Mann Kendall tests

Willem Vervoort, Michaela Dolk & Floris van Ogtrop

2017-03-15

```
# root dir
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(tidyr)
library(xts)
library(zoo)
library(Kendall)
library(mgcv)
library(oz)
library(ggplot2)
library(deseasonalize)
```

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 to Journal of Hydrology. This is the second document as part of the response to reviewers of earlier versions of the manuscript outlining in detail the analysis related to the mann-kendall statistical tests on the streamflow, rainfall and temperature data. The results of these tests are incorporated in the submitted manuscript as **Table 3**.

The key point of the the analysis on the weekly data is to provide a non-parametric trend test, which can deal with the correlated timeseries stream flow, rainfall and temperature data. To check, whether the weekly summaries bias the analysis, we will also run the Mann Kendall test on the daily data.

The data

Using the datasets that were developed earlier, we can load in the daily data for streamflow, rainfall and temperature. The difference is that the Mann Kendal analysis should concentrate on analysing the anomalies rather than the actual data, and therefore we have to resummarise to weekly data after calculating the anomalies.

The other thing that is missing is a dataset for the anomalies in weekly "maximum maximum temperature", rather than the weekly average maximum temperature.

A further point to note is that the observed data contain missing values, which remain in the analysis. The gridded data does not contain missing data as this is interpolated predicted data.

Deseasonalise the data

The first step is to calculate the anomalies by deseasonalising the data using the package `deseasonalize` in R: deseasonalize.

```
load("data/DailyDataIncludingGridded.Rdata")
load("data/ClimCh_project_MD.Rdata")


# daily observed flow
flow_deseas <- flow_zoo
# now assign to a new dataframe
for (i in (seq_along(Stations[,1]))) {
  foo <- flow_zoo[,i]
# replace missing values with mean flow
  bad <- is.na(foo)
  foo[bad] <- mean(foo,na.rm=T)
  flow_deseas[,i]  <- ds(as.ts(foo),ic="AIC")$z
# put NA values back
  flow_deseas[bad,i] <- NA
}
# daily obseverved rainfall
rain_deseas <- rain_zoo
# now assign to a new dataframe
for (i in seq_along(Stations[,1])) {
  foo <- rain_zoo[,i]
  # replace NA values with mean flow
  bad <- is.na(foo)
  foo[bad] <- mean(foo,na.rm=T)
  rain_deseas[,i]  <- ds(as.numeric(foo),ic="AIC")$z
  # put NA values back
  rain_deseas[bad,i] <- NA
}

# daily observed maximum temperature
maxT_deseas <- maxT_zoo
# now assign to a new dataframe
for (i in seq_along(Stations[,1])) {
  foo <- maxT_zoo[,i]
  # replace NA values with mean flow
  bad <- is.na(foo)
  foo[bad] <- mean(foo,na.rm=T)
  maxT_deseas[,i]  <- ds(as.numeric(foo),ic="AIC")$z
  # put NA values back
  maxT_deseas[bad,i] <- NA
}

# do the same for the gridded rainfall data
rain_griddeseas <- rain_zoo
#
for (i in seq_along(Stations[,1])) {
  foo <- GridRainAllDataout[GridRainAllDataout[,"Station"]==Stations[i,1],2]
  foo.z <- zoo(foo,  order.by=time(rain_zoo))
  rain_griddeseas[,i]  <- ds(as.numeric(foo),ic="AIC")$z
```

```
}
```

Summarise to weekly data

Similar to the original data the package `xts` can be used to summarise the data to weekly mean values. However, for the maximum temperature, we also calculate the "maximum" weekly maximum temperature, as this might be more meaningful than just the mean value.
In addition, the monthly anomalies are also generated so the difference between the trend analysis on a weekly and monthly scale can be checked.

```
# flow
flow_xts <- xts(flow_deseas,
                order.by=time(flow_deseas),
                frequency=1)
# weekly data
flow_weekly_xts <- apply.weekly(flow_xts, mean)
# monthly data
flow_monthly_xts <- apply.monthly(flow_xts,mean)

# rainfall
rainfall_xts <- xts(rain_deseas,
                    order.by=time(rain_deseas),
                    frequency=1)
# weekly data
rainfall_weekly_xts <- apply.weekly(rainfall_xts, mean)
# monthly data
rainfall_monthly_xts <- apply.monthly(rainfall_xts,mean)

# gridded rainfall
rainfall_grdxts <- xts(rain_griddeseas,
                    order.by=time(rain_griddeseas),
                    frequency=1)
# weekly data
rainfall_grdweekly_xts <- apply.weekly(rainfall_grdxts, mean)
# monthly data
rainfall_grdmonthly_xts <- apply.monthly(rainfall_grdxts,mean)


# maxT
maxT_xts <- xts(maxT_deseas,
                order.by=time(maxT_deseas),
                frequency=1)
# weekly data this calculates the "mean" maximum temperature
maxT_weekly_xts <- apply.weekly(maxT_xts, mean)
# monthly data
maxT_monthly_xts <- apply.monthly(maxT_xts,mean)

# Also calculate the "max" maximum temperature
# first substitute a large value into all NA values
maxT_xts2 <- maxT_xts
maxT_xts2 <- apply(maxT_xts2,2,function(x) ifelse(is.na(x)==T,99999,x))

# calculate the maximum by week and month
```

```
m_maxT_weekly_xts <- apply.weekly(maxT_xts2,function(x) apply(x,2,max,na.rm=T))
m_maxT_monthly_xts <- apply.monthly(maxT_xts2,function(x) apply(x,2,max,na.rm=T))

# now relace the 99999 values with NA again
m_maxT_weekly_xts <- as.xts(apply(m_maxT_weekly_xts,2,
                                  function(x) ifelse(x==99999,NA,x)))
m_maxT_monthly_xts <- as.xts(apply(m_maxT_monthly_xts,2,
                                   function(x) ifelse(x==99999,NA,x)))
```

Bootstrap Mann Kendall analysis

The idea here is to follow Westra et al. (2012) and calculate the field significance using a bootstrap analysis. This essentially check whether the significance is not due to the random sequence of years, i.e. is not a Type 1 error.

This starts will setting up the bootstrap for the weekly data for Mann Kendall analysis followed by the monthly analysis. In the bootstrap the sequence of years is randomly shuffled for each bootstrap iteration. First the data set needs to be split in years

```
split_flow <- split(flow_weekly_xts,"years")
split_rain <- split(rainfall_weekly_xts,"years")
split_grdrain <- split(rainfall_grdweekly_xts,"years")
split_maxT <- split(maxT_weekly_xts,"years")
split_mmaxT <- split(m_maxT_weekly_xts,"years")
```

To run the bootstrap repeatedly and make sure that the results do not vary slightly every time the analysis is run, the seed is set.

```
set.seed(10)
```

The loop is run over all the years (this creates 41 different data sets) and the Mann Kendall test is done on each resonstituted series and the results are saved.

Streamflow

Basically the same analysis is repeated for each of the data sets. Only the flow data is shown in the manuscript this is **Figure 5**.

```
MK_list <- list()
for (i in 1:500) {
  # reorganise the list elements
  series <- sample(1:nyears(flow_weekly_xts),nyears(flow_weekly_xts))
  for (j in 1:length(series)) {
    if (j==1) {
      new_df <- as.data.frame(split_flow[[series[j]]])
    } else {
  # rbind to dataframe
      new_df <- rbind(new_df,as.data.frame(split_flow[[series[j]]]))
    }
  }
  # run mann kendall on the columns and store the results
  mk_r <- apply(new_df,2,MannKendall)
  MK_list[[i]] <- do.call(cbind,mk_r)
```

```
}

MK_df <- do.call(rbind,MK_list)

# identify the p-values and the Mann Kendall tau
pvalues <- subset(MK_df, rownames(MK_df)=="sl")
tau <- subset(MK_df, rownames(MK_df)=="tau")
# find only the significant values
sig_set <- list()

for (i in 1:ncol(pvalues)) {
  #i <- 1
  set <- data.frame(pvalue=as.numeric(pvalues[,i]),
                    tau=as.numeric(tau[,i]),catch=rep(colnames(MK_df)[i],nrow(tau)))
  sig_set[[i]] <- set[set$pvalue < 0.05,]
}

# prepare a dataframe for plotting
sig_set_a <- do.call(rbind,sig_set)
sig_set_a$type <- rep("bootstrap",nrow(sig_set_a))

# do the Mann Kendall on the real data
real <- do.call(rbind,apply(flow_weekly_xts,2,MannKendall))
real_df_f <- data.frame(pvalue = as.numeric(real[,2]),
                      tau = as.numeric(real[,1]),
                      catch=rownames(real),type=rep("real",nrow(real)))
sig_set_f <- rbind(sig_set_a,real_df_f)

# Plot a histogram of taus
hp <- ggplot(sig_set_a, aes(x=tau)) + geom_histogram(binwidth=0.03,colour="white")

# Histogram of significant tau's, facetted by catchment
# With panels that have the same scaling,
# but different range (and therefore different physical sizes)
hp <- hp + facet_wrap(~ catch,ncol=5) + ggtitle("Streamflow")
# add a red point for the real slope from the data
p_value <- ifelse(real_df_f$pvalue<0.05,"< 0.05",">= 0.05")
hp <- hp + geom_point(data=real_df_f,aes(x=tau, y=0,colour=p_value),
                shape=16,size=5) + scale_colour_grey(start = 0, end = 0.6) +
  facet_wrap(~ catch,ncol=5)+ ggtitle("Deseasonalised Streamflow") #+
# show figure
print(hp)
```

```
# temporary saving for easy plotting later
save(hp,file="C:/Users/rver4657/ownCloud/Virtual Experiments/Manuscript drafts/MKdeseasonalisedStreamfl
#load("20160729_MKdeseasonalisedStreamflow.Rdata")

# publication quality figure for manuscript
tiff(paste("C:/Users/rver4657/ownCloud/Virtual Experiments/Manuscript",
           "Figure5_MKdeseasonalisedStreamflow.tif",sep="/"),
     width=12*480,height=10*480,
     compression = "lzw", res=600)
print(hp)
dev.off()
```

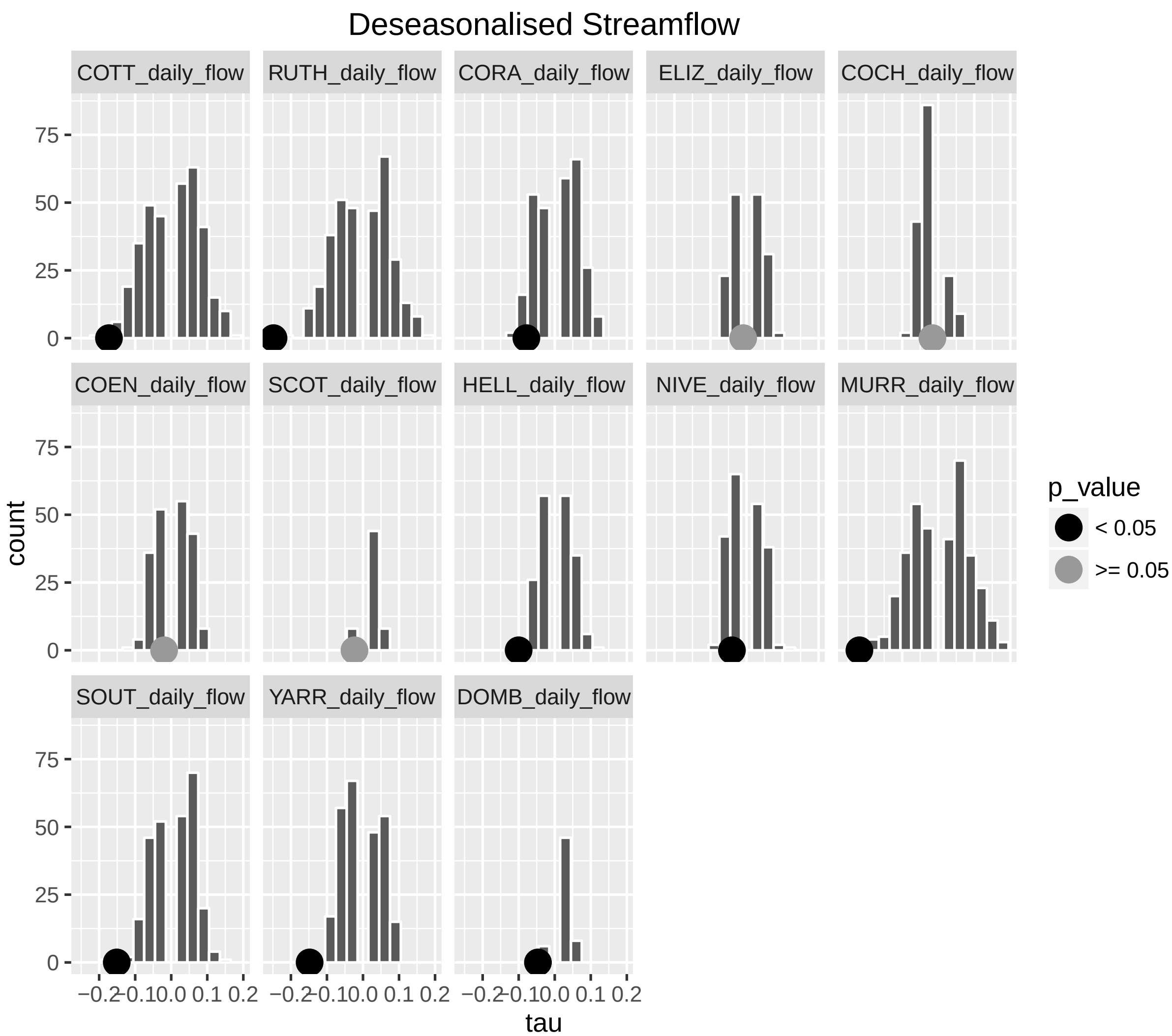

Figure 1: Results of the Bootstrap Mann Kendall analysis on the weekly streamflow data. The dot shows the actual result, while the histogram bars show the results of the bootstrap analysis

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

```r
# -------- end flow -----------------
```

rainfall

```r
MK_list <- list()
for (i in 1:500) {
  # reorganise the list elements
  series <- sample(1:nyears(rainfall_weekly_xts),nyears(rainfall_weekly_xts))
  for (j in 1:length(series)) {
    # j <- 1
    if (j==1) {
      new_df <- as.data.frame(split_rain[[series[j]]])
    } else {
      new_df <- rbind(new_df,as.data.frame(split_rain[[series[j]]]))
    }
  }
  # rbind to dataframe
  # run mann kendall on the columns and store the results
  mk_r <- apply(new_df,2,MannKendall)

  MK_list[[i]] <- do.call(cbind,mk_r)
}

MK_df <- do.call(rbind,MK_list)

# prepare a dataframe for plotting
pvalues <- subset(MK_df, rownames(MK_df)=="sl")
tau <- subset(MK_df, rownames(MK_df)=="tau")

sig_set <- list()

for (i in 1:ncol(pvalues)) {
  set <- data.frame(pvalue=as.numeric(pvalues[,i]),
                    tau=as.numeric(tau[,i]),catch=rep(colnames(MK_df)[i],nrow(tau)))
  sig_set[[i]] <- set[set$pvalue< 0.05,]
}

sig_set_a <- do.call(rbind,sig_set)
sig_set_a$type <- rep("bootstrap",nrow(sig_set_a))
real <- do.call(rbind,apply(rainfall_weekly_xts,2,MannKendall))
real_df_r <- data.frame(pvalue = as.numeric(real[,2]),
                      tau = as.numeric(real[,1]),
                      catch=rownames(real),type=rep("real",nrow(real)))
sig_set_f <- rbind(sig_set_a,real_df_r)
# A histogram of taus
hp <- ggplot(sig_set_a, aes(x=tau)) + geom_histogram(binwidth=0.03,colour="white")

# Histogram of significant tau's, facet wrapped by catchment
# With panels that have the same scaling, but different range
# (and therefore different physical sizes)
```

```
hp <- hp + facet_wrap(~ catch,ncol=5)
# add a red point for the real slope from the data
p_value <- ifelse(real_df_r$pvalue<0.05,"< 0.05",">= 0.05")
hp <- hp + geom_point(data=real_df_r,aes(x=tau, y=0,colour=p_value),
                shape=16,size=5) +
  scale_colour_brewer(palette="Set1") +
  facet_wrap(~ catch,ncol=5)+ ggtitle("Deseasonalised Rainfall") #+
print(hp)
```

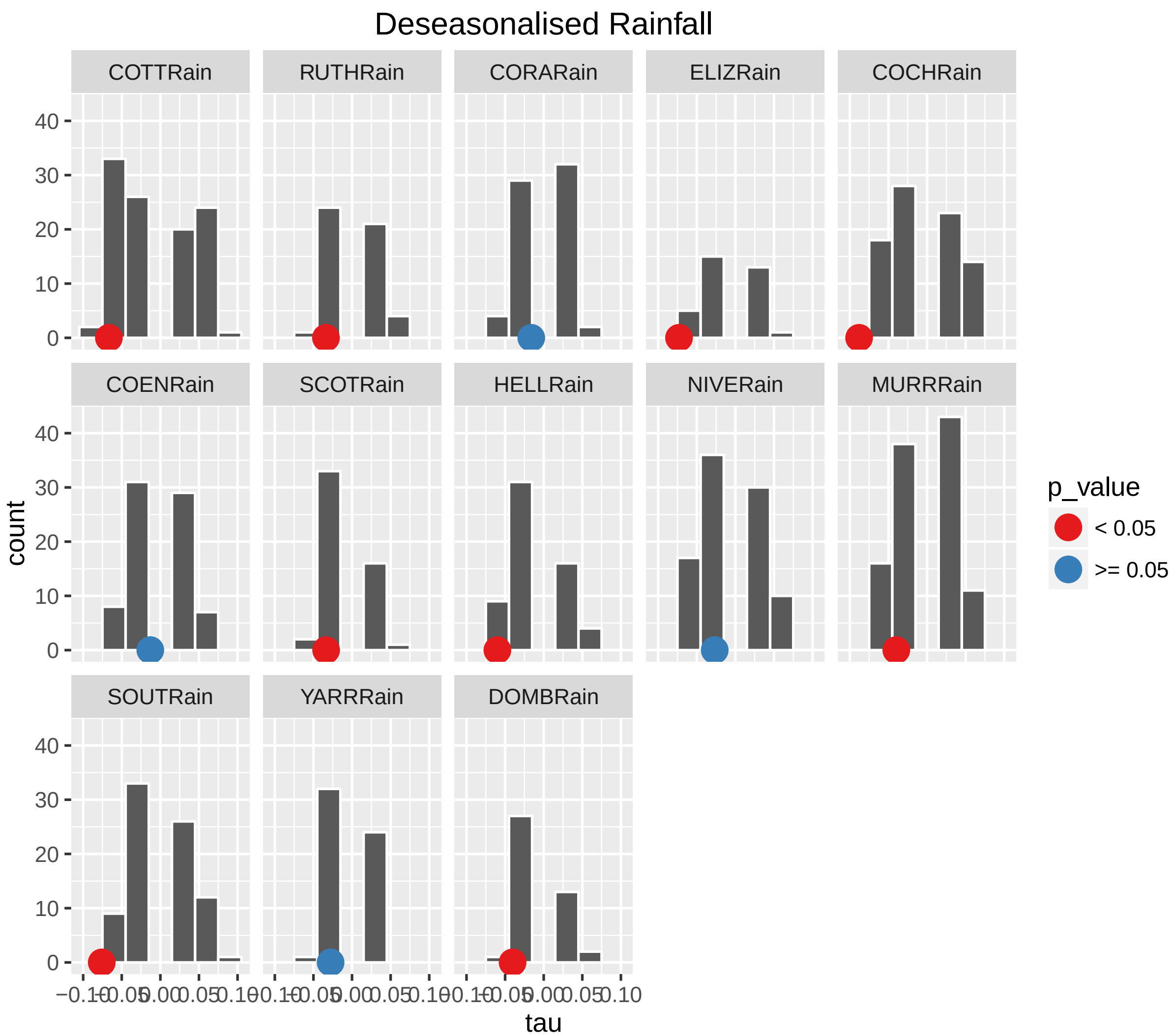

Figure 2: Results of the Bootstrap Mann Kendall analysis on the weekly rainfall data. The dot shows the actual result, while the histogram bars show the results of the bootstrap analysis

Gridded rainfall

```
MK_list <- list()
for (i in 1:500) {
  # reorganise the list elements
```

```r
  series <- sample(1:nyears(rainfall_grdweekly_xts),nyears(rainfall_grdweekly_xts))
  for (j in 1:length(series)) {
    if (j==1) {
      new_df <- as.data.frame(split_grdrain[[series[j]]])
    } else {
      new_df <- rbind(new_df,as.data.frame(split_grdrain[[series[j]]]))
    }
  }
  # rbind to dataframe
  # run mann kendall on the columns and store the results
  mk_r <- apply(new_df,2,MannKendall)

  MK_list[[i]] <- do.call(cbind,mk_r)
}

MK_df <- do.call(rbind,MK_list)

# prepare a dataframe for plotting
pvalues <- subset(MK_df, rownames(MK_df)=="sl")
tau <- subset(MK_df, rownames(MK_df)=="tau")

sig_set <- list()

for (i in 1:ncol(pvalues)) {
  set <- data.frame(pvalue=as.numeric(pvalues[,i]),
                    tau=as.numeric(tau[,i]),catch=rep(colnames(MK_df)[i],nrow(tau)))
  sig_set[[i]] <- set[set$pvalue< 0.05,]
}

sig_set_a <- do.call(rbind,sig_set)
sig_set_a$type <- rep("bootstrap",nrow(sig_set_a))
real <- do.call(rbind,apply(rainfall_grdweekly_xts,2,MannKendall))
real_df_grd_r <- data.frame(pvalue = as.numeric(real[,2]),
                      tau = as.numeric(real[,1]),
                      catch=rownames(real),type=rep("real",nrow(real)))
sig_set_f <- rbind(sig_set_a,real_df_grd_r)
# A histogram of taus
hp <- ggplot(sig_set_a, aes(x=tau)) + geom_histogram(binwidth=0.03,colour="white")

# Histogram of significant tau's, facet wrapped by catchment
# With panels that have the same scaling, but different range
# (and therefore different physical sizes)
hp <- hp + facet_wrap(~ catch,ncol=5)
# add a red point for the real slope from the data
p_value <- ifelse(real_df_grd_r$pvalue<0.05,"< 0.05",">= 0.05")
hp <- hp + geom_point(data=real_df_grd_r,aes(x=tau, y=0,colour=p_value),
               shape=16,size=5) +
  scale_colour_brewer(palette="Set1") +
  facet_wrap(~ catch,ncol=5)+ ggtitle("Deseasonalised Gridded Rainfall") #+
print(hp)
```

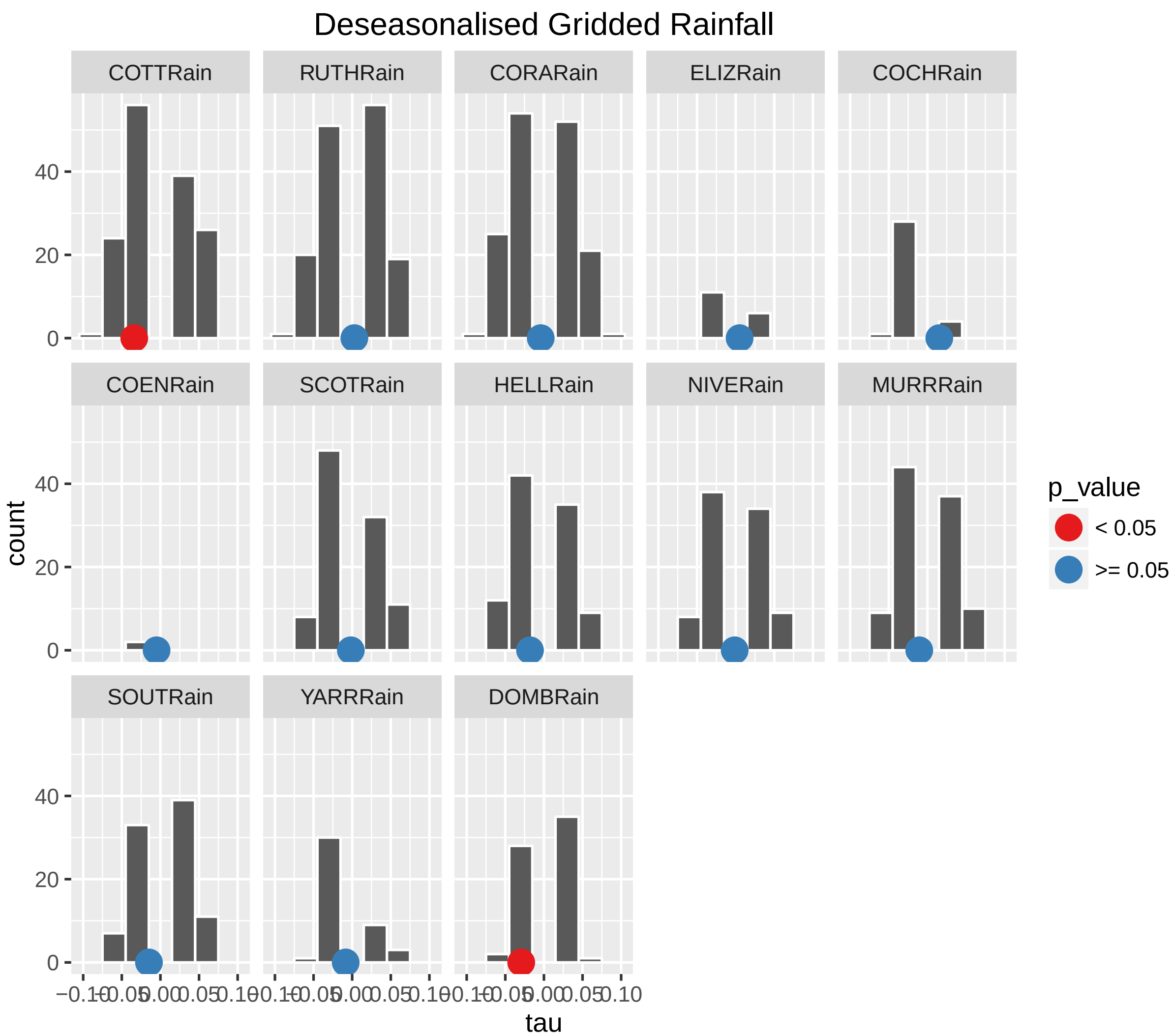

Figure 3: Results of the Bootstrap Mann Kendall analysis on the weekly gridded rainfall data. The dot shows the actual result, while the histogram bars show the results of the bootstrap analysis

Mean Maximum temperature

```
MK_list <- list()
for (i in 1:500) {
  # reorganise the list elements
  series <- sample(1:nyears(maxT_weekly_xts),nyears(maxT_weekly_xts))
  for (j in 1:length(series)) {
    if (j==1) {
      new_df <- as.data.frame(split_maxT[[series[j]]])
    } else {
      new_df <- rbind(new_df,as.data.frame(split_maxT[[series[j]]]))
    }
  }
  # rbind to dataframe
  # run mann kendall on the columns and store the results
  mk_r <- apply(new_df,2,MannKendall)

  MK_list[[i]] <- do.call(cbind,mk_r)
}

MK_df <- do.call(rbind,MK_list)

# prepare a dataframe for plotting
pvalues <- subset(MK_df, rownames(MK_df)=="sl")
tau <- subset(MK_df, rownames(MK_df)=="tau")

sig_set <- list()

for (i in 1:ncol(pvalues)) {
  set <- data.frame(pvalue=as.numeric(pvalues[,i]),
                    tau=as.numeric(tau[,i]),catch=rep(colnames(MK_df)[i],nrow(tau)))
  sig_set[[i]] <- set[set$pvalue<0.05,]
}

sig_set_a <- do.call(rbind,sig_set)
sig_set_a$type <- rep("bootstrap",nrow(sig_set_a))
real <- do.call(rbind,apply(maxT_weekly_xts,2,MannKendall))
real_df_maxT <- data.frame(pvalue = as.numeric(real[,2]),
                      tau = as.numeric(real[,1]),
                      catch=rownames(real),type=rep("real",nrow(real)))
sig_set_f <- rbind(sig_set_a,real_df_maxT)
# A histogram of taus
hp <- ggplot(sig_set_a, aes(x=tau)) +
  geom_histogram(binwidth=0.03,colour="white")

# Histogram of significant tau's, facet wrapped by catchment
# With panels that have the same scaling, but different range
# (and therefore different physical sizes)
hp <- hp + facet_wrap(~ catch,ncol=5)
# add a red point for the real slope from the data
p_value <- ifelse(real_df_maxT$pvalue<0.1,"< 0.05",">= 0.05")
hp <- hp + geom_point(data=real_df_maxT,aes(x=tau, y=0,colour=p_value),
                shape=16,size=5) +
```

```
  scale_colour_brewer(palette="Set1") +
  facet_wrap(~ catch,ncol=5)+ ggtitle("Deseasonalised Maximum Temperature") #+
  #scale_colour_discrete(name="p-value",
  #                      breaks=c("darkblue","red"),
  #                      labels=c("<0.1",">=0.1"))
print(hp)
```

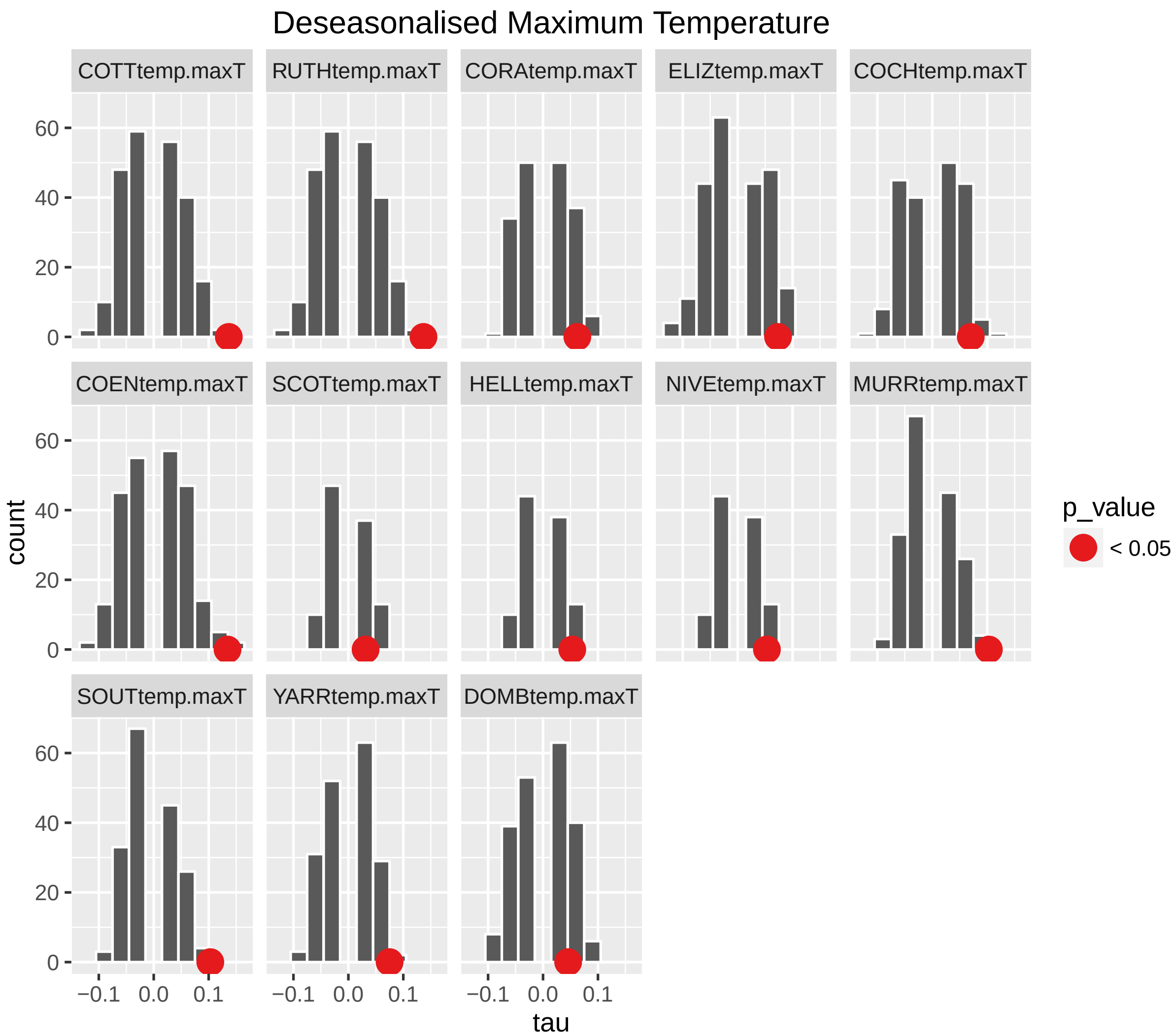

Figure 4: Results of the Bootstrap Mann Kendall analysis on the weekly average maximum temperature data. The dot shows the actual result, while the histogram bars show the results of the bootstrap analysis

Maximum Maximum temperature

This is the analysis for the weekly maximum in the maximum temperature as comparison to see if the assumption of the average maximum temperature is not biased.

```
MK_list <- list()
for (i in 1:500) {
  # reorganise the list elements
```

```
  series <- sample(1:nyears(m_maxT_weekly_xts),nyears(m_maxT_weekly_xts))
  for (j in 1:length(series)) {
    if (j==1) {
      new_df <- as.data.frame(split_mmaxT[[series[j]]])
    } else {
      new_df <- rbind(new_df,as.data.frame(split_mmaxT[[series[j]]]))
    }
  }
  # rbind to dataframe
  # run mann kendall on the columns and store the results
  mk_r <- apply(new_df,2,MannKendall)

  MK_list[[i]] <- do.call(cbind,mk_r)
}

MK_df <- do.call(rbind,MK_list)

# prepare a dataframe for plotting
pvalues <- subset(MK_df, rownames(MK_df)=="sl")
tau <- subset(MK_df, rownames(MK_df)=="tau")

sig_set <- list()

for (i in 1:ncol(pvalues)) {
  set <- data.frame(pvalue=as.numeric(pvalues[,i]),
                    tau=as.numeric(tau[,i]),catch=rep(colnames(MK_df)[i],nrow(tau)))
  sig_set[[i]] <- set[set$pvalue<0.05,]
}

sig_set_a <- do.call(rbind,sig_set)
sig_set_a$type <- rep("bootstrap",nrow(sig_set_a))
real <- do.call(rbind,apply(m_maxT_weekly_xts,2,MannKendall))
real_df_mmaxT <- data.frame(pvalue = as.numeric(real[,2]),
                      tau = as.numeric(real[,1]),
                      catch=rownames(real),type=rep("real",nrow(real)))
sig_set_f <- rbind(sig_set_a,real_df_mmaxT)
# A histogram of taus
hp <- ggplot(sig_set_a, aes(x=tau)) +
  geom_histogram(binwidth=0.03,colour="white")

# Histogram of significant tau's, facet wrapped by catchment
# With panels that have the same scaling, but different range
# (and therefore different physical sizes)
hp <- hp + facet_wrap(~ catch,ncol=5)
# add a red point for the real slope from the data
p_value <- ifelse(real_df_mmaxT$pvalue<0.1,"< 0.05",">= 0.05")
hp <- hp + geom_point(data=real_df_mmaxT,aes(x=tau, y=0,colour=p_value),
                shape=16,size=5) +
  scale_colour_brewer(palette="Set1") +
  facet_wrap(~ catch,ncol=5)+ ggtitle("Maximum Maximum Temperature") #+
print(hp)
```

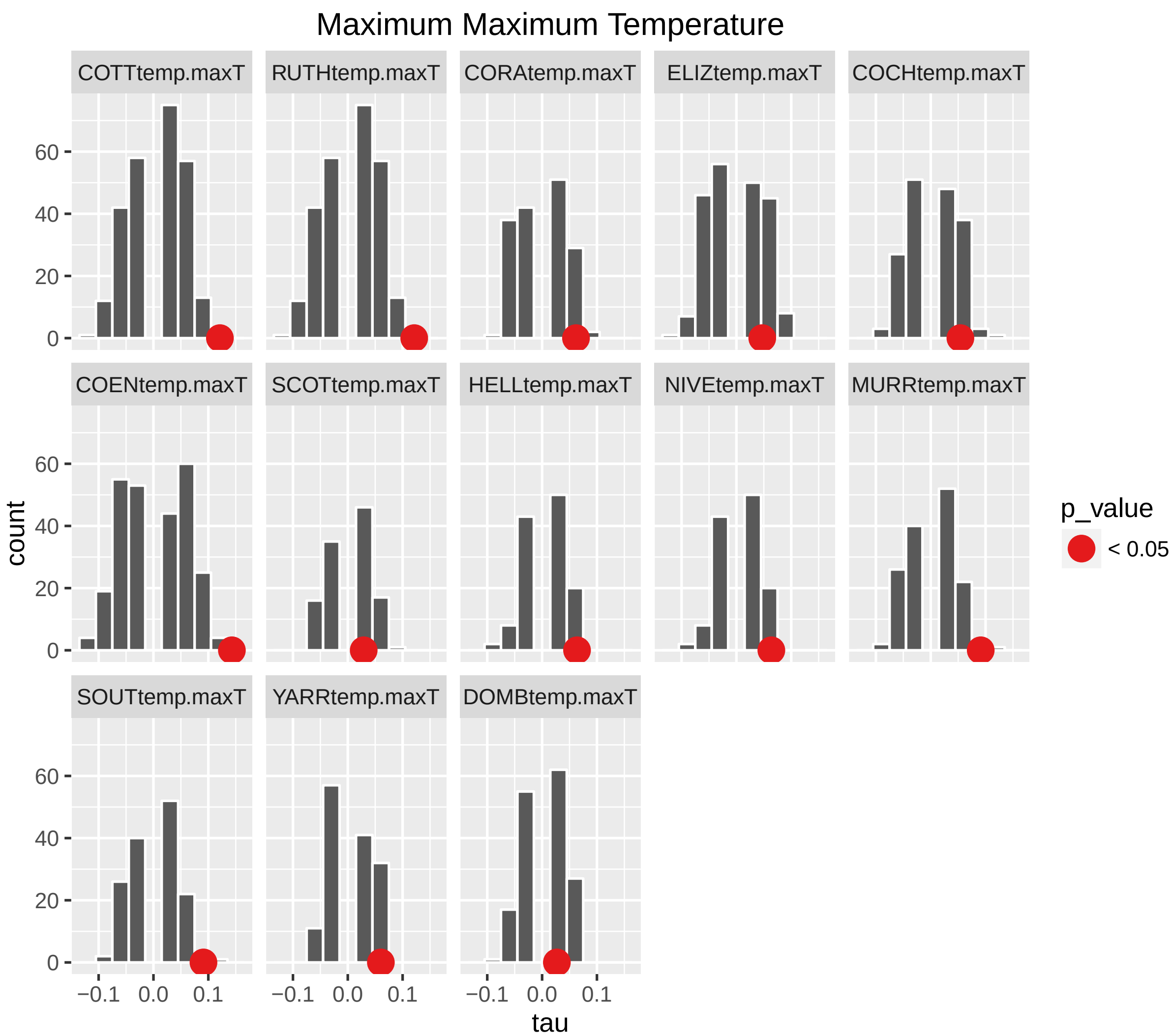

Figure 5: Results of the Bootstrap Mann Kendall analysis on the weekly maximum maximum temperature data. The dot shows the actual result, while the histogram bars show the results of the bootstrap analysis

Summary of weekly analysis in a table

This section brings together the data from all the weekly Mann Kendall analyses and puts this into a table. Some of these results are in **Table 3** in the manuscript.

```
weeklyTable <- data.frame(Catchment = Stations[,1],
                          "Streamflow tau" = real_df_f$tau,
                          "Streamflow p-value" = real_df_f$pvalue,
                          "Rainfall tau" = real_df_r$tau,
                          "Rainfall p-value" = real_df_r$pvalue,
                          "Grid Rainfall tau" = real_df_grd_r$tau,
                          "Grid Rainfall p-value" = real_df_grd_r$pvalue,
                          "Mean max T tau" = real_df_maxT$tau,
                          "Mean max T p-value" = real_df_maxT$pvalue,
                          "Max max T tau" = real_df_mmaxT$tau,
                          "Max max T p-value" = real_df_mmaxT$pvalue)
pander(weeklyTable,
       caption="Mann-Kendall test results on de-seasonalised weekly time series.
       p-values are  considered significant at the 5% level.")
```

Table 1: Mann-Kendall test results on de-seasonalised weekly time series. p-values are considered significant at the 5% level. (continued below)

Catchment	Streamflow.tau	Streamflow.p.value	Rainfall.tau
COTT	-0.1727	0	-0.06663
RUTH	-0.2482	0	-0.03374
CORA	-0.07862	5.013e-08	-0.01612
ELIZ	-0.009122	0.5271	-0.07297
COCH	-0.01612	0.2638	-0.08803
COEN	-0.01979	0.1701	-0.01303
SCOT	-0.02361	0.1017	-0.03356
HELL	-0.1002	3.795e-12	-0.06002
NIVE	-0.04036	0.005132	-0.02679
MURR	-0.2187	0	-0.03981
SOUT	-0.1511	0	-0.0759
YARR	-0.1481	0	-0.02777
DOMB	-0.04659	0.001237	-0.0403

Table 2: Table continues below

Rainfall.p.value	Grid.Rainfall.tau	Grid.Rainfall.p.value	Mean.max.T.tau
0.0002828	-0.03389	0.01879	0.1367
0.02436	0.00284	0.8439	0.1367
0.2672	-0.004133	0.7745	0.06245
3.538e-06	0.005064	0.7255	0.07348
9.714e-08	0.01536	0.2869	0.07001
0.3782	-0.004951	0.7314	0.1344
0.02054	-0.001741	0.904	0.03131
0.0001076	-0.01801	0.2118	0.05325
0.07223	-0.001295	0.9285	0.05325
0.006362	-0.01055	0.4647	0.103
6.43e-07	-0.01483	0.3039	0.103

Rainfall.p.value	Grid.Rainfall.tau	Grid.Rainfall.p.value	Mean.max.T.tau
0.05555	-0.008408	0.5599	0.07502
0.01096	-0.02938	0.04164	0.04572

Mean.max.T.p.value	Max.max.T.tau	Max.max.T.p.value
0	0.1212	0
0	0.1212	0
3.374e-05	0.0618	3.994e-05
3.576e-07	0.04713	0.001129
2.623e-06	0.0541	0.0002779
0	0.143	0
0.03294	0.02898	0.04834
0.0003325	0.06372	1.8e-05
0.0003325	0.06372	1.8e-05
0	0.09118	0
0	0.09118	0
2.384e-07	0.06034	2.921e-05
0.001757	0.02694	0.06534

Monthly Mann Kendall bootstrap analysis

This is a repeat of the weekly analysis to make sure that there is no major bias in the aggregation.

```
split_flow <- split(flow_monthly_xts,"years")
split_rain <- split(rainfall_monthly_xts,"years")
split_grdrain <- split(rainfall_grdmonthly_xts,"years")
split_maxT <- split(maxT_monthly_xts,"years")
split_mmaxT <- split(m_maxT_monthly_xts,"years")
```

The loop is again run over all the years (this creates 41 different data sets) and the Mann Kendall test is done on each resonstituted series and the results are saved.

Monthly Streamflow

Basically the same analysis is repeated for each of the data sets. None of this is shown in the manuscript.

```
MK_list <- list()
for (i in 1:500) {
  # reorganise the list elements
  series <- sample(1:nyears(flow_monthly_xts),nyears(flow_monthly_xts))
  for (j in 1:length(series)) {
    if (j==1) {
      new_df <- as.data.frame(split_flow[[series[j]]])
    } else {
  # rbind to dataframe
      new_df <- rbind(new_df,as.data.frame(split_flow[[series[j]]]))
    }
  }
  # run mann kendall on the columns and store the results
  mk_r <- apply(new_df,2,MannKendall)
```

```r
  MK_list[[i]] <- do.call(cbind,mk_r)
}

MK_df <- do.call(rbind,MK_list)

# identify the p-values and the Mann Kendall tau
pvalues <- subset(MK_df, rownames(MK_df)=="sl")
tau <- subset(MK_df, rownames(MK_df)=="tau")
# find only the significant values
sig_set <- list()

for (i in 1:ncol(pvalues)) {
  #i <- 1
  set <- data.frame(pvalue=as.numeric(pvalues[,i]),
                     tau=as.numeric(tau[,i]),catch=rep(colnames(MK_df)[i],nrow(tau)))
  sig_set[[i]] <- set[set$pvalue < 0.05,]
}

# prepare a dataframe for plotting
sig_set_a <- do.call(rbind,sig_set)
sig_set_a$type <- rep("bootstrap",nrow(sig_set_a))

# do the Mann Kendall on the real data
real <- do.call(rbind,apply(flow_monthly_xts,2,MannKendall))
real_df_f_m <- data.frame(pvalue = as.numeric(real[,2]),
                          tau = as.numeric(real[,1]),
                          catch=rownames(real),type=rep("real",nrow(real)))
sig_set_f <- rbind(sig_set_a,real_df_f_m)

# Plot a histogram of taus
hp <- ggplot(sig_set_a, aes(x=tau)) + geom_histogram(binwidth=0.03,colour="white")

# Histogram of significant tau's, facetted by catchment
# With panels that have the same scaling,
# but different range (and therefore different physical sizes)
hp <- hp + facet_wrap(~ catch,ncol=5)
# add a red point for the real slope from the data
p_value <- ifelse(real_df_f_m$pvalue<0.05,"< 0.05",">= 0.05")
hp <- hp + geom_point(data=real_df_f_m,aes(x=tau, y=0,colour=p_value),
                shape=16,size=5) + scale_colour_grey(start = 0, end = 0.6) +
  facet_wrap(~ catch,ncol=5)+ ggtitle("Deseasonalised Monthly Streamflow") #+
# show figure
print(hp)
```

rainfall

```r
MK_list <- list()
for (i in 1:500) {
  # reorganise the list elements
  series <- sample(1:nyears(rainfall_monthly_xts),nyears(rainfall_monthly_xts))
  for (j in 1:length(series)) {
    # j <- 1
```

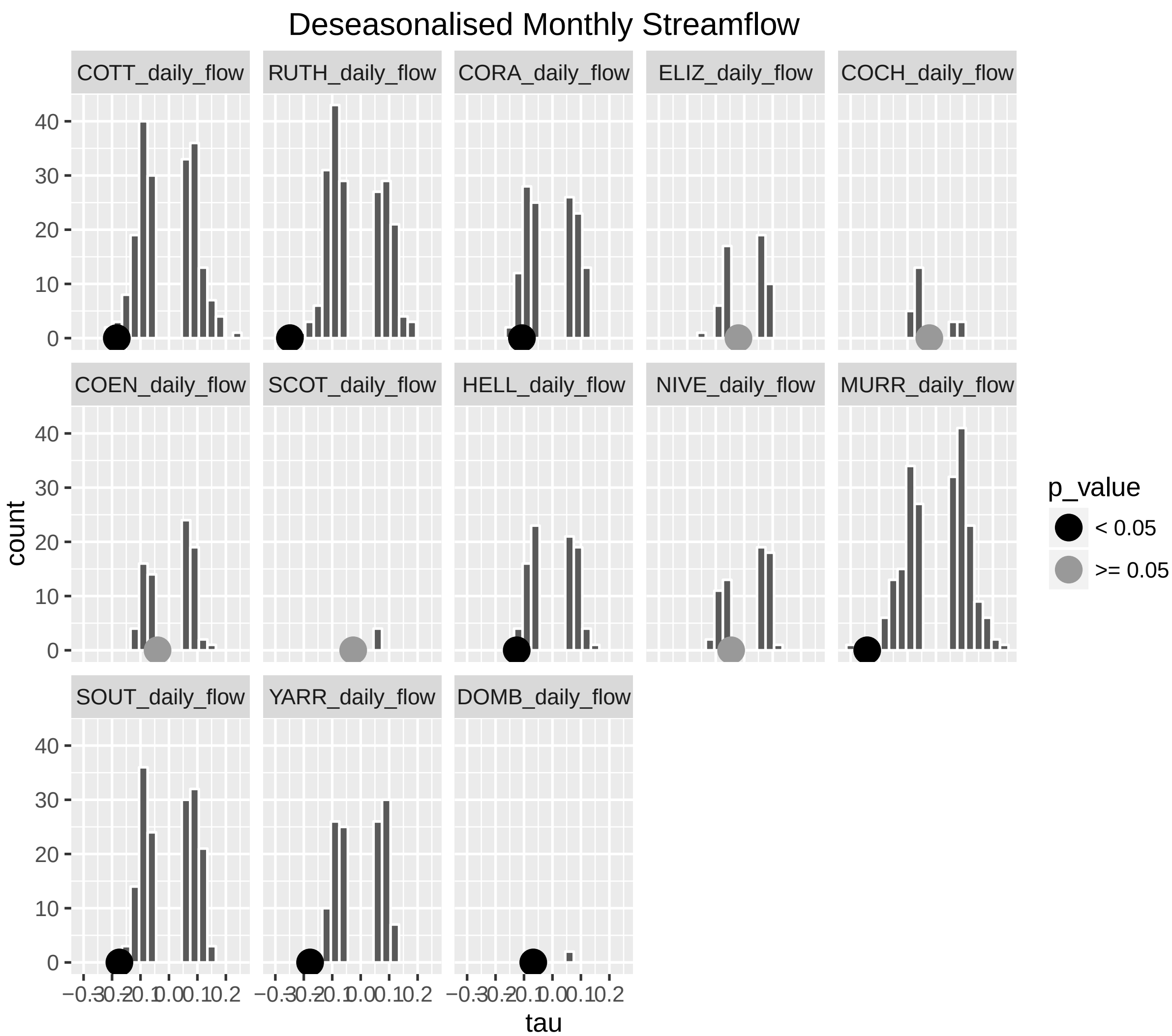

Figure 6: Results of the Bootstrap Mann Kendall analysis on the monthly streamflow data. The dot shows the actual result, while the histogram bars show the results of the bootstrap analysis

```r
    if (j==1) {
      new_df <- as.data.frame(split_rain[[series[j]]])
    } else {
      new_df <- rbind(new_df,as.data.frame(split_rain[[series[j]]]))
    }
  }
  # rbind to dataframe
  # run mann kendall on the columns and store the results
  mk_r <- apply(new_df,2,MannKendall)

  MK_list[[i]] <- do.call(cbind,mk_r)
}

MK_df <- do.call(rbind,MK_list)

# prepare a dataframe for plotting
pvalues <- subset(MK_df, rownames(MK_df)=="sl")
tau <- subset(MK_df, rownames(MK_df)=="tau")

sig_set <- list()

for (i in 1:ncol(pvalues)) {
  set <- data.frame(pvalue=as.numeric(pvalues[,i]),
                    tau=as.numeric(tau[,i]),catch=rep(colnames(MK_df)[i],nrow(tau)))
  sig_set[[i]] <- set[set$pvalue< 0.05,]
}

sig_set_a <- do.call(rbind,sig_set)
sig_set_a$type <- rep("bootstrap",nrow(sig_set_a))
real <- do.call(rbind,apply(rainfall_monthly_xts,2,MannKendall))
real_df_r_m <- data.frame(pvalue = as.numeric(real[,2]),
                      tau = as.numeric(real[,1]),
                      catch=rownames(real),type=rep("real",nrow(real)))
sig_set_f <- rbind(sig_set_a,real_df_r_m)
# A histogram of taus
hp <- ggplot(sig_set_a, aes(x=tau)) + geom_histogram(binwidth=0.03,colour="white")

# Histogram of significant tau's, facet wrapped by catchment
# With panels that have the same scaling, but different range
# (and therefore different physical sizes)
hp <- hp + facet_wrap(~ catch,ncol=5)
# add a red point for the real slope from the data
p_value <- ifelse(real_df_r_m$pvalue<0.05,"< 0.05",">= 0.05")
hp <- hp + geom_point(data=real_df_r_m,aes(x=tau, y=0,colour=p_value),
               shape=16,size=5) +
  scale_colour_brewer(palette="Set1") +
  facet_wrap(~ catch,ncol=5)+ ggtitle("Deseasonalised Monthly Rainfall") #+
print(hp)
```

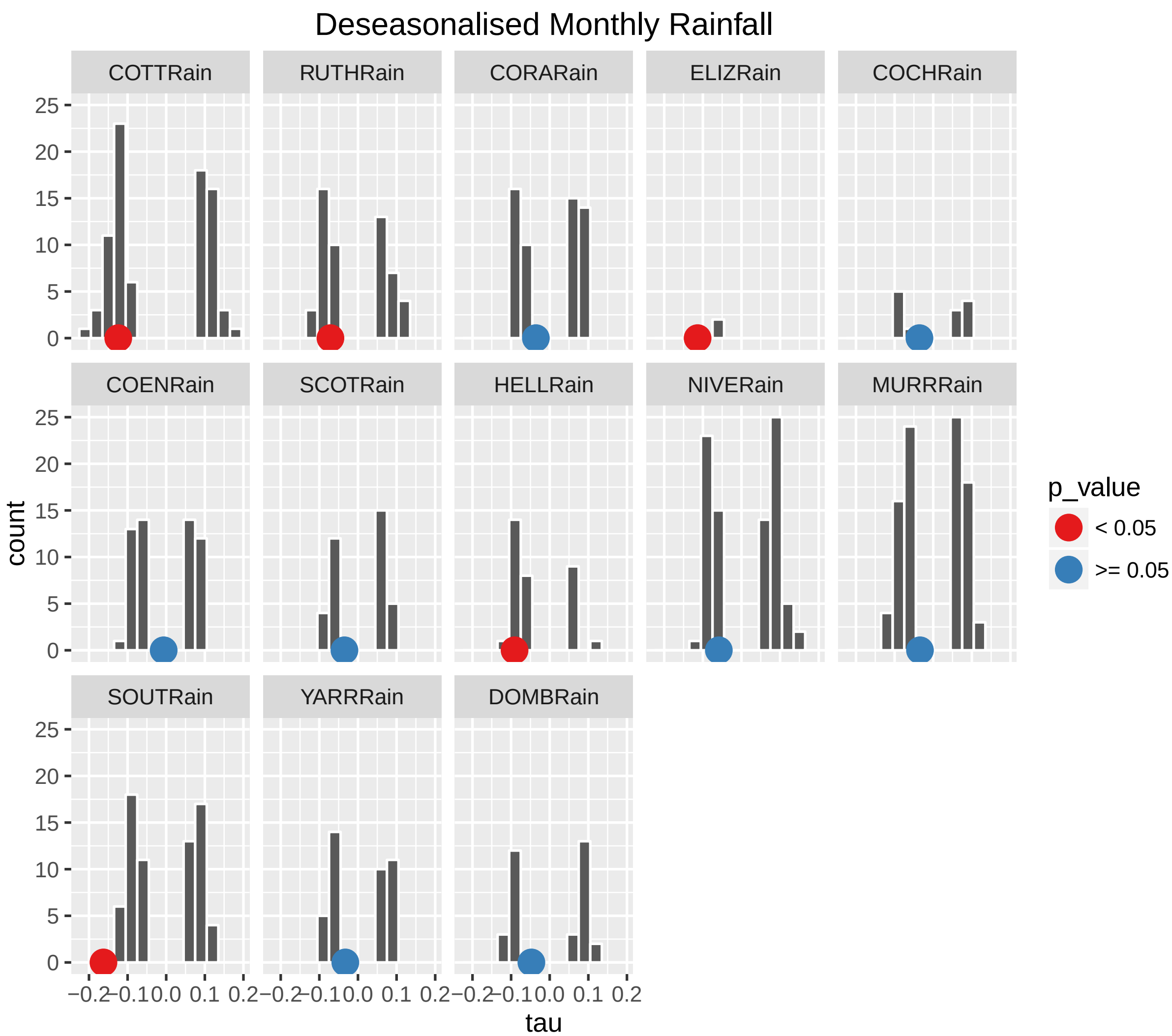

Figure 7: Results of the Bootstrap Mann Kendall analysis on the monthly rainfall data. The dot shows the actual result, while the histogram bars show the results of the bootstrap analysis

Gridded rainfall

```
MK_list <- list()
for (i in 1:500) {
  # reorganise the list elements
  series <- sample(1:nyears(rainfall_grdmonthly_xts),nyears(rainfall_grdmonthly_xts))
  for (j in 1:length(series)) {
    if (j==1) {
      new_df <- as.data.frame(split_grdrain[[series[j]]])
    } else {
      new_df <- rbind(new_df,as.data.frame(split_grdrain[[series[j]]]))
    }
  }
  # rbind to dataframe
  # run mann kendall on the columns and store the results
  mk_r <- apply(new_df,2,MannKendall)

  MK_list[[i]] <- do.call(cbind,mk_r)
}

MK_df <- do.call(rbind,MK_list)

# prepare a dataframe for plotting
pvalues <- subset(MK_df, rownames(MK_df)=="sl")
tau <- subset(MK_df, rownames(MK_df)=="tau")

sig_set <- list()

for (i in 1:ncol(pvalues)) {
  set <- data.frame(pvalue=as.numeric(pvalues[,i]),
                    tau=as.numeric(tau[,i]),catch=rep(colnames(MK_df)[i],nrow(tau)))
  sig_set[[i]] <- set[set$pvalue< 0.05,]
}

sig_set_a <- do.call(rbind,sig_set)
sig_set_a$type <- rep("bootstrap",nrow(sig_set_a))
real <- do.call(rbind,apply(rainfall_grdmonthly_xts,2,MannKendall))
real_df_grdr_m <- data.frame(pvalue = as.numeric(real[,2]),
                      tau = as.numeric(real[,1]),
                      catch=rownames(real),type=rep("real",nrow(real)))
sig_set_f <- rbind(sig_set_a,real_df_grdr_m)
# A histogram of taus
hp <- ggplot(sig_set_a, aes(x=tau)) + geom_histogram(binwidth=0.03,colour="white")

# Histogram of significant tau's, facet wrapped by catchment
# With panels that have the same scaling, but different range
# (and therefore different physical sizes)
hp <- hp + facet_wrap(~ catch,ncol=5)
# add a red point for the real slope from the data
p_value <- ifelse(real_df_grdr_m$pvalue<0.05,"< 0.05",">= 0.05")
hp <- hp + geom_point(data=real_df_grdr_m,aes(x=tau, y=0,colour=p_value),
                shape=16,size=5) +
  scale_colour_brewer(palette="Set1") +
```

```
  facet_wrap(~ catch,ncol=5)+ ggtitle("Deseasonalised Gridded Monthly Rainfall") #+
print(hp)
```

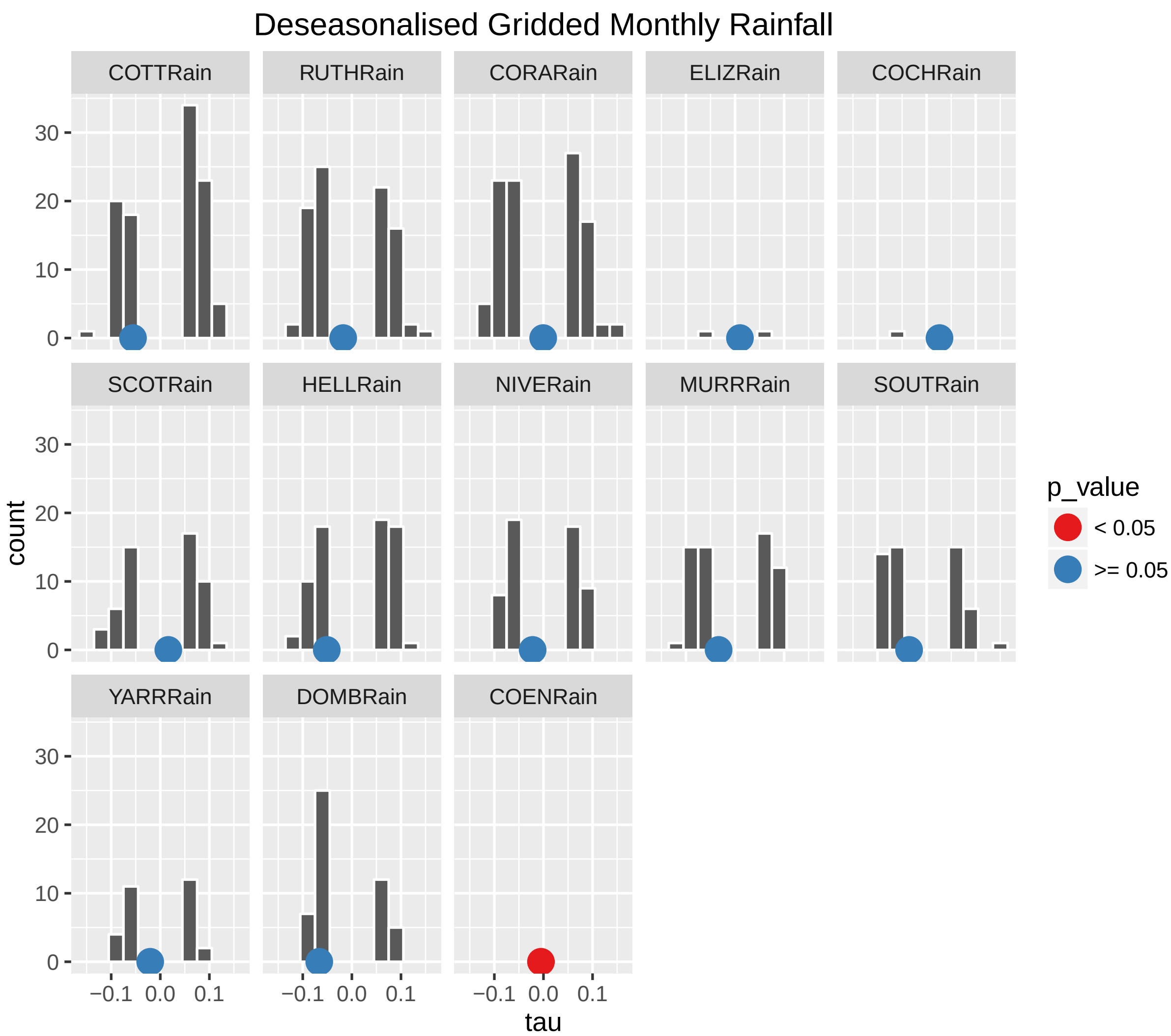

Figure 8: Results of the Bootstrap Mann Kendall analysis on the weekly gridded rainfall data. The dot shows the actual result, while the histogram bars show the results of the bootstrap analysis

Mean Maximum temperature

```
MK_list <- list()
for (i in 1:500) {
  # reorganise the list elements
  series <- sample(1:nyears(maxT_monthly_xts),nyears(maxT_monthly_xts))
  for (j in 1:length(series)) {
    if (j==1) {
      new_df <- as.data.frame(split_maxT[[series[j]]])
    } else {
      new_df <- rbind(new_df,as.data.frame(split_maxT[[series[j]]]))
```

```r
    }
  }
  # rbind to dataframe
  # run mann kendall on the columns and store the results
  mk_r <- apply(new_df,2,MannKendall)

  MK_list[[i]] <- do.call(cbind,mk_r)
}

MK_df <- do.call(rbind,MK_list)

# prepare a dataframe for plotting
pvalues <- subset(MK_df, rownames(MK_df)=="sl")
tau <- subset(MK_df, rownames(MK_df)=="tau")

sig_set <- list()

for (i in 1:ncol(pvalues)) {
  set <- data.frame(pvalue=as.numeric(pvalues[,i]),
                    tau=as.numeric(tau[,i]),catch=rep(colnames(MK_df)[i],nrow(tau)))
  sig_set[[i]] <- set[set$pvalue<0.05,]
}

sig_set_a <- do.call(rbind,sig_set)
sig_set_a$type <- rep("bootstrap",nrow(sig_set_a))
real <- do.call(rbind,apply(maxT_monthly_xts,2,MannKendall))
real_df_maxT_m <- data.frame(pvalue = as.numeric(real[,2]),
                      tau = as.numeric(real[,1]),
                      catch=rownames(real),type=rep("real",nrow(real)))
sig_set_f <- rbind(sig_set_a,real_df_maxT_m)
# A histogram of taus
hp <- ggplot(sig_set_a, aes(x=tau)) +
  geom_histogram(binwidth=0.03,colour="white")

# Histogram of significant tau's, facet wrapped by catchment
# With panels that have the same scaling, but different range
# (and therefore different physical sizes)
hp <- hp + facet_wrap(~ catch,ncol=5)
# add a red point for the real slope from the data
p_value <- ifelse(real_df_maxT_m$pvalue<0.1,"< 0.05",">= 0.05")
hp <- hp + geom_point(data=real_df_maxT_m,aes(x=tau, y=0,colour=p_value),
                shape=16,size=5) +
  scale_colour_brewer(palette="Set1") +
  facet_wrap(~ catch,ncol=5)+ ggtitle("Deseasonalised Monthly Maximum Temperature") #+
  #scale_colour_discrete(name="p-value",
  #                     breaks=c("darkblue","red"),
  #                     labels=c("<0.1",">=0.1"))
print(hp)
```

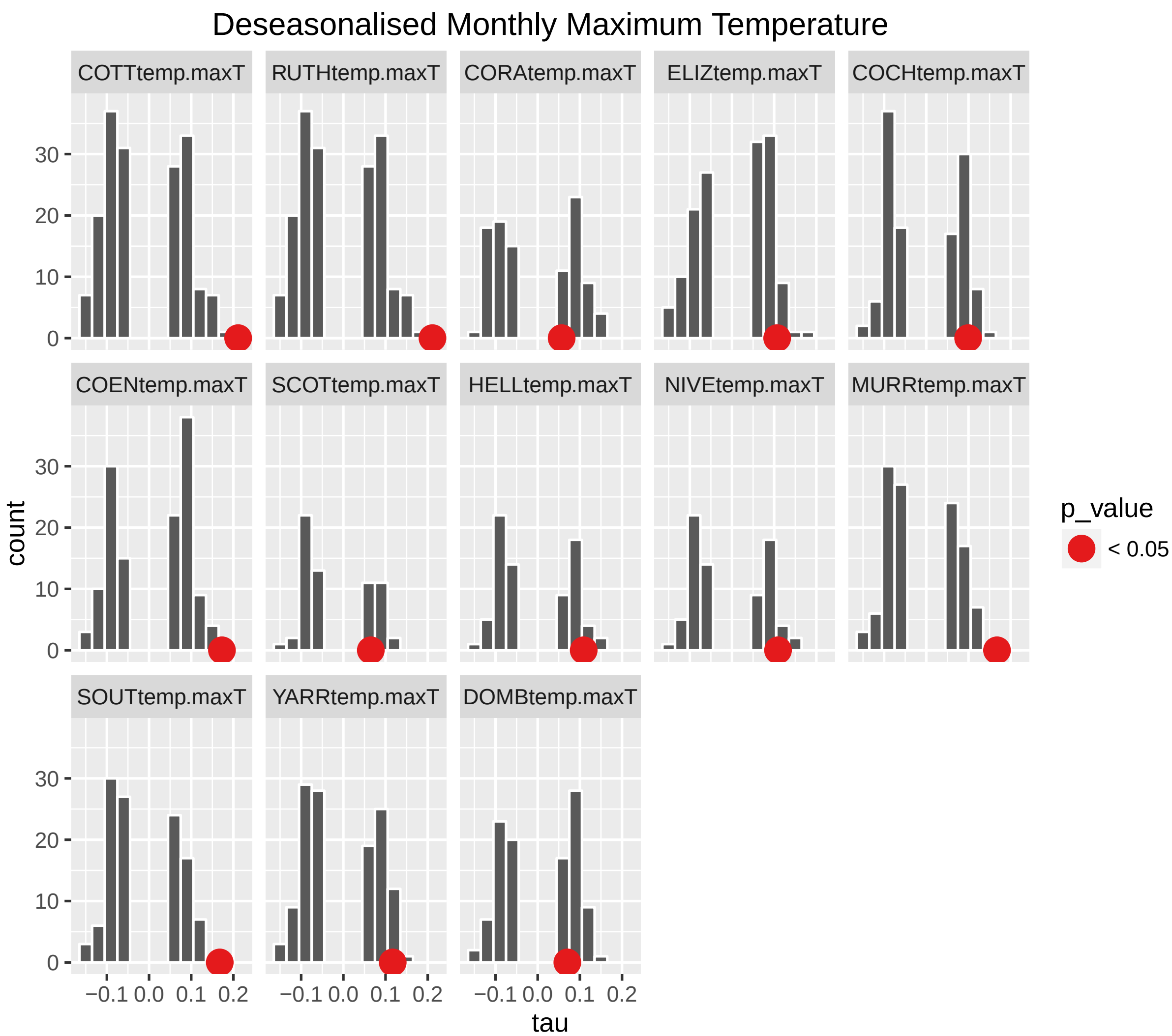

Figure 9: Results of the Bootstrap Mann Kendall analysis on the monthly average maximum temperature data. The dot shows the actual result, while the histogram bars show the results of the bootstrap analysis

Maximum Maximum temperature

This is the analysis for the monthly maximum in the maximum temperature as comparison to see if the assumption of the average maximum temperature is not biased.

```
MK_list <- list()
for (i in 1:500) {
  # reorganise the list elements
  series <- sample(1:nyears(m_maxT_monthly_xts),nyears(m_maxT_monthly_xts))
  for (j in 1:length(series)) {
    if (j==1) {
      new_df <- as.data.frame(split_mmaxT[[series[j]]])
    } else {
      new_df <- rbind(new_df,as.data.frame(split_mmaxT[[series[j]]]))
    }
  }
  # rbind to dataframe
  # run mann kendall on the columns and store the results
  mk_r <- apply(new_df,2,MannKendall)

  MK_list[[i]] <- do.call(cbind,mk_r)
}

MK_df <- do.call(rbind,MK_list)

# prepare a dataframe for plotting
pvalues <- subset(MK_df, rownames(MK_df)=="sl")
tau <- subset(MK_df, rownames(MK_df)=="tau")

sig_set <- list()

for (i in 1:ncol(pvalues)) {
  set <- data.frame(pvalue=as.numeric(pvalues[,i]),
                    tau=as.numeric(tau[,i]),catch=rep(colnames(MK_df)[i],nrow(tau)))
  sig_set[[i]] <- set[set$pvalue<0.05,]
}

sig_set_a <- do.call(rbind,sig_set)
sig_set_a$type <- rep("bootstrap",nrow(sig_set_a))
real <- do.call(rbind,apply(m_maxT_monthly_xts,2,MannKendall))
real_df_mmaxT_m <- data.frame(pvalue = as.numeric(real[,2]),
                      tau = as.numeric(real[,1]),
                      catch=rownames(real),type=rep("real",nrow(real)))
sig_set_f <- rbind(sig_set_a,real_df_mmaxT_m)
# A histogram of taus
hp <- ggplot(sig_set_a, aes(x=tau)) +
  geom_histogram(binwidth=0.03,colour="white")

# Histogram of significant tau's, facet wrapped by catchment
# With panels that have the same scaling, but different range
# (and therefore different physical sizes)
hp <- hp + facet_wrap(~ catch,ncol=5)
# add a red point for the real slope from the data
p_value <- ifelse(real_df_mmaxT_m$pvalue<0.1,"< 0.05",">= 0.05")
```

```
hp <- hp + geom_point(data=real_df_mmaxT_m,aes(x=tau, y=0,colour=p_value),
               shape=16,size=5) +
  scale_colour_brewer(palette="Set1") +
  facet_wrap(~ catch,ncol=5)+ ggtitle("Monthly Maximum Maximum Temperature") #+
print(hp)
```

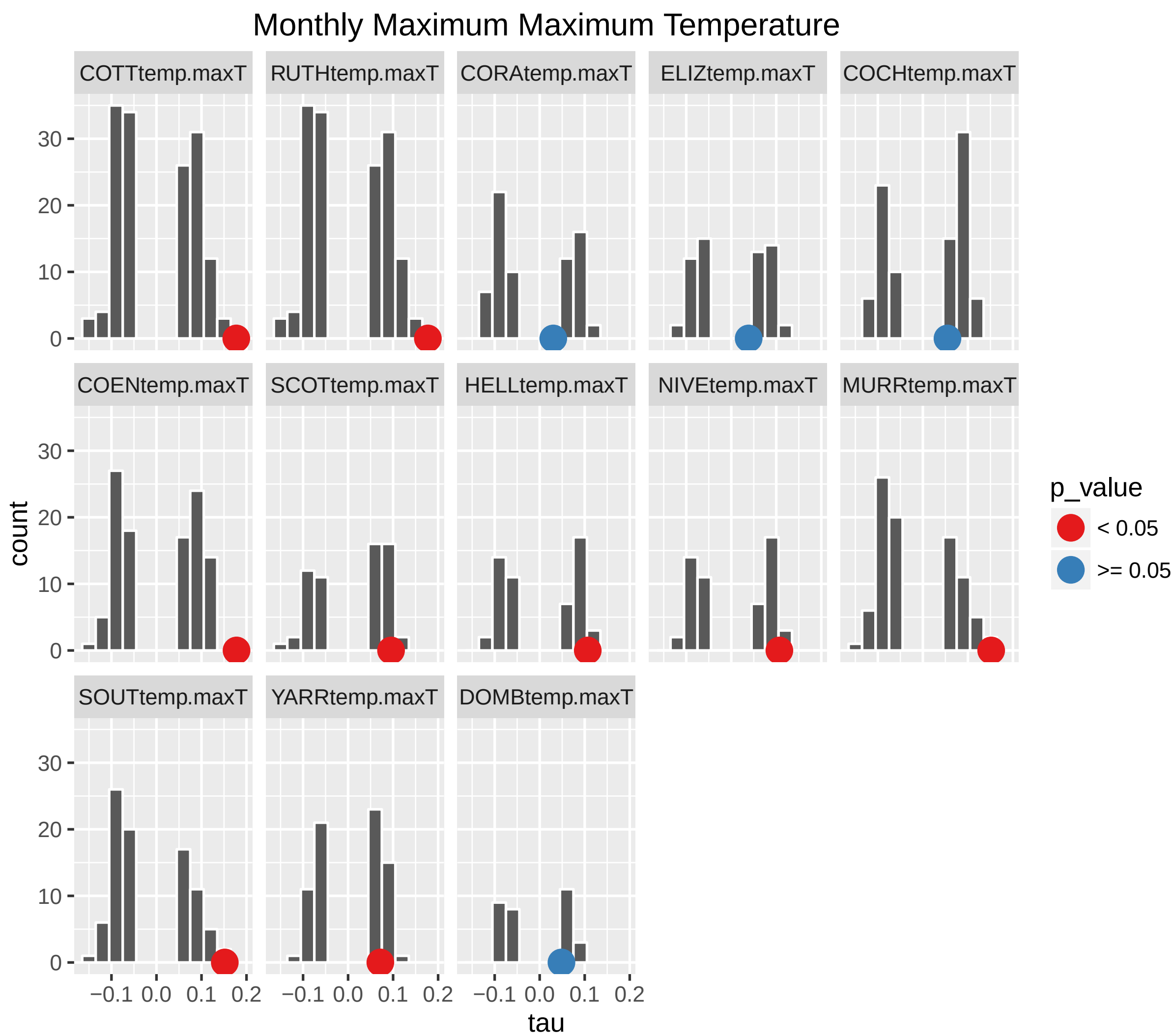

Figure 10: Results of the Bootstrap Mann Kendall analysis on the weekly maximum maximum temperature data. The dot shows the actual result, while the histogram bars show the results of the bootstrap analysis

Summary of monthly analysis in a table

This section brings together the data from all the monthlyly Mann Kendall analyses and puts this into a table. These results are only for comparison and are not repeated in the manuscript.

```
monthlyTable <- data.frame(Catchment = Stations[,1],
                           "Streamflow tau" = real_df_f_m$tau,
                           "Streamflow p-value" = real_df_f_m$pvalue,
                           "Rainfall tau" = real_df_r_m$tau,
```

```
                           "Rainfall p-value" = real_df_r_m$pvalue,
                           "Grid Rainfall tau" = real_df_grdr_m$tau,
                           "Grid Rainfall p-value" = real_df_grdr_m$pvalue,
                           "Mean max T tau" = real_df_maxT_m$tau,
                           "Mean max T p-value" = real_df_maxT_m$pvalue,
                           "Max max T tau" = real_df_mmaxT_m$tau,
                           "Max max T p-value" = real_df_mmaxT_m$pvalue)
pander(monthlyTable,
       caption="Mann-Kendall test results on de-seasonalised monthly time series.
       p-values are  considered significant at the 5% level.")
```

Table 4: Mann-Kendall test results on de-seasonalised monthly time series. p-values are considered significant at the 5% level. (continued below)

Catchment	Streamflow.tau	Streamflow.p.value	Rainfall.tau
COTT	-0.1833	1.235e-09	-0.1242
RUTH	-0.2489	1.568e-16	-0.07129
CORA	-0.1074	0.000372	-0.03587
ELIZ	-0.02029	0.502	-0.1136
COCH	-0.02331	0.4397	-0.03568
COEN	-0.04006	0.1842	-0.006566
SCOT	-0.02641	0.3814	-0.03507
HELL	-0.1256	3.125e-05	-0.09105
NIVE	-0.04597	0.1276	-0.05863
MURR	-0.2416	1.14e-15	-0.03432
SOUT	-0.1742	7.703e-09	-0.1625
YARR	-0.1779	4.054e-09	-0.03274
DOMB	-0.06746	0.02534	-0.04746

Table 5: Table continues below

Rainfall.p.value	Grid.Rainfall.tau	Grid.Rainfall.p.value	Mean.max.T.tau
0.007312	-0.05556	0.06555	0.2113
0.03217	-0.0177	0.5575	0.2113
0.2478	-0.0005796	0.9849	0.05681
0.0007232	0.009979	0.741	0.1071
0.3286	0.02616	0.3859	0.09922
0.8357	-0.005084	0.8664	0.1729
0.2505	0.01638	0.5874	0.06507
0.009725	-0.05115	0.08998	0.1091
0.07885	-0.02206	0.4648	0.1091
0.2737	-0.03347	0.2674	0.1676
1.142e-06	-0.0357	0.2367	0.1676
0.2833	-0.02055	0.4959	0.1171
0.2034	-0.0664	0.02773	0.07035

Mean.max.T.p.value	Max.max.T.tau	Max.max.T.p.value
0	0.1773	0
0	0.1773	0

Mean.max.T.p.value	Max.max.T.tau	Max.max.T.p.value
0.09182	0.03023	0.3699
0.0004582	0.03859	0.2069
0.003699	0.05465	0.1099
1.192e-07	0.1778	0
0.04312	0.09546	0.003005
0.001151	0.1068	0.00146
0.001151	0.1068	0.00146
0	0.1517	7.153e-07
0	0.1517	7.153e-07
0.0001096	0.07166	0.01788
0.02627	0.04839	0.1264