

AERO: preliminary analysis of Niger measles data, 1995-2005

```
dim(d)
```

```
## [1] 40 573
```

```
rownames(d)<-d[,1]  
d<-d[,-1]  
rownames(d)
```

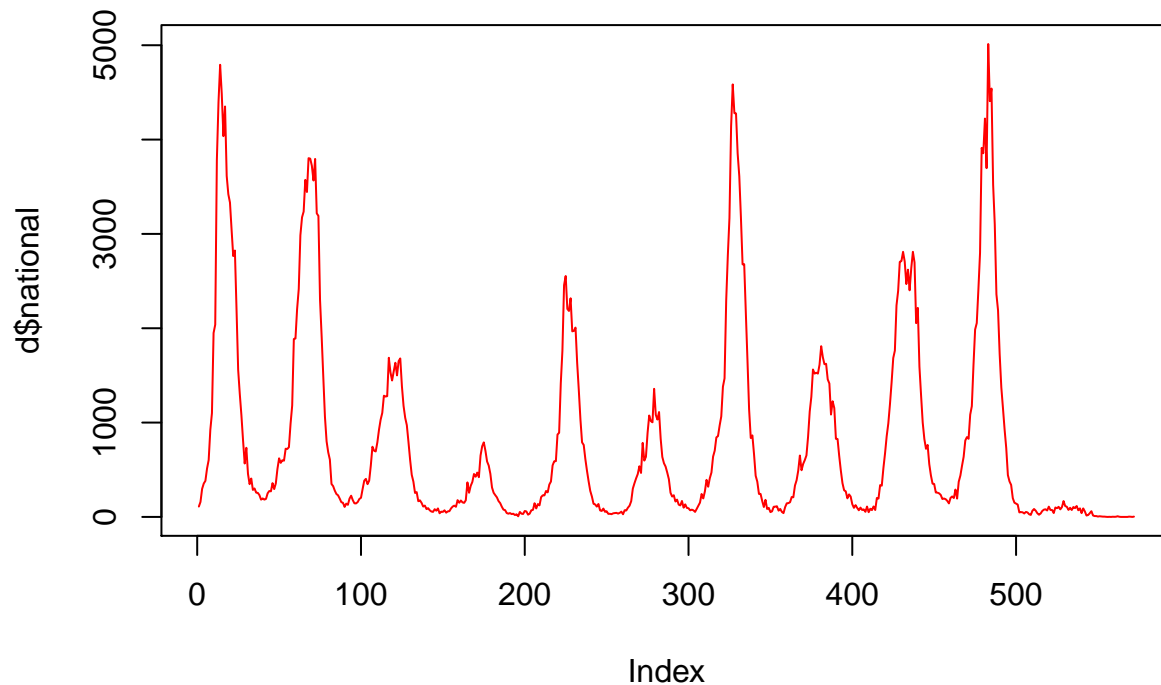
```
## [1] "Agadez (City)" "Tchirozerine" "Arlit"  
## [4] "Bilma" "Diffa" "Maine-Sora"  
## [7] "N'Guimi" "Birni N'Gaour" "Dogon-Doutchi"  
## [10] "Dosso" "Gaya" "Loga"  
## [13] "Aguie" "Dakoro" "Guidan-Roundji"  
## [16] "Madarounfa" "Maradi (City)" "Mayahi"  
## [19] "Tessaoua" "Filingue" "Kollo"  
## [22] "Ouallam" "Say" "Tera"  
## [25] "Tillabery" "Abalak" "Birni N'Konni"  
## [28] "Bouza" "Illela" "Keita"  
## [31] "Madaoua" "Tahoua" "Tchin-Tabaraden"  
## [34] "Goure" "Magaria" "Matamey"  
## [37] "Miria" "Tannout" "Zinder (City)"  
## [40] "Niamey (City)"
```

The data frame contains 40 regions and 572 time points spanning 11 years (presumably weeks starting 1 Jan 1995, since $11 \times 52 = 572$). First, transpose the data frame and name the weeks 1 through 572

```
d<-as.data.frame(t(d))  
rownames(d)<-1:572
```

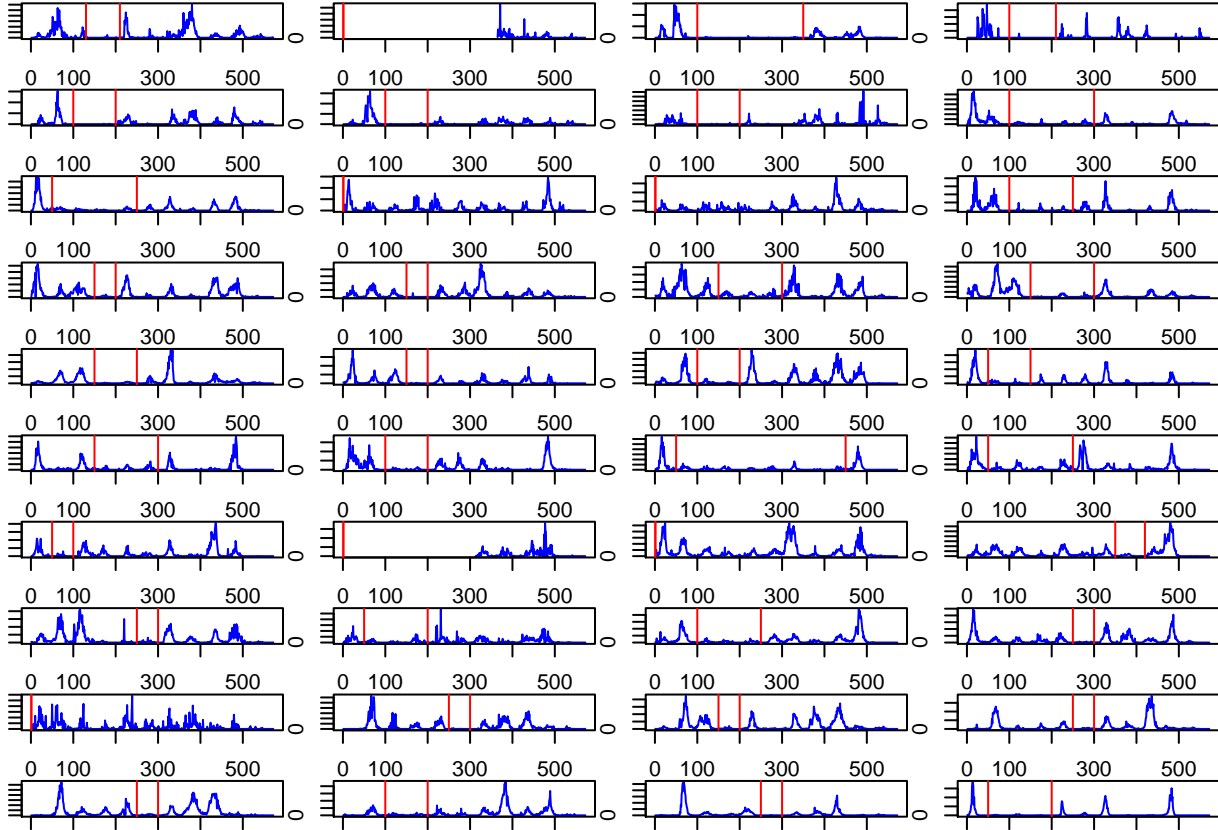
Each entry in the data frame is (presumably) the number of reported measles cases in that region in that week. Now we add a column for the country-level reported measles cases and inspect this

```
d$national<-rowSums(d,na.rm=T)  
plot(d$national,type="l",col="red")
```



We observe a seasonal (annual) incidence pattern with considerable variation in total cases per year. Many regions have >1 year intervals between large outbreaks. This is posited to be due to weak coupling between regions, resulting in a distribution of import times >1 year. During this time, regions may be building up susceptible numbers (mainly through births) driving the system to criticality.

```
fromto<-matrix(c(130,210,1,1,100,350,100,210,100,200,100,200,100,200,100,300,50,250,1,1,1,1,100,250,150,
par(mfrow=c(10,4),mar=c(1,1,1,1))
for (i in 1:40){
  plot(d[,i],type="l",col="blue")
  abline(v=fromto[i,1],col="red")
  abline(v=fromto[i,2],col="red")
}
```



Obtain drift time series and randomized versions of same. Drift is time series we suspect may have R_0 drifting towards 1. They are the red intervals in blue time series above, excluding those with NA, or exclusion by visual inspection (e.g annual peaks), or if very few cases (max incidence <10).

```
lappend <- function(lst, obj) {
  lst[[length(lst)+1]] <- obj
  return(lst)
}

inter.epi.2<-list()
for (i in 1:40){
  if ((is.na(min(d[,i]))==F)&(fromto[i,1]>1)&(max(d[fromto[i,1]:fromto[i,2],i])>=10)){
    this.ts<-d[fromto[i,1]:fromto[i,2],i]
    #this.ts<-tail(this.ts,50)
    inter.epi.2<-lappend(inter.epi.2,this.ts)
  }
}

non.inter.epi.2<-list() #chops up time series
for (i in 1:length(inter.epi.2)){
  new.order<-sample(1:length(inter.epi.2[[i]]),replace=F)
  scrambled<-inter.epi.2[[i]][order(new.order)]
  non.inter.epi.2<-lappend(non.inter.epi.2,scrambled)
}
```

Now we see if the Kendall's tau statistic for a given EWS statistic is different for time series of the two lists

```
library(VGAM)
```

```
## Loading required package: stats4
```

```
## Loading required package: splines
```

```
library(spaero)
cf.tau<-NULL
ny<-1
for (i in 1:length(inter.epi.2)){
  nw<-length(inter.epi.2[[i]])
  st1<-get_stats(inter.epi.2[[i]],center_kernel="uniform",center_trend="local_constant",center_bandwidth=ny)
  cf.tau<-rbind(cf.tau,c(1,kendall.tau(seq(1,ny*nw,1),st1$stats$variance),kendall.tau(seq(1,ny*nw,1),st1$stats$variance)))
}
for (i in 1:length(non.inter.epi.2)){
  nw<-length(non.inter.epi.2[[i]])
  st1<-get_stats(non.inter.epi.2[[i]],center_kernel="uniform",center_trend="local_constant",center_bandwidth=ny)
  cf.tau<-rbind(cf.tau,c(0,kendall.tau(seq(1,ny*nw,1),st1$stats$variance),kendall.tau(seq(1,ny*nw,1),st1$stats$variance)))
}

cf.tau<-as.data.frame(cf.tau)
names(cf.tau)<-c("label",names(st1$stats))
cf.tau.orig<-cf.tau
my.aucs<-NULL
for (i in 1:9){
  tmp<-cf.tau[,c(1,(i+1))]
  real.name<-names(tmp)[2]
  names(tmp)[2]<- "stat.of.interest"
  tmp<-tmp[order(tmp$stat.of.interest),]
  pos1<-which(tmp$label==1)#position in ranked list of 1's
  n1<-length(pos1)#number of 1's in vector
  pos0<-which(tmp$label==0)
  n0<-length(pos0)
  auc<-0
  for (j in 1:n1){
    auc<-auc+(pos1[j]-j)
  }
  auc<-auc/(n0*n1)
  my.aucs<-rbind(my.aucs,c(real.name, auc))
}

my.aucs<-as.data.frame(my.aucs)
names(my.aucs)<-c("EWS", "AUC")
my.aucs$AUC<-as.numeric(as.character(my.aucs$AUC))
```

```
bstrap.num<-100
for (b in 1:(bstrap.num-1)){# add to original to make n=10,100 bootstrap
  my.subsample.nonepi<-sample(1:length(non.inter.epi.2),replace=T)
  non.inter.epi.3<-non.inter.epi.2[my.subsample.nonepi]
  my.subsample.epi<-sample(1:length(inter.epi.2),replace=T)
  inter.epi.3<-inter.epi.2[my.subsample.epi]
```

```

cf.tau<-NULL
for (i in 1:length(inter.epi.3)){
  nw<-length(inter.epi.3[[i]])
  st1<-get_stats(inter.epi.3[[i]],center_kernel="uniform",center_trend="local_constant",center_bandwidth=1)
  cf.tau<-rbind(cf.tau,c(1,kendall.tau(seq(1,ny*nw,1),st1$stats$variance),kendall.tau(seq(1,ny*nw,1),st1$stats$variance)))
}
for (i in 1:length(non.inter.epi.3)){
  nw<-length(non.inter.epi.3[[i]])
  st1<-get_stats(non.inter.epi.3[[i]],center_kernel="uniform",center_trend="local_constant",center_bandwidth=1)
  cf.tau<-rbind(cf.tau,c(0,kendall.tau(seq(1,ny*nw,1),st1$stats$variance),kendall.tau(seq(1,ny*nw,1),st1$stats$variance)))
}

cf.tau<-as.data.frame(cf.tau)
names(cf.tau)<-c("label",names(st1$stats))

my.aucs.2<-NULL
for (i in 1:9){
  tmp<-cf.tau[,c(1,(i+1))]
  real.name<-names(tmp)[2]
  names(tmp)[2]<-"stat.of.interest"
  tmp<-tmp[order(tmp$stat.of.interest),]
  pos1<-which(tmp$label==1)#position in ranked list of 1's
  pos0<-which(tmp$label==0)
  n1<-length(pos1)#number of 1's in vector
  n0<-length(pos0)
  auc<-0
  for (j in 1:n1){
    auc<-auc+(pos1[j]-j)
  }
  auc<-auc/(n0*n1)
  my.aucs.2<-rbind(my.aucs.2,c(real.name,auc))
}
my.aucs.2<-as.data.frame(my.aucs.2)
names(my.aucs.2)<-c("EWS","AUC")
my.aucs.2$AUC<-as.numeric(as.character(my.aucs.2$AUC))
my.aucs<-merge(my.aucs,my.aucs.2,by="EWS")

}

par(mfrow=c(1,1),mar=c(5,5,5,1))

ews.list<-list()
my.aucs.vals<-t(my.aucs[,2:(bstrap.num+1)])
for (i in 1:9){
  ews.list<-lappend(ews.list,unname(my.aucs.vals[,i]))
}

boxplot(ews.list,names=NA,cex.axis=1.0,las=1,ylim=c(0,1),col="gray",ylab="AUC")
lines(c(0.12,9.88),c(0.5,0.5),col="red")
text(1:9,rep(0.2,9),my.aucs$EWS,srt=90,col="gray20",cex=0.7)

```

