wrap(~ station,ncol=5) + ylim(c(-2,2)) +
  theme(legend.position="bottom") +
  guides(col = guide_legend(nrow = 3))
p <- p + ggtitle("Rainfall") +
  theme(plot.title = element_text(lineheight=.8, face="bold"))+
  xlab("Probability") +
  theme(axis.title.x = element_text(face="bold",  size=16),
        axis.text.x  = element_text(size=12)) +
  ylab("Scaled Difference") +
  theme(axis.title.y = element_text(face="bold",  size=16),
        axis.text.y  = element_text(size=12)) +
 scale_colour_manual(values=c("black", "gray33", "gray66")) +
  theme(legend.text = element_text( size = 12))+
  theme(legend.title = element_text(size=14, face="bold")) +
  theme(strip.text.x = element_text(size=16))
print(p)
```

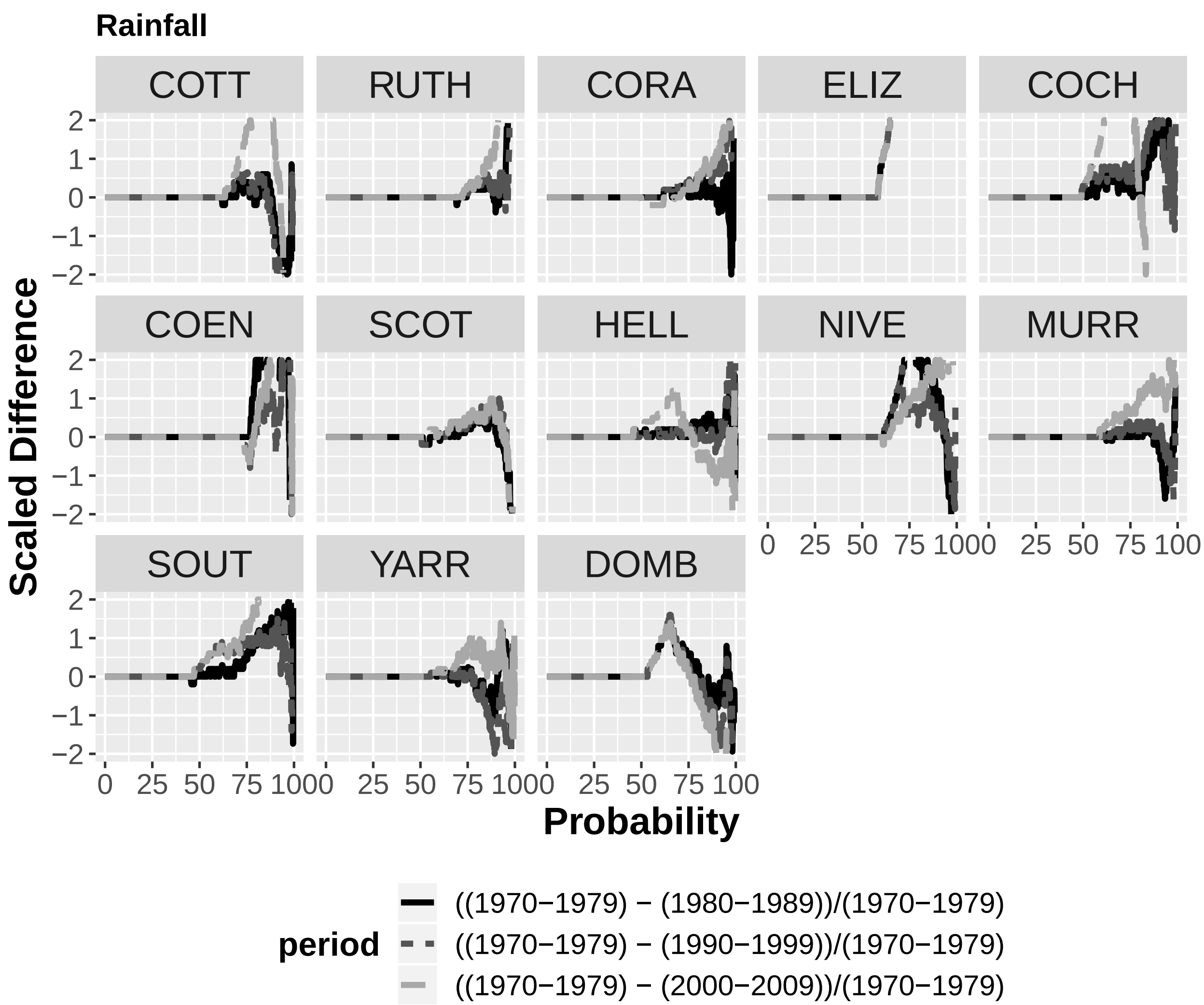

Figure 2: Difference in rainfal duration curves based on daily data split by decade

```
## Warning: Removed 64 rows containing missing values (geom_path).
```

daily gridded rainfall

```
# gridded rainfall
grrain_df <- data.frame(decade=flow_decade,
                        as.data.frame(gridRain_zoo))

# now calculate the rainfall duration curve for each catchment by decade
# tapply this across the columns and produce a list
# do this for each decade
# an empty list for the decades
FDCs <- list()
for (i in 1:4) {
```

```r
  FDCs[[i]] <- apply(subset(grrain_df,
      as.character(grrain_df$decade)==as.character(decades[i]))[,2:14],
                      2,FDC_gen)
}

# plotting
plot.list <- vector("list", length=13)
temp <- list()
for (j in 1:length(periods)) {
  for (i in seq_along(Stations[,1])) {
    temp[[i]] <- (FDCs[[1]][[i]]$flow -
                    FDCs[[j+1]][[i]]$flow)#/mean(FDCs[[1]][[i]]$flow,
                                                #na.rm=T)
  }
  fdc <- do.call(cbind,temp)
  colnames(fdc) <- Stations[,1]
  plot.list[[j]] <- melt(data.frame(fdc = fdc,
                               period = rep(periods[j],nrow(fdc))),
                               id.vars="period")
}
# bring back together to a dataframe
plot_df_r <- do.call(rbind,plot.list)
# add the probabilities
plot_df_r <- cbind(plot_df_r,probs=rep(FDCs[[1]][[1]]$probs,13*3))
colnames(plot_df_r)[2] <- "station"
levels(plot_df_r$station) <- Stations[,1]

# make a plot
p <- ggplot(plot_df_r, aes(x = probs, y = value)) +
  geom_line(aes(linetype=period, colour=period),size=1.2) +
  facet_wrap(~ station,ncol=5) + ylim(c(-2,2)) +
  theme(legend.position="bottom") +
  guides(col = guide_legend(nrow = 3))
p <- p + ggtitle("gridded Rainfall") +
  theme(plot.title = element_text(lineheight=.8, face="bold"))+
  xlab("Probability") +
  theme(axis.title.x = element_text(face="bold",  size=16),
        axis.text.x  = element_text(size=12)) +
  ylab("Scaled Difference") +
  theme(axis.title.y = element_text(face="bold",  size=16),
        axis.text.y  = element_text(size=12)) +
 scale_colour_manual(values=c("black", "gray33", "gray66")) +
  theme(legend.text = element_text( size = 12))+
  theme(legend.title = element_text(size=14, face="bold")) +
  theme(strip.text.x = element_text(size=16))
print(p)
```

```
## Warning: Removed 4 rows containing missing values (geom_path).
```

```r
# #publication quality
# tiff("../manuscript/Figure7_RainfallFDCDifference.tif",
#      width=16*480, height=12*480,
```

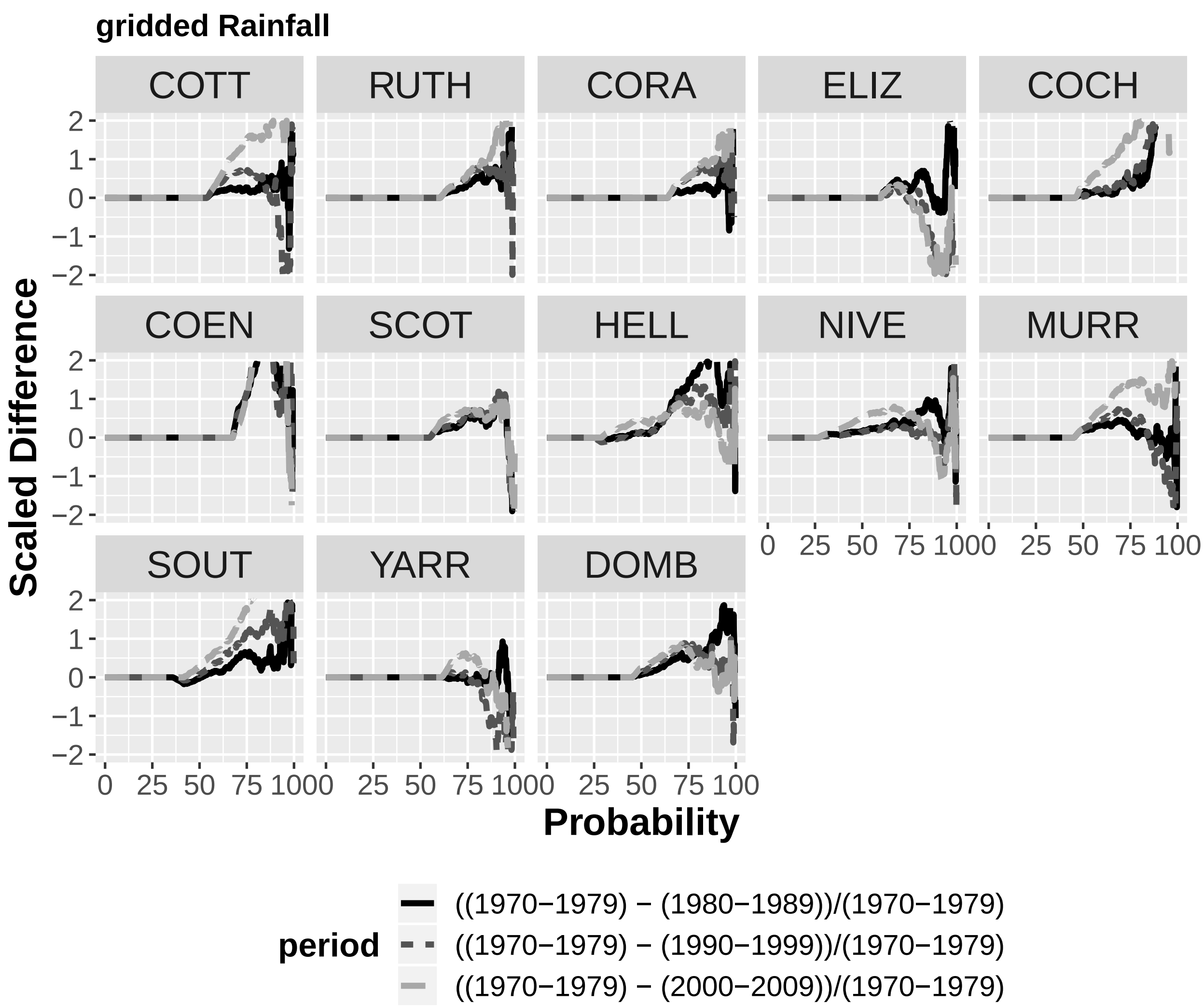

Figure 3: Difference in gridded rainfal duration curves based on daily gridded data split by decade

```
#      res=600, compression="lzw")
# print(p)
# dev.off()
```

These results indicate differences between the distributions, but there is no immediate consistent pattern between the different decades or between the flow and the rainfall patterns. So while there could be changes in the distribution of the rainfall over time. There is no direct indication that this has consistently changed the distribution of the flow in time (at least not on the daily time scale).

Weekly data Flow Duration Curves

flow data

```
# Calculate the flow duration curve for each catchment by decade
# function to generate FDC, use quantile decades have different data lenghts (leap years)
FDC_gen_w <- function(DATA) {
   FDC <- data.frame(probs = seq(0,1,length=100)*100,
                     flow = quantile(DATA,probs=seq(0,1,length=100),
                                     na.rm=T))
    return(FDC)
}

# tapply this across the columns and produce a list
# do this for each decade
decades <- c("70_80", "80_90", "90_00", "00_10")
# an empty list for the decades
FDCs <- list()
for (i in 1:4) {
  temp <- subset(flow_rain_maxT_weekly,
                        as.character(flow_rain_maxT_weekly$Decade)==
                        as.character(decades[i]))
  FDCs[[i]] <- tapply(temp[,"Flow"],
                      list(Station=temp$Station),FDC_gen_w)
}

# plotting
plot.list <- vector("list", length=13)
temp <- list()
for (j in 1:length(periods)) {
  for (i in seq_along(Stations[,1])) {
    temp[[i]] <- (FDCs[[1]][[i]]$flow -
                    FDCs[[j+1]][[i]]$flow)#/mean(FDCs[[1]][[i]]$flow,
                                              #   na.rm=T)
  }
  fdc <- do.call(cbind,temp)
  colnames(fdc) <- Stations[,1]
  plot.list[[j]] <- melt(data.frame(fdc = fdc,
                                period = rep(periods[j],nrow(fdc))),
                                id.vars="period")
}
```

```r
# bring back together to a dataframe
plot.df <- do.call(rbind,plot.list)
# add the probabilities
plot.df <- cbind(plot.df,probs=rep(FDCs[[1]][[1]]$probs,13*3))
colnames(plot.df)[2] <- "station"
levels(plot.df$station) <- Stations[,1]

# make a plot
p <- ggplot(plot.df, aes(x = probs, y = value)) +
  geom_line(aes(linetype=period, colour=period),size=1.2) +
  facet_wrap(~ station,ncol=5) + ylim(c(-2,2)) +
  theme(legend.position="bottom") +
  guides(col = guide_legend(nrow = 3))
p <- p + ggtitle("Weekly Streamflow") +
  theme(plot.title = element_text(lineheight=.8, face="bold")) +
  xlab("Probability") +
  theme(axis.title.x = element_text(face="bold",  size=16),
        axis.text.x  = element_text(size=12)) +
  ylab("Scaled Difference") +
  theme(axis.title.y = element_text(face="bold",  size=16),
        axis.text.y  = element_text(size=12)) +
 scale_colour_manual(values=c("black", "gray33", "gray66")) +
  theme(legend.text = element_text( size = 12))+
  theme(legend.title = element_text(size=14, face="bold")) +
  theme(strip.text.x = element_text(size=16))
print(p)
```

```
## Warning: Removed 12 rows containing missing values (geom_path).
```

```r
# # publication quality
# tiff("../manuscript/Figure6_StreamflowFDCDifference.tif",
#      width=16*480, height=12*480,
#      res=600, compression="lzw")
# print(p)
# dev.off()
```

##Weekly Rainfall

```r
# Calculate the rainfall duration curve for each catchment by decade
for (i in 1:4) {
  temp <- subset(flow_rain_maxT_weekly,
                        as.character(flow_rain_maxT_weekly$Decade)==
                        as.character(decades[i]))
  FDCs[[i]] <- tapply(temp[,"Rain"],
                      list(Station=temp$Station),FDC_gen_w)
}

# plotting
plot.list <- vector("list", length=13)
temp <- list()
for (j in 1:length(periods)) {
  for (i in seq_along(Stations[,1])) {
```

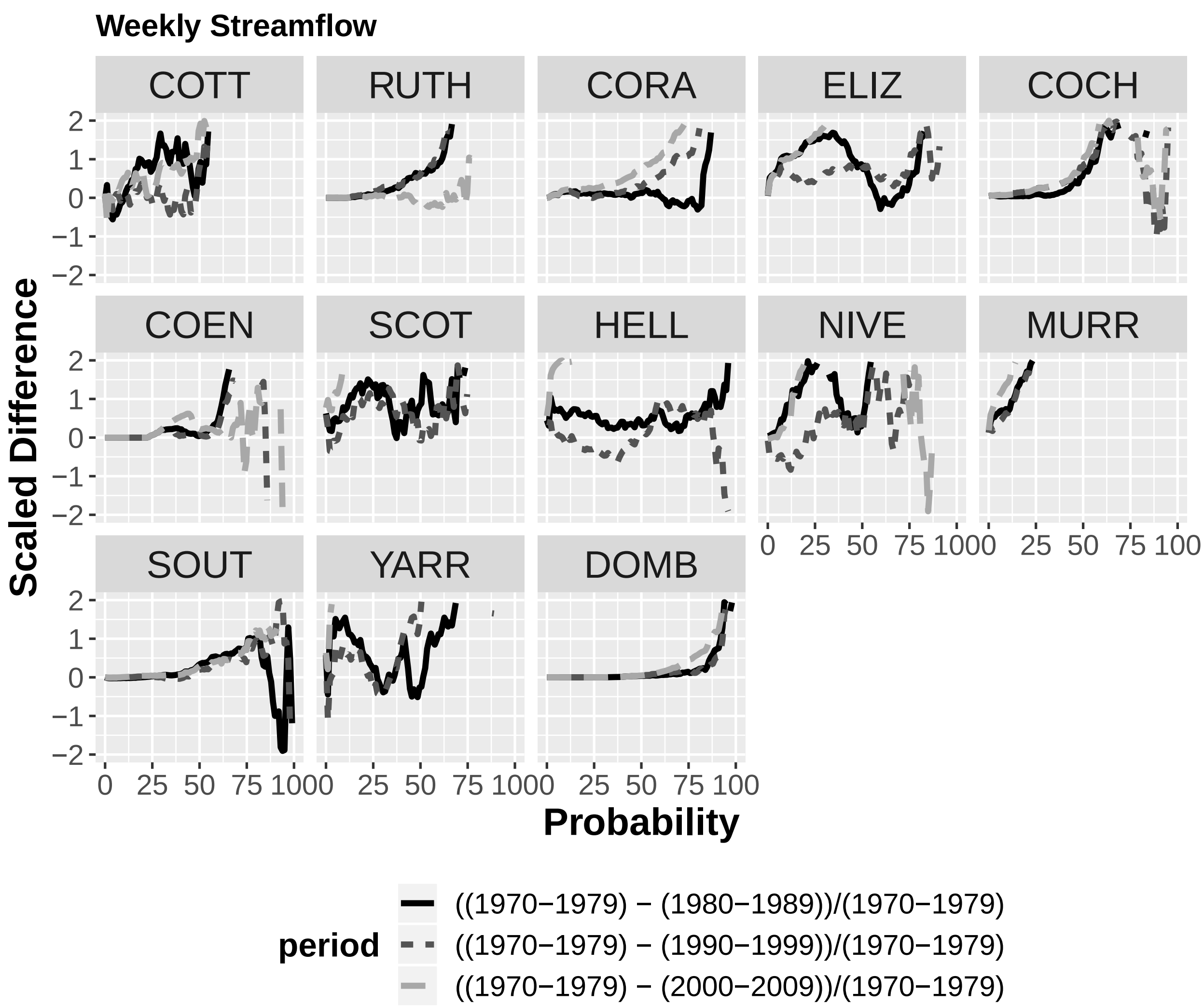

Figure 4: Difference in flow duration curves based on weekly data split by decade

```
    temp[[i]] <- (FDCs[[1]][[i]]$flow -
                    FDCs[[j+1]][[i]]$flow)#/mean(FDCs[[1]][[i]]$flow,
                                                #    na.rm=T)
  }
  fdc <- do.call(cbind,temp)
  colnames(fdc) <- Stations[,1]
  plot.list[[j]] <- melt(data.frame(fdc = fdc,
                                period = rep(periods[j],nrow(fdc))),
                                id.vars="period")
}
# bring back together to a dataframe
plot.df <- do.call(rbind,plot.list)
# add the probabilities
plot.df <- cbind(plot.df,probs=rep(FDCs[[1]][[1]]$probs,13*3))
colnames(plot.df)[2] <- "station"
levels(plot.df$station) <- Stations[,1]

# make a plot
p <- ggplot(plot.df, aes(x = probs, y = value)) +
  geom_line(aes(linetype=period, colour=period),size=1.2) +
  facet_wrap(~ station,ncol=5) + ylim(c(-2,2)) +
  theme(legend.position="bottom") +
  guides(col = guide_legend(nrow = 3))
p <- p + ggtitle("Weekly Rainfall") +
  theme(plot.title = element_text(lineheight=.8, face="bold")) +
  xlab("Probability") +
  theme(axis.title.x = element_text(face="bold",  size=16),
        axis.text.x  = element_text(size=12)) +
  ylab("Scaled Difference") +
  theme(axis.title.y = element_text(face="bold",  size=16),
        axis.text.y  = element_text(size=12)) +
 scale_colour_manual(values=c("black", "gray33", "gray66")) +
  theme(legend.text = element_text( size = 12))+
  theme(legend.title = element_text(size=14, face="bold")) +
  theme(strip.text.x = element_text(size=16))
print(p)
```

```
## Warning: Removed 6 rows containing missing values (geom_path).
```

```
# # publication quality
# tiff("../manuscript/Figure6_StreamflowFDCDifference.tif",
#      width=16*480, height=12*480,
#      res=600, compression="lzw")
# print(p)
# dev.off()
```

Weekly gridded rainfall

```
# Calculate the gridrainfall duration curve for each catchment by decade
for (i in 1:4) {
```

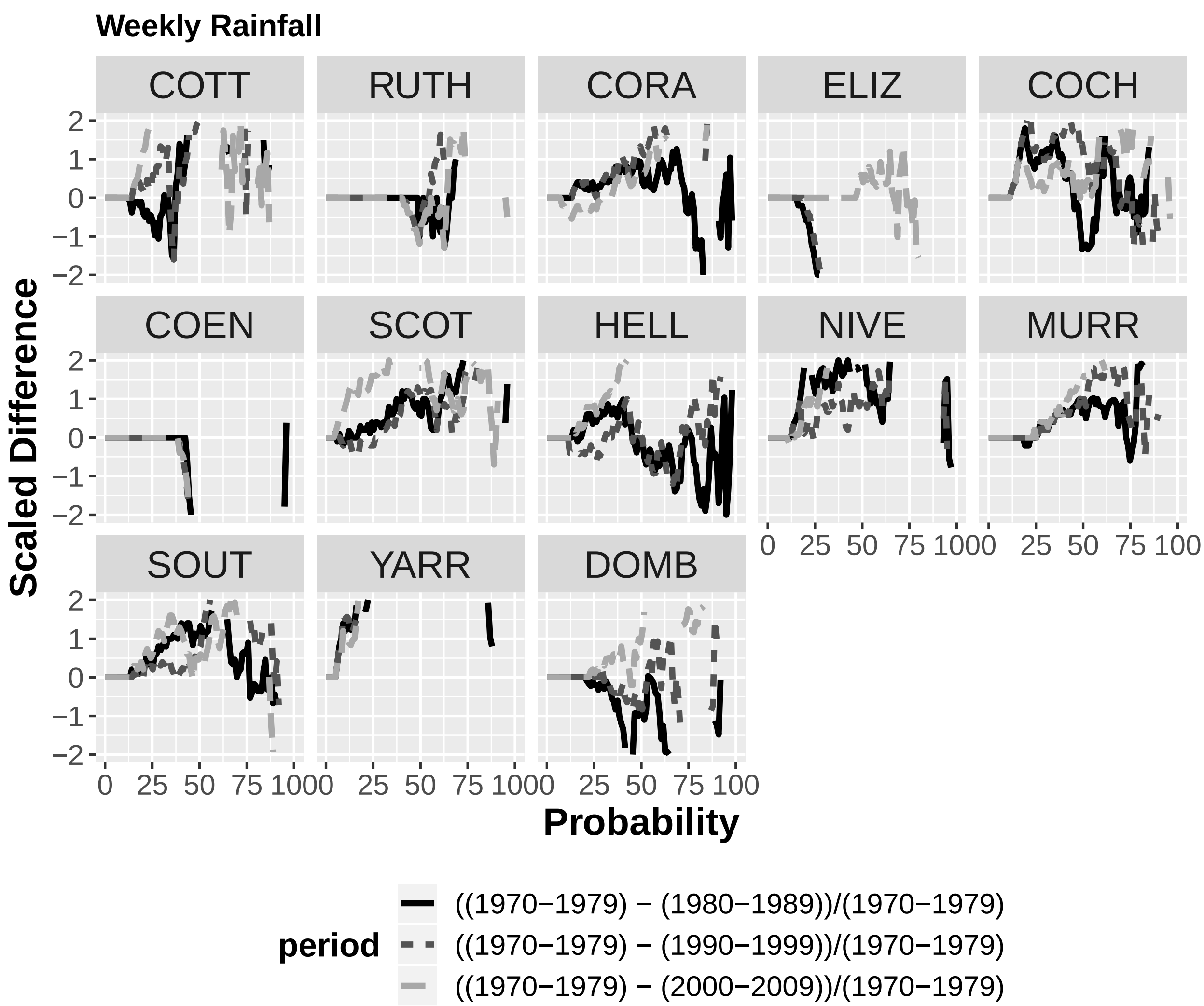

Figure 5: Difference in rainfall duration curves based on weekly data split by decade

```r
  temp <- subset(flow_rain_maxT_weekly,
                      as.character(flow_rain_maxT_weekly$Decade)==
                      as.character(decades[i]))
  FDCs[[i]] <- tapply(temp[,"gridRain"],
                     list(Station=temp$Station),FDC_gen_w)
}

# plotting
plot.list <- vector("list", length=13)
temp <- list()
for (j in 1:length(periods)) {
  for (i in seq_along(Stations[,1])) {
    temp[[i]] <- (FDCs[[1]][[i]]$flow -
                  FDCs[[j+1]][[i]]$flow)#/mean(FDCs[[1]][[i]]$flow,
                                        #     na.rm=T)
  }
  fdc <- do.call(cbind,temp)
  colnames(fdc) <- Stations[,1]
  plot.list[[j]] <- melt(data.frame(fdc = fdc,
                              period = rep(periods[j],nrow(fdc))),
                              id.vars="period")
}
# bring back together to a dataframe
plot.df <- do.call(rbind,plot.list)
# add the probabilities
plot.df <- cbind(plot.df,probs=rep(FDCs[[1]][[1]]$probs,13*3))
colnames(plot.df)[2] <- "station"
levels(plot.df$station) <- Stations[,1]

# make a plot
p <- ggplot(plot.df, aes(x = probs, y = value)) +
  geom_line(aes(linetype=period, colour=period),size=1.2) +
  facet_wrap(~ station,ncol=5) + ylim(c(-2,2)) +
  theme(legend.position="bottom") +
  guides(col = guide_legend(nrow = 3))
p <- p + ggtitle("Gridded Rainfall") +
  theme(plot.title = element_text(lineheight=.8, face="bold")) +
  xlab("Probability") +
  theme(axis.title.x = element_text(face="bold",  size=16),
        axis.text.x  = element_text(size=12)) +
  ylab("Scaled Difference") +
  theme(axis.title.y = element_text(face="bold",  size=16),
        axis.text.y  = element_text(size=12)) +
 scale_colour_manual(values=c("black", "gray33", "gray66")) +
  theme(legend.text = element_text( size = 12))+
  theme(legend.title = element_text(size=14, face="bold")) +
  theme(strip.text.x = element_text(size=16))
print(p)
```

```
## Warning: Removed 12 rows containing missing values (geom_path).
```

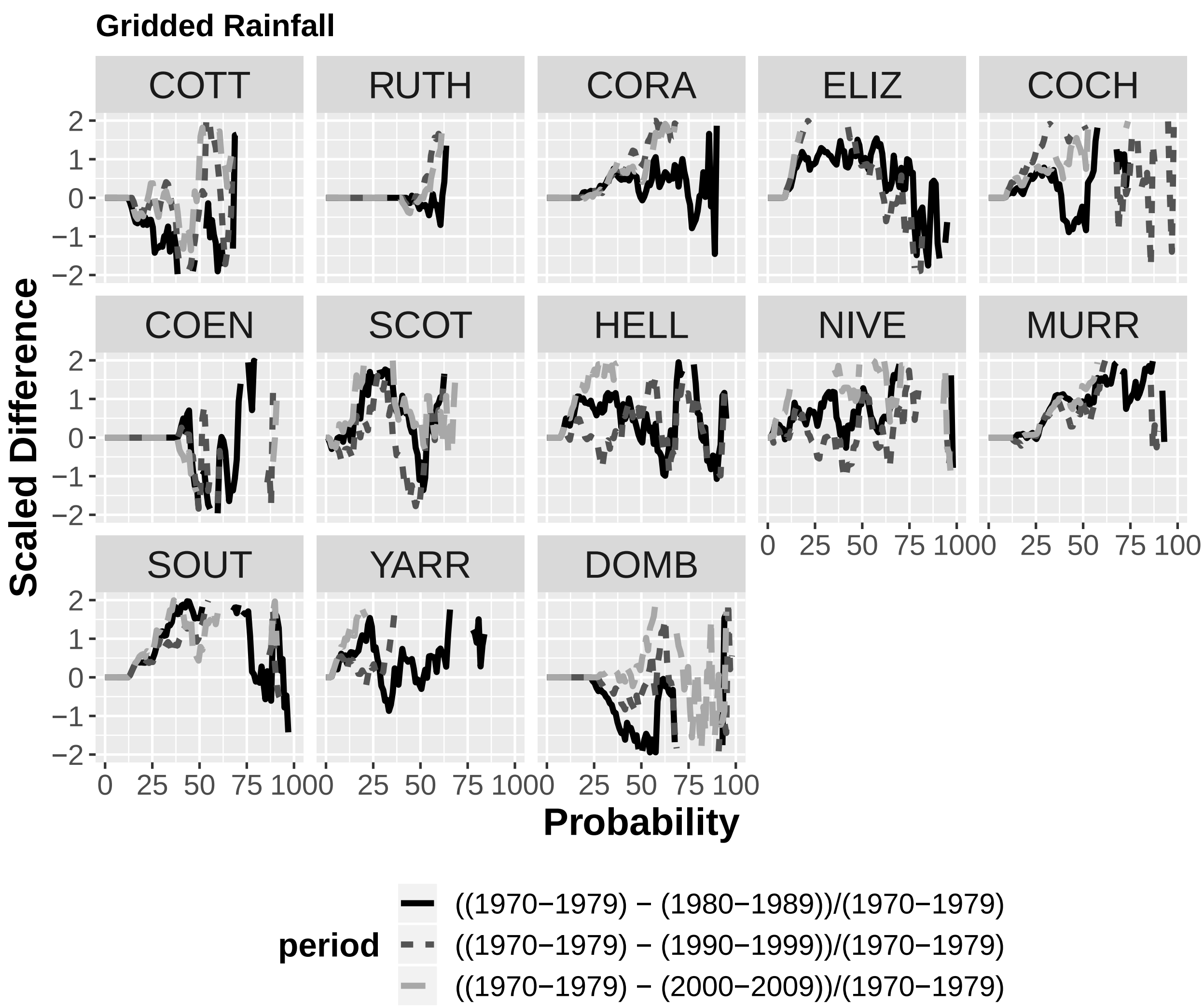

Figure 6: Difference in gridded rainfall duration curves based on weekly data split by decade

```r
# # publication quality
# tiff("../manuscript/Figure6_StreamflowFDCDifference.tif",
#      width=16*480, height=12*480,
#      res=600, compression="lzw")
# print(p)
# dev.off()
```

Apart from the fact that the weekly data are less noisy than the daily data, there is little difference between the weekly data analysis and the daily flow duration curve analysis. There is no consistent trend in the changes in rainfall or flow distributions.

#Double mass curves
Finally as a comparison to the duration curves, we can run douoble mass curves for the daily flow and rainfall data. This involves plotting the cumulative sums of the data for one decade against the cumulative sums for another decade.

flow data

```r
# Calculate the double mass curve for each catchment by decade
# apply cumulative sum across the columns and produce a list
# do this for each decade
# set all missing values to 0
flow_df2 <- flow_df
flow_df2[is.na(flow_df2)==T] <- 0
decades <- c("70-80","80-90", "90-00", "00-10")

MDCs <- list()
for (i in 1:4) {
  MDCs[[i]] <- apply(subset(flow_df2,
               as.character(flow_df2$decade)==
                 as.character(decades[i]))[,2:14],
                   2,cumsum)
}

# plotting
plot.list <- vector("list", length=3)
for (j in seq_along(periods)) {
  mdc = melt(MDCs[[j+1]][1:3652,])[,2:3]
  plot.list[[j]] <- cbind(mdc=mdc, period = rep(periods[j],3652*13))
  colnames(  plot.list[[j]]) <- c("Station", "cumulative_flow","period")
  levels(  plot.list[[j]]$Station) <- Stations[,1]
}
# bring back together to a dataframe
plot.df <- do.call(rbind,plot.list)
# add the the first decade as x-axis
plot.df <- cbind(plot.df,
               base_decade=rep(melt(MDCs[[1]][1:3652,])[,3],3))

# make a plot
p <- ggplot(plot.df, aes(x = base_decade, y = cumulative_flow)) +
  geom_line(aes(linetype=period, colour=period),size=1.2) +
  facet_wrap(~ Station,ncol=5, scales="free") +
```

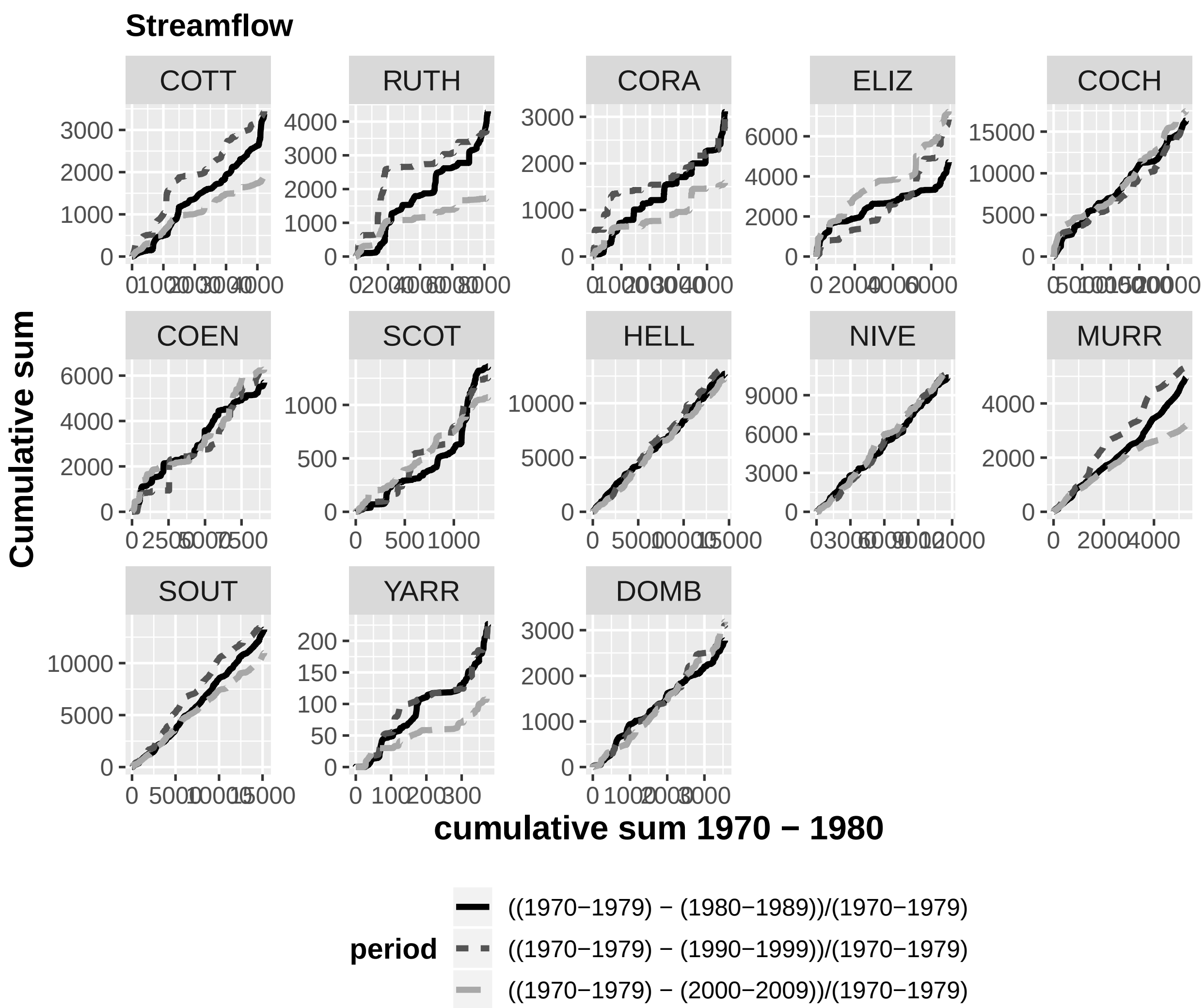

Figure 7: Double mass curves based on daily flow data split by decade

```
  theme(legend.position="bottom") +
  guides(col = guide_legend(nrow = 3))
p <- p + ggtitle("Streamflow") +
  theme(plot.title = element_text(lineheight=.8, face="bold")) +
  xlab("cumulative sum 1970 - 1980") +
  theme(axis.title.x = element_text(face="bold",  size=14),
        axis.text.x  = element_text(size=10)) +
  ylab("Cumulative sum") +
  theme(axis.title.y = element_text(face="bold",  size=14),
        axis.text.y  = element_text(size=10)) +
 scale_colour_manual(values=c("black", "gray33", "gray66")) +
 theme(legend.text = element_text( size = 10))+
  theme(legend.title = element_text(size=12, face="bold")) +
  theme(strip.text.x = element_text(size=12))
print(p)
```

```
# # publication quality
# tiff("../manuscript/Figure6_StreamflowFDCDifference.tif",
#      width=16*480, height=12*480,
#      res=600, compression="lzw")
# print(p)
# dev.off()
```

Double mass rainfall

```
# Calculate the double mass curve for each catchment by decade
# apply cumulative sum across the columns and produce a list
# do this for each decade
# set all missing values to 0
rain_df2 <- rain_df
rain_df2[is.na(rain_df2)==T] <- 0
# calculate sumulatve sums
MDCs <- list()
for (i in 1:4) {
  MDCs[[i]] <- apply(subset(rain_df2,
                as.character(rain_df2$decade)==
                  as.character(decades[i]))[,2:14],
                      2,cumsum)
}

# plotting
plot.list <- vector("list", length=3)
for (j in seq_along(periods)) {
  mdc = melt(MDCs[[j+1]][1:3652,])[,2:3]
  plot.list[[j]] <- cbind(mdc=mdc, period = rep(periods[j],3652*13))
  colnames(  plot.list[[j]]) <- c("Station", "cumulative_rain","period")
  levels(  plot.list[[j]]$Station) <- Stations[,1]
}
# bring back together to a dataframe
plot.df <- do.call(rbind,plot.list)
# add the the first decade as x-axis
plot.df <- cbind(plot.df,
                base_decade=rep(melt(MDCs[[1]][1:3652,])[,3],3))

# make a plot
p <- ggplot(plot.df, aes(x = base_decade,
                         y = cumulative_rain)) +
  geom_line(aes(linetype=period, colour=period),size=1.2) +
  facet_wrap(~ Station,ncol=5, scales="free") +
  theme(legend.position="bottom") +
  guides(col = guide_legend(nrow = 3))
p <- p + ggtitle("Station rainfall") +
  theme(plot.title = element_text(lineheight=.8, face="bold")) +
  xlab("cumulative sum 1970 - 1980") +
  theme(axis.title.x = element_text(face="bold",  size=14),
        axis.text.x  = element_text(size=10)) +
  ylab("Cumulative sum") +
```

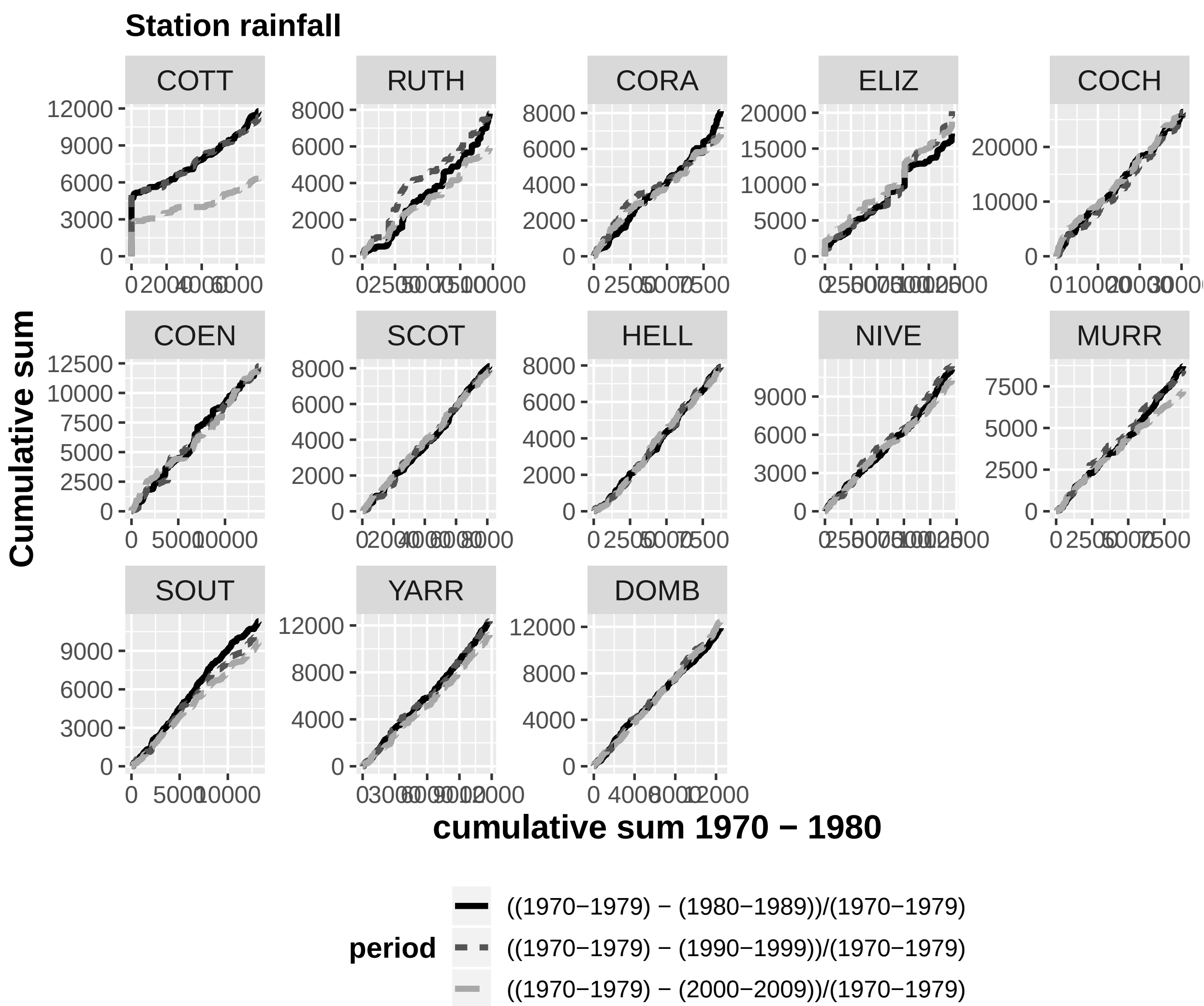

Figure 8: Double mass curves based on daily station rainfall data split by decade

```
  theme(axis.title.y = element_text(face="bold",  size=14),
        axis.text.y  = element_text(size=10)) +
 scale_colour_manual(values=c("black", "gray33", "gray66")) +
  theme(legend.text = element_text( size = 10))+
  theme(legend.title = element_text(size=12, face="bold")) +
  theme(strip.text.x = element_text(size=12))
print(p)
```

```
# # publication quality
# tiff("../manuscript/Figure6_StreamflowFDCDifference.tif",
#      width=16*480, height=12*480,
#      res=600, compression="lzw")
# print(p)
# dev.off()
```

Again, this indicates variation between the decades, but no consistent trend between the decades or for groups of stations. In general this indicates that the could be changes in the distribution of rainfall, but that this change is currently not strong enough to be consistently reflected in changes in the frequency curve or the double mass curve. There seems to be a change at the end of the decades

Double mass gridded rainfall

```
# Calculate the double mass curve for each catchment by decade
# apply cumulative sum across the columns and produce a list
# do this for each decade
# set all missing values to 0
grrain_df2 <- grrain_df
#rain_df2[is.na(rain_df2)==T] <- 0
# calculate sumulatve sums
MDCs <- list()
for (i in 1:4) {
  MDCs[[i]] <- apply(subset(grrain_df2,
                as.character(grrain_df2$decade)==
                  as.character(decades[i]))[,2:14],
                      2,cumsum)
}

# plotting
plot.list <- vector("list", length=3)
for (j in seq_along(periods)) {
  mdc = melt(MDCs[[j+1]][1:3652,])[,2:3]
  plot.list[[j]] <- cbind(mdc=mdc, period = rep(periods[j],3652*13))
  colnames(  plot.list[[j]]) <- c("Station", "cumulative_rain","period")
  levels(  plot.list[[j]]$Station) <- Stations[,1]
}
# bring back together to a dataframe
plot.df <- do.call(rbind,plot.list)
# add the the first decade as x-axis
plot.df <- cbind(plot.df,
                base_decade=rep(melt(MDCs[[1]][1:3652,])[,3],3))

# make a plot
p <- ggplot(plot.df, aes(x = base_decade,
                         y = cumulative_rain)) +
  geom_line(aes(linetype=period, colour=period),size=1.2) +
  facet_wrap(~ Station,ncol=5, scales="free") +
  theme(legend.position="bottom") +
  guides(col = guide_legend(nrow = 3))
p <- p + ggtitle("Gridded rainfall") +
  theme(plot.title = element_text(lineheight=.8, face="bold")) +
  xlab("cumulative sum 1970 - 1980") +
  theme(axis.title.x = element_text(face="bold",  size=14),
        axis.text.x  = element_text(size=10)) +
  ylab("Cumulative sum") +
  theme(axis.title.y = element_text(face="bold",  size=14),
        axis.text.y  = element_text(size=10)) +
 scale_colour_manual(values=c("black", "gray33", "gray66")) +
  theme(legend.text = element_text( size = 10))+
```

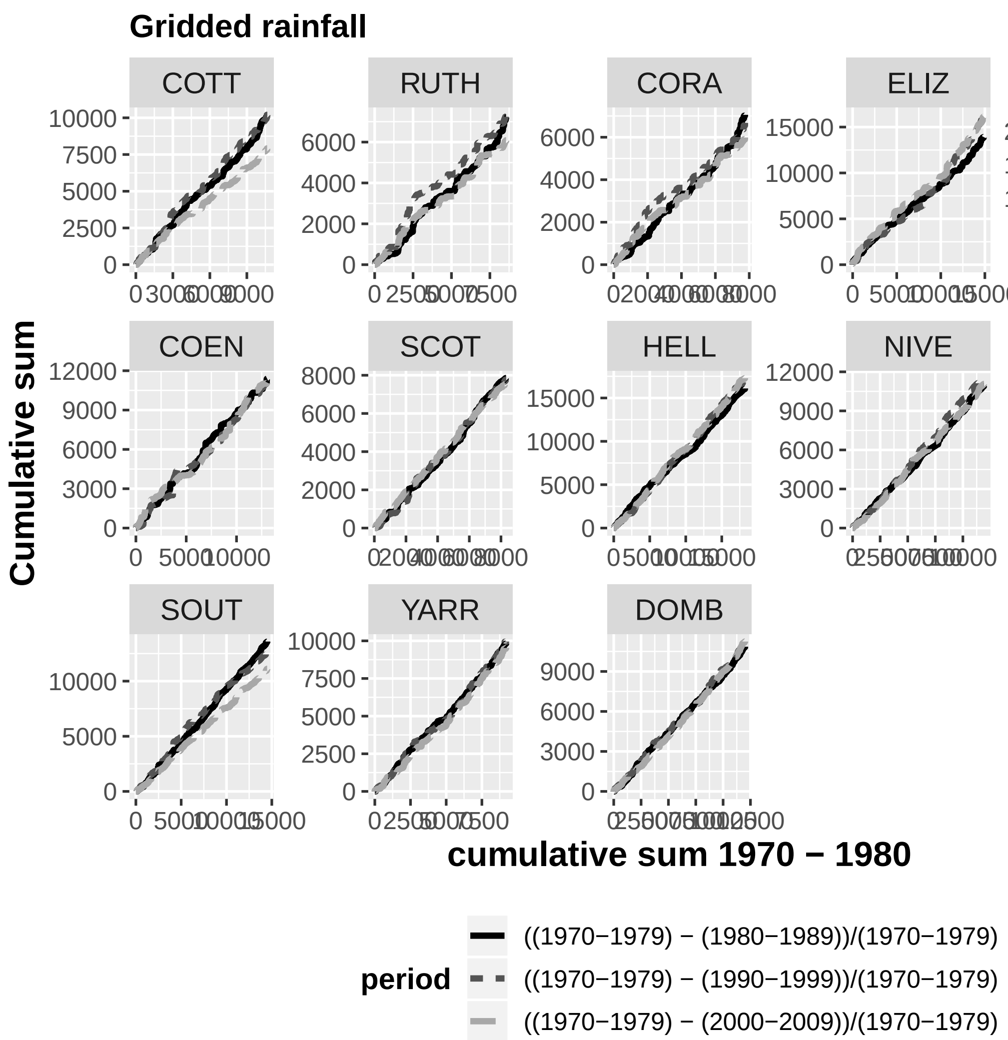

Figure 9: Double mass curves based on daily gridded rainfall data split by decade

```r
  theme(legend.title = element_text(size=12, face="bold")) +
  theme(strip.text.x = element_text(size=12))
print(p)
```

```r
# # publication quality
# tiff("../manuscript/Figure6_StreamflowFDCDifference.tif",
#      width=16*480, height=12*480,
#      res=600, compression="lzw")
# print(p)
# dev.off()
```

The only thing that shows up in most of the southern catchments is a consistenly lower decade between 2000 - 2009, but this is often preceded by a wetter decade between 1990 - 1999, and a drier decade between 1980 - 1989. This suggests a more periodic change rather than a consistent trend.