Q1: Are lake and terrestrial primary productivity coherent?

Jonathan Walter, Grace Wilkinson, Rachel Fleck, Michael Pace

4/17/2019

This document organizes for openness and reproducibility analyses of the temporal coherence of interannual variation in lake primary productivity with terrestrial primary productivity in the landscape surrounding the lake.

Data import

Data produced in 'ms1__prep.Rmd' are loaded.

```
load("~/Box Sync/NSF EAGER Synchrony/Data/RData files/ms1_analysis_inprogress1_v10873_2.RData")
```

```
any(sapply(analysislakes$lakedata, function(x){any(is.infinite(x))}))
```

```
## [1] FALSE
```

```
any(sapply(analysislakes$lakedata, function(x){any(is.na(x))}))
```

```
## [1] TRUE
```

```
which(sapply(analysislakes$lakedata, function(x){any(is.na(x))}))
```

```
## 7545 7790 7970 8271
##   70   72   76   77
```

```
analysislakes$lakeinfo[which(sapply(analysislakes$lakedata, function(x){any(is.na(x))})),]
```

```
##      lagoslakeid         gnis_name  nhd_lat  nhd_long lake_area_ha
## 7445        7545        Alton Pond 41.44294 -71.71835    17.843701
## 7689        7790     Watchaug Pond 41.38381 -71.69161   232.391660
## 7867        7970       Yawgoo Pond 41.51113 -71.57300    60.724131
## 8165        8271 Meadow Brook Pond 41.44110 -71.69034     9.808244
##      lake_perim_meters nhd_ftype nhd_fcode hu4_zoneid hu12_zoneid
## 7445          3746.121       390     39004     HU4_10  HU12_16347
## 7689          8397.096       390     39010     HU4_10  HU12_17513
## 7867          3195.857       390     39010     HU4_10  HU12_17512
## 8165          2030.864       390     39004     HU4_10  HU12_17513
##      state_zoneid elevation_m start  end
## 7445      State_8     14.3702  1989 2010
## 7689      State_8     11.1300  1989 2010
## 7867      State_8     36.1000  1989 2010
## 8165      State_8     16.2900  1989 2010
```

```
# image(accndvi)
# points(lakepts.prj[which(sapply(analysislakes$lakedata, function(x){any(is.na(x))})),])

dbuff[which(sapply(analysislakes$lakedata, function(x){any(is.na(x))}))]
```

```
## [1] 2500.000 4939.589 2500.000 2500.000
```

```
analysislakes$lakeinfo<-analysislakes$lakeinfo[!sapply(analysislakes$lakedata, function(x){any(is.na(x)
analysislakes$lakedata<-analysislakes$lakedata[!sapply(analysislakes$lakedata, function(x){any(is.na(x)

analysislakes$lakeinfo$tslength<-analysislakes$lakeinfo$end-analysislakes$lakeinfo$start+1
# analysislakes$lakedata<-analysislakes$lakedata[!analysislakes$lakeinfo$tslength < 20]
# analysislakes$lakeinfo<-analysislakes$lakeinfo[!analysislakes$lakeinfo$tslength < 20,]
```

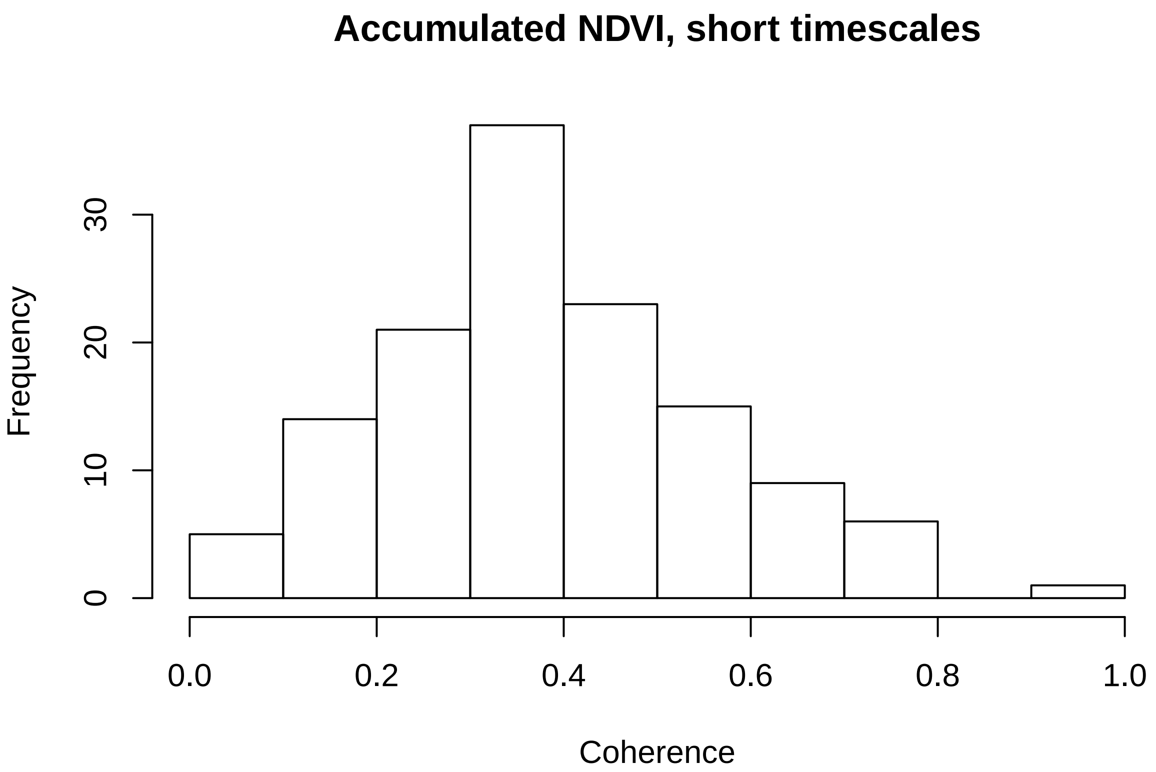
```
source("~/GitHub/AquaTerrSynch/AnalysisCode/bandtest_coh.R")

tsranges<-rbind(c(2,4),c(4,Inf),c(2,Inf))

coh.chlaXaccndvi<-NULL
#coh.chlaXmaxndvi<-NULL

for(lind in 1:length(analysislakes$lakedata)){
  lakedat.ii<-cleandat(analysislakes$lakedata[[lind]], as.numeric(colnames(analysislakes$lakedata[[lind
  chlaXaccndvi<-coh(lakedat.ii[1,], lakedat.ii[2,], as.numeric(colnames(analysislakes$lakedata[[lind]])
                   norm="powall", sigmethod="fast", nrand=10000)
#  chlaXmaxndvi<-coh(lakedat.ii[1,], lakedat.ii[3,], as.numeric(colnames(analysislakes$lakedata[[lind]]
#                    norm="powall", sigmethod="fast", nrand=10000)
  for(rind in 1:nrow(tsranges)){
    chlaXaccndvi<-bandtest.coh(chlaXaccndvi, tsranges[rind,])
    #chlaXmaxndvi<-bandtest.coh(chlaXmaxndvi, tsranges[rind,])
  }
  coh.chlaXaccndvi<-rbind(coh.chlaXaccndvi, c(t(as.matrix(chlaXaccndvi$bandp[,3:5]))))
#  coh.chlaXmaxndvi<-rbind(coh.chlaXmaxndvi, c(t(as.matrix(chlaXmaxndvi$bandp[,3:5]))))

}

coh.chlaXaccndvi<-as.data.frame(coh.chlaXaccndvi)
#coh.chlaXmaxndvi<-as.data.frame(coh.chlaXmaxndvi)

colnames(coh.chlaXaccndvi)<-paste0("accndvi",c("p.ts1","phi.ts1","coh.ts1","p.ts2","phi.ts2","coh.ts2",
#colnames(coh.chlaXmaxndvi)<-paste0("maxndvi",c("p.ts1","phi.ts1","coh.ts1","p.ts2","phi.ts2","coh.ts2"

coh.chlaXaccndvi$lagoslakeid<-analysislakes$lakeinfo$lagoslakeid
#coh.chlaXmaxndvi$lagoslakeid<-analysislakes$lakeinfo$lagoslakeid
```

```
#looking for especially coherent lakes did not return easy-to-interpret examples. Proceeding with simul

tmax=50
res=0.1
tt=seq(1,tmax,res)

p1<-2
sig1<-sin(seq(0,2*pi*tmax/p1,length.out=length(tt)))
p2<-10
sig2<-sin(seq(0,2*pi*tmax/p2,length.out=length(tt)))

comb1<-sig1+0.7*sig2+3.5
comb2<-sig1+-0.7*sig2


laymat<-matrix(1,nrow=2,ncol=3)
```

```
laymat[2,]<-2:4

sig3<-sig2[tt<=20]
sig4<-sig3*0.9
sig5<-sin(seq(-pi/2,2*pi*20/p2-(pi/2),length.out=length(tt[tt<=20])))
sig6<-sig3*-1

tiff("~/Box Sync/NSF EAGER Synchrony/Manuscripts/1_CoherenceSpatialVariation/fig1_pedagogical.tif", uni
     res=300, width=6.5, height=4)

layout(laymat)
par(mar=c(1.5,1.5,2,1.5), mgp=c(1,1,0), oma=c(2,2,0,0))

plot(NA,NA,ylim=c(-2,5.2),xlim=range(tt), xlab="", ylab="", xaxt="n",yaxt="n")
lines(tt,comb1,lwd=2)
lines(tt,comb2,lwd=2,col="red")
axis(1, at=c(0,10,20,30,40,50),labels=NA)
axis(2, at=c(-1,1.5,4), labels=NA)
mtext("Timescale specific relationship",3,line=0.25)
text(0.4,4.9,"a)",cex=1.2)

plot(NA,NA,ylim=c(-1,1),xlim=c(0,20),xaxt="n",yaxt="n",xlab="",ylab="")
axis(1, at=c(0,10,20), labels=NA)
axis(2, at=c(-1,0,1), labels=NA)
lines(tt[tt<=20],sig3,lwd=2)
lines(tt[tt<=20],sig4,lwd=2,col="red")
mtext(expression(paste(phi," = 0")))
text(1,0.9,"b)",cex=1.2)

plot(NA,NA,ylim=c(-1,1),xlim=c(0,20),xaxt="n",yaxt="n",xlab="",ylab="")
axis(1, at=c(0,10,20), labels=NA)
axis(2, at=c(-1,0,1), labels=NA)
lines(tt[tt<=20],sig3,lwd=2)
lines(tt[tt<=20],sig5,lwd=2,col="red")
mtext(expression(paste(phi," = ",pi,"/2")))
text(1,0.9,"c)",cex=1.2)

plot(NA,NA,ylim=c(-1,1),xlim=c(0,20),xaxt="n",yaxt="n",xlab="",ylab="")
axis(1, at=c(0,10,20), labels=NA)
axis(2, at=c(-1,0,1), labels=NA)
lines(tt[tt<=20],sig3,lwd=2)
lines(tt[tt<=20],sig6,lwd=2,col="red")
mtext(expression(paste(phi," = ",pi)))
text(1,0.9,"d)",cex=1.2)


mtext("Time", 1, outer=T)
mtext("Signal", 2, outer=T)

dev.off()
```

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

```r
#short timescales
hist(coh.chlaXaccndvi$accndvicoh.ts1, main="Accumulated NDVI, short timescales", xlab="Coherence", ylab=
```




```r
#hist(coh.chlaXmaxndvi$maxndvicoh.ts1, main="Maximum NDVI, short timescales", xlab="Coherence", ylab="F

quantile(coh.chlaXaccndvi$accndvicoh.ts1)
```

```
##         0%        25%        50%        75%       100%
## 0.06293777 0.26880179 0.36586451 0.49214074 0.92134749
```

```r
#quantile(coh.chlaXmaxndvi$maxndvicoh.ts1)

alpha=0.05
sum(coh.chlaXaccndvi$accndvip.ts1<alpha)/nrow(coh.chlaXaccndvi)
```

```
## [1] 0.06870229
```

```r
#sum(coh.chlaXmaxndvi$maxndvip.ts1<alpha)/nrow(coh.chlaXmaxndvi)

print(cbind(coh.chlaXaccndvi$lagoslakeid, coh.chlaXaccndvi$accndvip.ts1)[coh.chlaXaccndvi$accndvip.ts1<
```

```
##         [,1]       [,2]
##  [1,]   2851 0.03249675
##  [2,]   3370 0.00009999
##  [3,]   6075 0.04199580
##  [4,]   6547 0.00809919
##  [5,]   7523 0.02039796
##  [6,]  28836 0.00679932
##  [7,]  72641 0.00629937
##  [8,] 133500 0.02809719
##  [9,]   4909 0.03279672
```

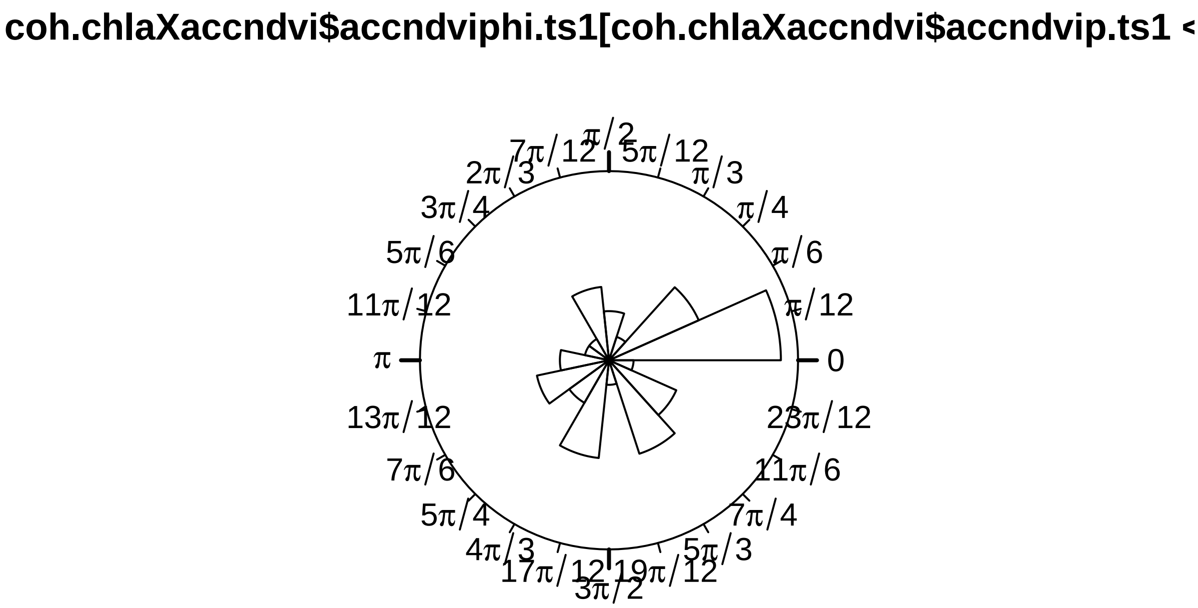
```r
print(cbind(coh.chlaXaccndvi$lagoslakeid, coh.chlaXaccndvi$accndvip.ts2)[coh.chlaXaccndvi$accndvip.ts2<a
```

```
##         [,1]       [,2]
##  [1,]   3370 0.01429857
##  [2,]   3834 0.00849915
##  [3,]   4243 0.00009999
##  [4,]   4416 0.01069893
##  [5,]   4434 0.02659734
##  [6,]   5895 0.02249775
##  [7,]   6199 0.03799620
##  [8,]   7523 0.02549745
##  [9,]   8369 0.03119688
## [10,]  39489 0.04539546
## [11,] 115040 0.00779922
## [12,] 133500 0.02279772
## [13,]  14815 0.00349965
## [14,] 102115 0.00289971
## [15,]   5463 0.03079692
```

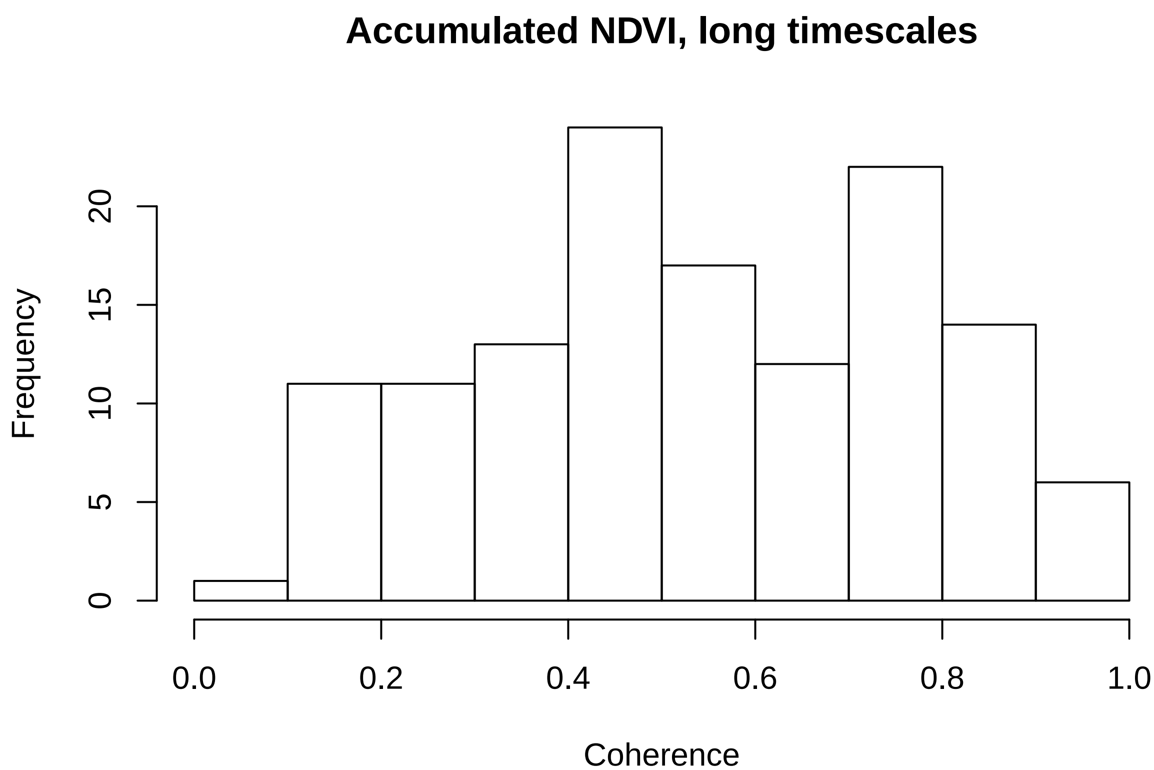
```r
cor(coh.chlaXaccndvi$accndvicoh.ts1,coh.chlaXaccndvi$accndvicoh.ts2)
```

```
## [1] 0.02535991
```

```r
# print(coh.chlaXaccndvi$accndviphi.ts1[coh.chlaXaccndvi$accndvip.ts1<alpha]/pi) #only pattern is that
# print(coh.chlaXmaxndvi$maxndviphi.ts1[coh.chlaXmaxndvi$maxndvip.ts1<alpha]/pi)

phicls<-c(-1,-.75,-0.25,0.25,0.75,1)

# hist(coh.chlaXaccndvi$accndviphi.ts1[coh.chlaXaccndvi$accndvip.ts1<0.2]/pi, main="Accumulated NDVI, s
rose(coh.chlaXaccndvi$accndviphi.ts1[coh.chlaXaccndvi$accndvip.ts1<0.3], unit="radian",
     breaks=seq(0,2*pi,length.out=16))
```




```r
#hist(coh.chlaXmaxndvi$maxndviphi.ts1[coh.chlaXmaxndvi$maxndvip.ts1<0.2]/pi, main="Maximum NDVI, short

#long timescales
hist(coh.chlaXaccndvi$accndvicoh.ts2, main="Accumulated NDVI, long timescales", xlab="Coherence", ylab=
```




```r
#hist(coh.chlaXmaxndvi$maxndvicoh.ts2, main="Maximum NDVI, long timescales", xlab="Coherence", ylab="Fr

quantile(coh.chlaXaccndvi$accndvicoh.ts2)
```

```
##         0%        25%        50%        75%       100%
## 0.07654579 0.37832224 0.53015425 0.73194394 0.95309993
```

```r
#quantile(coh.chlaXmaxndvi$maxndvicoh.ts2)

alpha=0.05
sum(coh.chlaXaccndvi$accndvip.ts2<alpha)/nrow(coh.chlaXaccndvi)
```

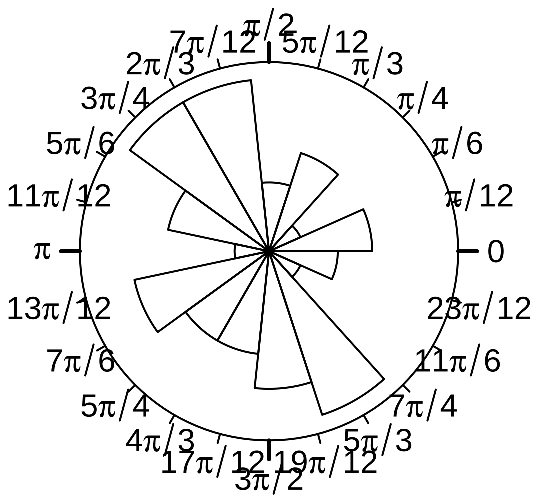
```
## [1] 0.1145038
```

```r
#sum(coh.chlaXmaxndvi$maxndvip.ts2<alpha)/nrow(coh.chlaXmaxndvi)

print(coh.chlaXaccndvi$accndviphi.ts2[coh.chlaXaccndvi$accndvip.ts2<alpha]/pi)
```

```
##  [1]  0.73992662 -0.90124264  0.53888310  0.27521830 -0.95606597
##  [6]  0.52712388 -0.68599147 -0.91773670 -0.29662433 -0.40541240
## [11] -0.22793834 -0.04840174 -0.57269988 -0.52827164  0.88106993
```

```r
#print(coh.chlaXmaxndvi$maxndviphi.ts2[coh.chlaXmaxndvi$maxndvip.ts2<alpha]/pi)

# hist(coh.chlaXaccndvi$accndviphi.ts2[coh.chlaXaccndvi$accndvip.ts2<0.2]/pi, main="Accumulated NDVI, l
rose(coh.chlaXaccndvi$accndviphi.ts2[coh.chlaXaccndvi$accndvip.ts2<0.3], unit="radian",
     breaks=seq(0,2*pi,length.out=16))
```

coh.chlaXaccndvi$accndviphi.ts2[coh.chlaXaccndvi$accndvip.ts2 <



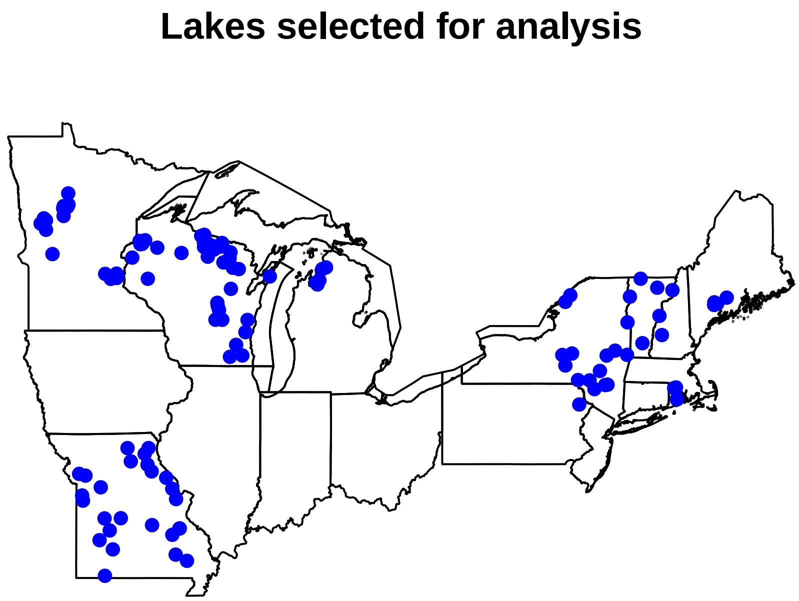

```
#hist(coh.chlaXmaxndvi$maxndviphi.ts1[coh.chlaXmaxndvi$maxndvicoh.ts2>0.6]/pi, main="Maximum NDVI, shor

tiff("~/Box Sync/NSF EAGER Synchrony/Manuscripts/1_CoherenceSpatialVariation/fig2_distributions_2019103
     res=300, width=6.5, height=6.5)

par(mar=c(3,3,2,1),mgp=c(1.7,0.5,0),mfrow=c(2,2),cex.main=0.9)

hist(coh.chlaXaccndvi$accndvicoh.ts1, main="Short timescale coherence", xlab="Coherence", ylab="Frequen
text(par()$usr[1]+.05,0.95*par()$usr[4],"a)")
hist(coh.chlaXaccndvi$accndvicoh.ts2, main="Long timescale coherence", xlab="Coherence", ylab="Frequency
text(par()$usr[1]+.05,0.95*par()$usr[4],"b)")

par(mar=c(1,1,2,1))
rose(coh.chlaXaccndvi$accndviphi.ts1[coh.chlaXaccndvi$accndvip.ts1<0.3], unit="radian", col="lightgrey"
     breaks=c(0,pi/4,pi/2,3*pi/4,pi,5*pi/4,3*pi/2,7*pi/4,2*pi), main="Short timescale phases",
       at=c(0,pi/4,pi/2,3*pi/4,pi,-3*pi/4,-pi/2,-pi/4))
text(0.9*par()$usr[1],0.95*par()$usr[4],"c)")
rose(coh.chlaXaccndvi$accndviphi.ts2[coh.chlaXaccndvi$accndvip.ts2<0.3], unit="radian", col="lightgrey"
     breaks=c(0,pi/4,pi/2,3*pi/4,pi,5*pi/4,3*pi/2,7*pi/4,2*pi), main="Long timescale phases",
     at=c(0,pi/4,pi/2,3*pi/4,pi,-3*pi/4,-pi/2,-pi/4))
text(0.9*par()$usr[1],0.95*par()$usr[4],"d)")

dev.off()
```

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

```
states<-readOGR("~/Box Sync/NSF EAGER Synchrony/Data/statesp020.shp")
```

```
## OGR data source with driver: ESRI Shapefile
## Source: "/Users/jonathanwalter/Box Sync/NSF EAGER Synchrony/Data/statesp020.shp", layer: "statesp020
## with 2895 features
## It has 9 fields
## Integer64 fields read as strings:  STATESP020 DAY_ADM YEAR_ADM
```

```
getstates<-c("Minnesota", "Iowa", "Wisconsin", "Illinois", "Missouri", "Michigan", "Indiana", "Ohio", "
lagosstates<-states[states@data$STATE %in% getstates,]

plot(lagosstates, main="Lakes selected for analysis")
points(analysislakes$lakeinfo$nhd_long, analysislakes$lakeinfo$nhd_lat, pch=16, cex=1, col="blue")
```

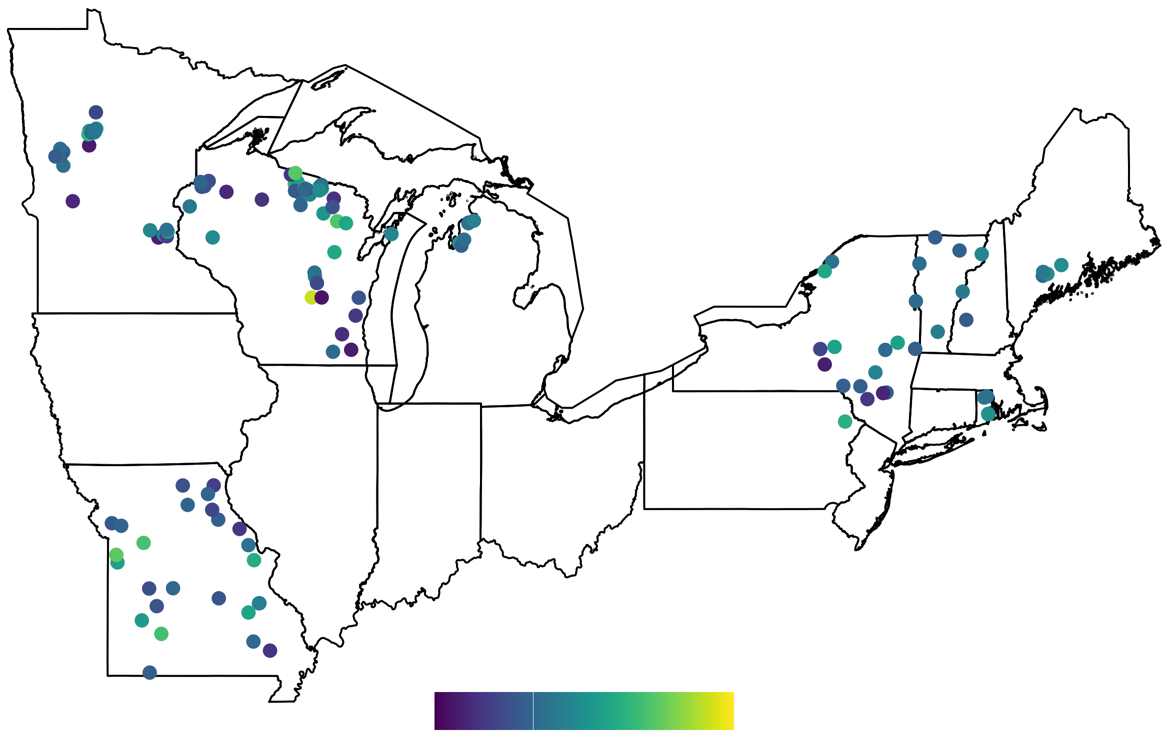



```
cohplotdata<-left_join(analysislakes$lakeinfo, coh.chlaXaccndvi, by="lagoslakeid")

pal<-viridis(100)

par(mar=c(1,0,2,0))

plot(lagosstates, main="Lakes by short timescale coherence")
par()$usr
```

```
## [1] -98.49241 -65.70056  35.45947  49.92124
```

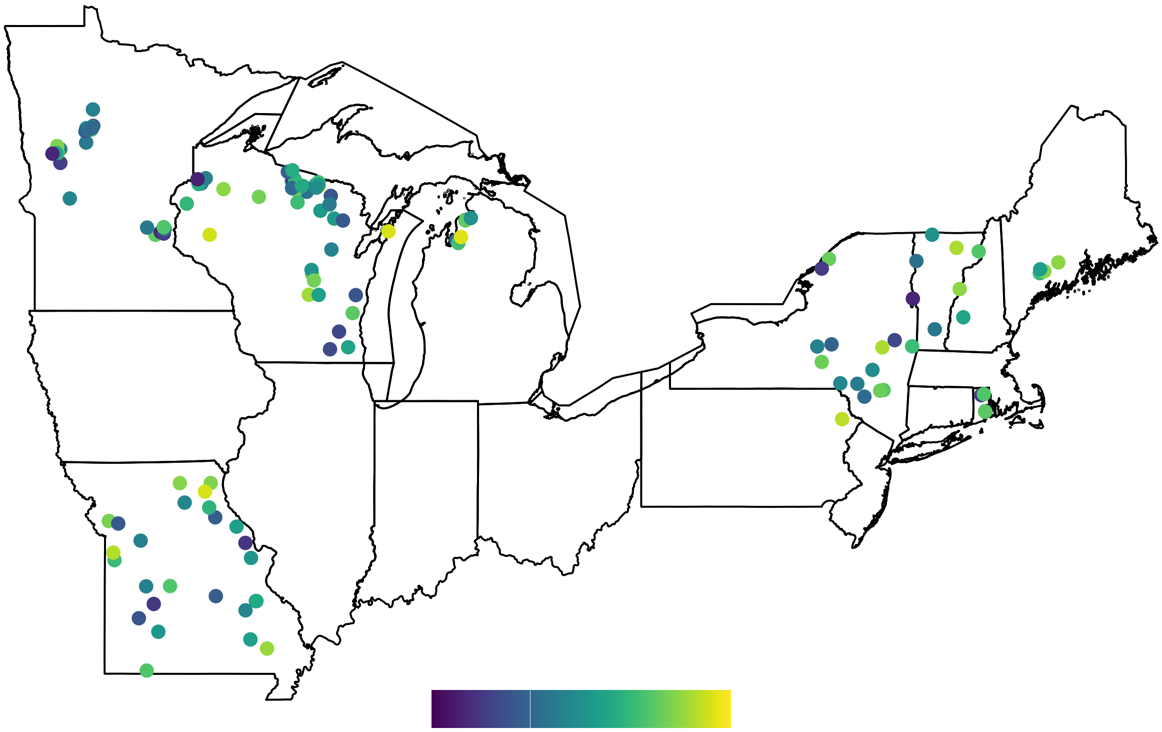
```
points(cohplotdata$nhd_long, cohplotdata$nhd_lat, pch=16, cex=1, col=pal[round(cohplotdata$accndvicoh.ts
colorbar.plot(x=mean(par("usr")[1:2]),y=par("usr")[3],strip=1:100,col=pal,horizontal = T)
```

Lakes by short timescale coherence



```
plot(lagosstates, main="Lakes by long timescale coherence")
points(cohplotdata$nhd_long, cohplotdata$nhd_lat, pch=16, cex=1, col=pal[round(cohplotdata$accndvicoh.ts
colorbar.plot(x=mean(par("usr")[1:2]),y=par("usr")[3],strip=1:100,col=pal,horizontal = T)
```

Lakes by long timescale coherence



```
laymat=matrix(1,nrow=2,ncol=13)
laymat[2,]<-2
laymat[,13]<-3

tiff("~/Box Sync/NSF EAGER Synchrony/Manuscripts/1_CoherenceSpatialVariation/fig3_coherencemap.tif", uni
     res=300, width=6.5, height=7.5)

layout(laymat)
par(mar=c(0,0,1.5,0))

plot(lagosstates, main="Short timescale coherence")
points(cohplotdata$nhd_long, cohplotdata$nhd_lat, pch=16, cex=1.5, col=pal[round(cohplotdata$accndvicoh
text(0.99*par()$usr[1],0.99*par()$usr[4],"a)",cex=1.5)

plot(lagosstates, main="Long timescale coherence")
points(cohplotdata$nhd_long, cohplotdata$nhd_lat, pch=16, cex=1.5, col=pal[round(cohplotdata$accndvicoh
text(0.99*par()$usr[1],0.99*par()$usr[4],"b)",cex=1.5)

par(mar=c(5,1,5,1))
image(matrix(1:100,nrow=1),col=pal,xaxt="n",yaxt="n")
axis(2,at=seq(0,1,0.2))

dev.off()
```

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

```
dt<-lagosne_load("1.087.3")
```

```r
dt.conn<-dt$buffer500m.conn
dt.conn<-dt.conn[,!grepl("sum_lengthm",colnames(dt.conn))]
dt.conn<-dt.conn[,colnames(dt.conn)!="buffer500m_nhdid"]


dt.chag<-dt$hu12.chag
dt.chag<-dt.chag[,!grepl("_min",colnames(dt.chag))]
dt.chag<-dt.chag[,!grepl("_max",colnames(dt.chag))]
dt.chag<-dt.chag[,!grepl("_ha",colnames(dt.chag))]
dt.chag<-dt.chag[,!colnames(dt.chag)=="borderhu12s"]
dt.chag$hu12_dep_no3_tavg_mean<-rowMeans(dt.chag[,grepl("hu12_dep_no3",colnames(dt.chag)) &
                                                   grepl("_mean",colnames(dt.chag))])
dt.chag$hu12_dep_no3_tavg_std<-rowMeans(dt.chag[,grepl("hu12_dep_no3",colnames(dt.chag)) &
                                                   grepl("_std",colnames(dt.chag))])
dt.chag$hu12_dep_so4_tavg_mean<-rowMeans(dt.chag[,grepl("hu12_dep_so4",colnames(dt.chag)) &
                                                    grepl("_mean", colnames(dt.chag))])
dt.chag$hu12_dep_so4_tavg_std<-rowMeans(dt.chag[,grepl("hu12_dep_so4",colnames(dt.chag)) &
                                                    grepl("_std", colnames(dt.chag))])
dt.chag$hu12_dep_totaln_tavg_mean<-rowMeans(dt.chag[,grepl("hu12_dep_totaln",colnames(dt.chag)) &
                                                      grepl("_mean", colnames(dt.chag))])
dt.chag$hu12_dep_totaln_tavg_std<-rowMeans(dt.chag[,grepl("hu12_dep_totaln",colnames(dt.chag)) &
                                                      grepl("_std", colnames(dt.chag))])
dt.chag<-dt.chag[,!(grepl("hu12_dep",colnames(dt.chag)) & grepl("_19",colnames(dt.chag)))]
dt.chag<-dt.chag[,!(grepl("hu12_dep",colnames(dt.chag)) & grepl("_20",colnames(dt.chag)))]
dt.chag<-dt.chag[,!(grepl("_std",colnames(dt.chag)))]
dt.chag<-dt.chag[,!grepl("surficialgeology",colnames(dt.chag))]


dt.geo<-dt$lakes.geo
dt.geo<-dt$lakes.geo[,!colnames(dt.geo) %in% c("state_zoneid","iws_zoneid","edu_zoneid","county_zoneid"
dt.geo<-dt.geo[,!grepl("_count",colnames(dt.geo))]

dt.lulc<-dt$hu12.lulc
dt.lulc<-dt.lulc[,!grepl("_ha_",colnames(dt.lulc))]
dt.lulc<-dt.lulc[,!grepl("_nlcd1992_",colnames(dt.lulc))]
dt.lulc<-dt.lulc[,!grepl("_nlcd2006_",colnames(dt.lulc))]
dt.lulc<-dt.lulc[,!grepl("_nlcd2001_",colnames(dt.lulc))]
dt.lulc<-dt.lulc[,colnames(dt.lulc)!="hu12_damdensity_pointsperha"]
dt.lulc<-dt.lulc[,colnames(dt.lulc)!="hu12_damdensity_pointcount"]
dt.lulc<-dt.lulc[,colnames(dt.lulc)!="hu12_roaddensity_sum_lengthm"]
dt.lulc<-dt.lulc[,!grepl("_min",colnames(dt.lulc))]
dt.lulc<-dt.lulc[,!grepl("_max",colnames(dt.lulc))]
dt.lulc<-dt.lulc[,!grepl("_std",colnames(dt.lulc))]

#depth
depth<-lagosne_select(table="lakes_limno", vars=c("lagoslakeid","maxdepth"))
depth<-depth[depth$lagoslakeid %in% analysislakes$lakeinfo$lagoslakeid,] #use max depth because it's mo

#growing season Chlorophyll-a
chla<-lagosne_select(table="epi_nutr", vars=c("lagoslakeid","samplemonth","chla"))
chla<-chla[chla$lagoslakeid %in% analysislakes$lakeinfo$lagoslakeid,]
gs.chla<-chla[chla$samplemonth %in% 5:9,]
avg.chla<-aggregate(chla ~ lagoslakeid, data=gs.chla, FUN=mean, na.rm=T)
```

```r
#Chlorophyll-a TSI class
#TSI(CHL) = 9.81 ln(CHL) + 30.6
tsi.chl<-data.frame(lagoslakeid=avg.chla$lagoslakeid, tsi=9.81 * log(avg.chla$chla) + 30.6)
tsi.chl$tsi.cat<-rep("lake",nrow(tsi.chl))

tsi.chl$tsi.cat[tsi.chl$tsi < 40]<-"oligotrophic"
tsi.chl$tsi.cat[tsi.chl$tsi >=40 & tsi.chl$tsi < 50]<-"mesotrophic"
tsi.chl$tsi.cat[tsi.chl$tsi >=50 & tsi.chl$tsi < 70]<-"eutrophic"
tsi.chl$tsi.cat[tsi.chl$tsi >= 70] <-"hypereutrophic"

#CV of terrestrial NDVI
cv.accndvi<-NULL
for(lake in 1:length(analysislakes$lakedata)){
  tmp<-analysislakes$lakedata[[lake]][rownames(analysislakes$lakedata[[lake]])=="avhrrdata",]

  cv.accndvi<-c(cv.accndvi, sd(tmp)/mean(tmp))
 # rm(tmp)
}
cv.accndvi<-data.frame(lagoslakeid=as.numeric(names(analysislakes$lakedata)), cv.accndvi=cv.accndvi)

#shoreline development ratio
sdev<-analysislakes$lakeinfo$lake_perim_meters/(2*sqrt(pi*analysislakes$lakeinfo$lake_area_ha*10000))
shoredev<-data.frame(lagoslakeid=analysislakes$lakeinfo$lagoslakeid,shoredev=sdev)

preds<-analysislakes$lakeinfo[,colnames(analysislakes$lakeinfo) %in% c("lagoslakeid","end","start")]
preds$tslength<-preds$end-preds$start + 1
preds<-left_join(preds, dt.geo, by="lagoslakeid")
preds<-left_join(preds, dt.conn, by="lagoslakeid")
preds<-left_join(preds, dt.chag, by="hu12_zoneid")
preds<-left_join(preds, dt.lulc, by="hu12_zoneid")
preds<-left_join(preds, avg.chla, by="lagoslakeid")
preds<-left_join(preds, cv.accndvi, by="lagoslakeid")
preds<-left_join(preds, depth, by="lagoslakeid")
preds<-left_join(preds, shoredev, by="lagoslakeid")

#remove any variables that don't vary in this dataset
# for(nn in 1:ncol(preds)){
#   print(colnames(preds)[nn])
#   print(unique(preds[,nn]))
# }

preds<-preds[,!colnames(preds) %in% c("hu12_surficialgeology_beach_pct",
                                      "hu12_surficialgeology_colluv_pct",
                                      "hu12_surficialgeology_grus_pct",
                                      "hu12_surficialgeology_other_pct",
                                      "hu12_surficialgeology_solif_pct",
                                      "hu12_surficialgeology_till_oth_pct",
                                      "hu12_nlcd2011_pct_0")]

# modvars.conn<-left_join(pred.conn, coh.chlaXaccndvi, by="lagoslakeid")
# modvars.chag<-left_join(pred.chag, coh.chlaXaccndvi, by="lagoslakeid")
```

```r
#huc2 and huc4 watershed codes
huc_codes<-read.csv("~/GitHub/AquaTerrSynch/AnalysisCode/match_huc_codes.csv", colClasses = 'character'

#state info
states<-lagosne_select(table="state", vars=c("state_zoneid","state_name"))

for(nn in 1:ncol(preds)){

  if(is.factor(preds[,nn])){
    preds[,nn]<-factor(preds[,nn])
  }

}
```

```r
rfdat.cohst<-left_join(coh.chlaXaccndvi[,c(10,3)], preds)
```

```
## Joining, by = "lagoslakeid"
```

```r
rfdat.cohst<-rfdat.cohst[,!colnames(rfdat.cohst) %in% c("lagoslakeid","start","end","lakes_nhdid","hu12_
rfdat.cohst<-rfdat.cohst[,!grepl("borderhu12s",colnames(rfdat.cohst))]

for(nn in 1:ncol(rfdat.cohst)){
  if(is.character(rfdat.cohst[,nn])){
    rfdat.cohst[,nn]<-as.factor(rfdat.cohst[,nn])
  }
}

cf.cohst<-party::cforest(accndvicoh.ts1 ~ ., data=rfdat.cohst, controls=cforest_control(ntree=50000,minc

varimp.coh.st<-varimp(cf.cohst)
print(varimp.coh.st[order(varimp.coh.st, decreasing=T)][1:10])
```

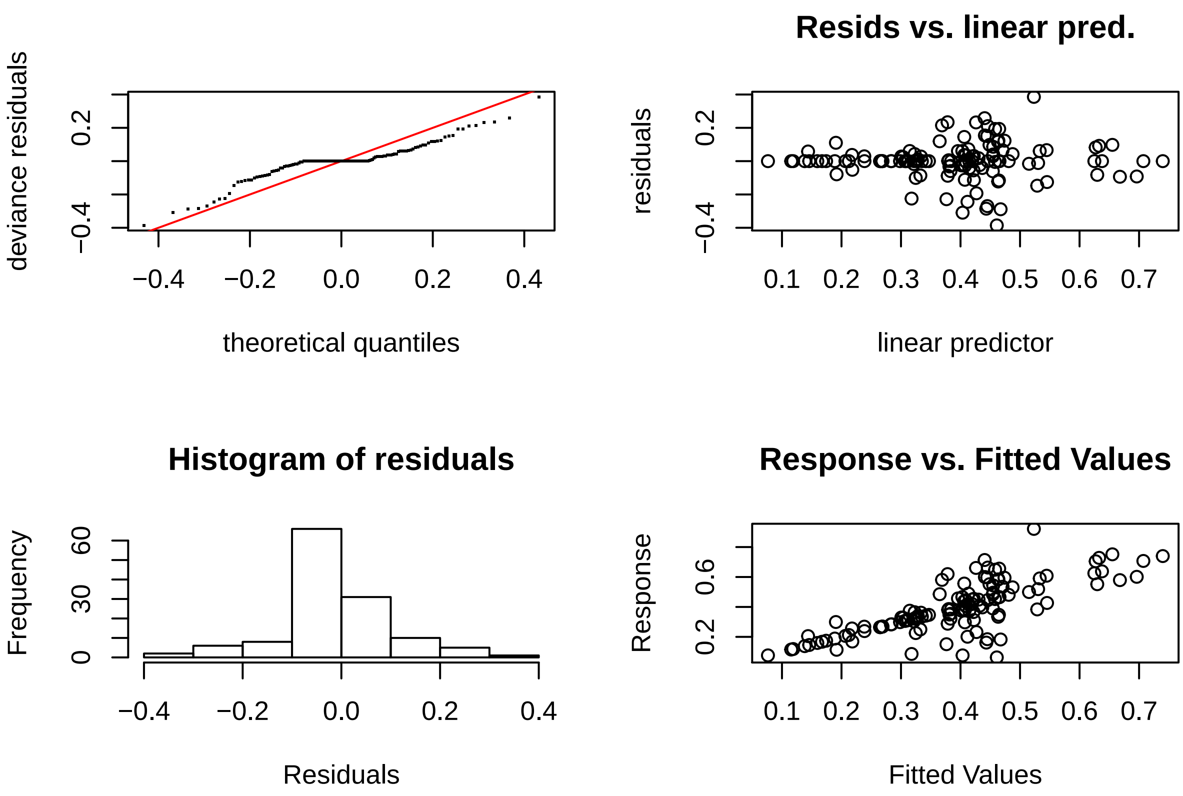
```
##                                    cv.accndvi
##                                  0.0005710355
##                                    hu8_zoneid
##                                  0.0005568739
##                      hu12_dep_totaln_tavg_mean
##                                  0.0004949402
##                           hu12_nlcd2011_pct_90
##                                  0.0001939783
##                           hu12_nlcd2011_pct_95
##                                  0.0001609831
## wlconnections_scrubshrubwetlands_shoreline_km
##                                  0.0001473959
##        wlconnections_allwetlands_shoreline_km
##                                  0.0001400305
##                                hu12_slope_mean
##                                  0.0001274735
##                                  hu12_tri_mean
##                                  0.0001274102
##   wlconnections_forestedwetlands_shoreline_km
##                                  0.0000901705
```

```r
#hist(predcoh.st)
#hist(modvars.accndvi$accndvicoh.ts1)
```

```
predcoh.st<-predict(cf.cohst, newdata=rfdat.cohst,type="response")
# plot(predcoh.st, rfdat.cohst$accndvicoh.ts1, xlab="predicted", ylab="empirical", main="Coherence, sho
#       xlim=c(0,1), ylim=c(0,1))
# abline(a=0,b=1)
cor.test(predcoh.st,rfdat.cohst$accndvicoh.ts1)
```

```
##
##  Pearson's product-moment correlation
##
## data:  predcoh.st and rfdat.cohst$accndvicoh.ts1
## t = 17.438, df = 129, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.7782621 0.8826028
## sample estimates:
##       cor
## 0.8379308
```

```
lwgt<-preds$tslength/mean(preds$tslength)

gam.cohst<-gam(accndvicoh.ts1 ~ hu8_zoneid + s(cv.accndvi) + s(hu12_dep_totaln_tavg_mean) +
               s(hu12_nlcd2011_pct_90) + s(wlconnections_scrubshrubwetlands_shoreline_km), data=rfdat.c
gam.check(gam.cohst)
```




```
##
## Method: GCV   Optimizer: magic
## Smoothing parameter selection converged after 15 iterations.
## The RMS GCV score gradient at convergence was 4.447258e-08 .
## The Hessian was positive definite.
```

```
## Model rank =  101 / 101
##
## Basis dimension (k) checking results. Low p-value (k-index<1) may
## indicate that k is too low, especially if edf is close to k'.
##
##                                                     k'  edf k-index p-value
## s(cv.accndvi)                                     9.00 1.00    0.97    0.32
## s(hu12_dep_totaln_tavg_mean)                      9.00 1.00    1.09    0.79
## s(hu12_nlcd2011_pct_90)                           9.00 1.00    1.16    0.96
## s(wlconnections_scrubshrubwetlands_shoreline_km) 9.00 1.45    0.90    0.14
```

```r
concurvity(gam.cohst)
```

```
##          para s(cv.accndvi) s(hu12_dep_totaln_tavg_mean)
## worst       1     0.9780469                    0.9995309
## observed    1     0.7757388                    0.9743131
## estimate    1     0.6615465                    0.9589468
##          s(hu12_nlcd2011_pct_90)
## worst                  0.9936800
## observed               0.9675487
## estimate               0.9220433
##          s(wlconnections_scrubshrubwetlands_shoreline_km)
## worst                                           1.0000000
## observed                                        0.9488984
## estimate                                        0.9741744
```

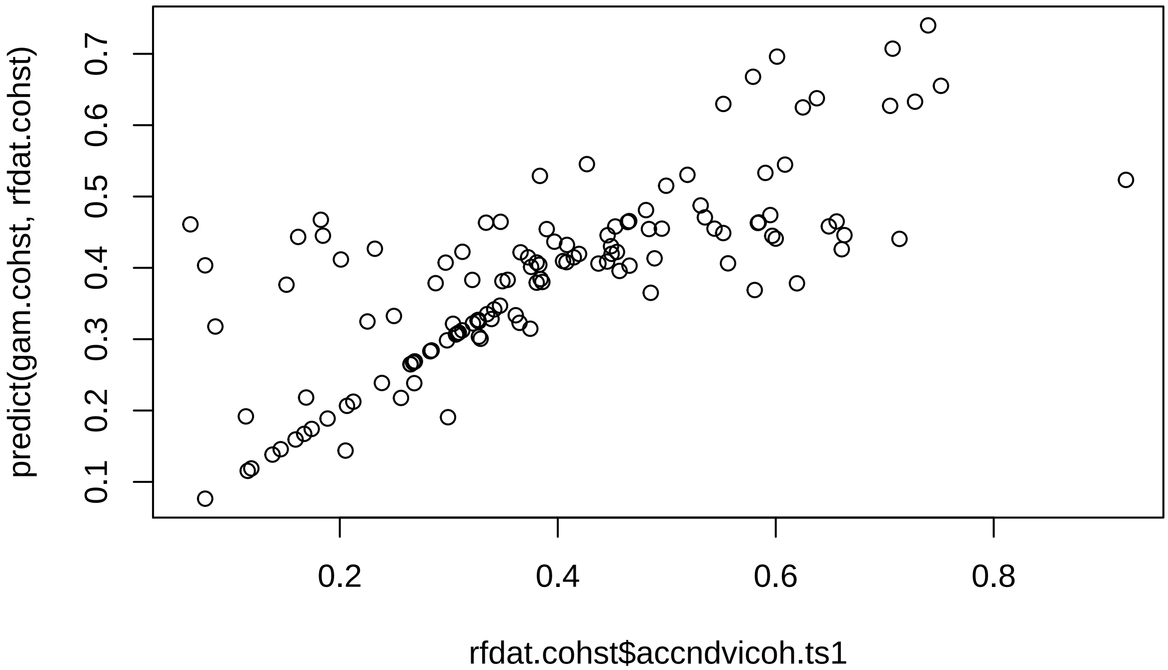
```r
summary(gam.cohst)
```

```
##
## Family: gaussian
## Link function: identity
##
## Formula:
## accndvicoh.ts1 ~ hu8_zoneid + s(cv.accndvi) + s(hu12_dep_totaln_tavg_mean) +
##     s(hu12_nlcd2011_pct_90) + s(wlconnections_scrubshrubwetlands_shoreline_km)
##
## Parametric coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        0.57612    0.18972   3.037  0.00355 **
## hu8_zoneidHU8_13  -0.31582    0.22944  -1.376  0.17383
## hu8_zoneidHU8_141 -0.26463    0.21346  -1.240  0.21994
## hu8_zoneidHU8_142 -0.32628    0.28855  -1.131  0.26269
## hu8_zoneidHU8_143 -0.26387    0.18153  -1.454  0.15131
## hu8_zoneidHU8_160 -0.39395    0.23444  -1.680  0.09812 .
## hu8_zoneidHU8_161 -0.28358    0.17450  -1.625  0.10942
## hu8_zoneidHU8_163 -0.34553    0.19179  -1.802  0.07666 .
## hu8_zoneidHU8_164 -0.22037    0.27533  -0.800  0.42667
## hu8_zoneidHU8_165 -0.19884    0.24103  -0.825  0.41267
## hu8_zoneidHU8_167  0.07213    0.23228   0.311  0.75723
## hu8_zoneidHU8_168 -0.36755    0.23811  -1.544  0.12799
## hu8_zoneidHU8_169  0.05673    0.26060   0.218  0.82843
## hu8_zoneidHU8_170 -0.08418    0.21860  -0.385  0.70155
## hu8_zoneidHU8_173 -0.34463    0.26824  -1.285  0.20384
## hu8_zoneidHU8_177 -0.28252    0.30374  -0.930  0.35606
## hu8_zoneidHU8_181 -0.29777    0.30503  -0.976  0.33292
```

```
## hu8_zoneidHU8_186 -0.09077    0.29792  -0.305  0.76167
## hu8_zoneidHU8_204 -0.11262    0.21556  -0.522  0.60330
## hu8_zoneidHU8_23  -0.21810    0.24453  -0.892  0.37602
## hu8_zoneidHU8_24  -0.32460    0.25769  -1.260  0.21272
## hu8_zoneidHU8_256 -0.08987    0.22557  -0.398  0.69177
## hu8_zoneidHU8_325 -0.32064    0.21407  -1.498  0.13947
## hu8_zoneidHU8_331 -0.42798    0.28454  -1.504  0.13783
## hu8_zoneidHU8_335 -0.23998    0.21111  -1.137  0.26019
## hu8_zoneidHU8_343 -0.05678    0.24185  -0.235  0.81521
## hu8_zoneidHU8_346 -0.46909    0.25742  -1.822  0.07344 .
## hu8_zoneidHU8_350 -0.07386    0.27018  -0.273  0.78551
## hu8_zoneidHU8_370 -0.05145    0.26575  -0.194  0.84713
## hu8_zoneidHU8_382 -0.46263    0.25876  -1.788  0.07888 .
## hu8_zoneidHU8_39  -0.31752    0.27108  -1.171  0.24613
## hu8_zoneidHU8_40  -0.32863    0.25241  -1.302  0.19795
## hu8_zoneidHU8_428 -0.29904    0.25225  -1.185  0.24054
## hu8_zoneidHU8_429  0.11952    0.26411   0.453  0.65253
## hu8_zoneidHU8_431 -0.13000    0.28765  -0.452  0.65296
## hu8_zoneidHU8_434 -0.17999    0.27648  -0.651  0.51754
## hu8_zoneidHU8_435 -0.11599    0.26584  -0.436  0.66420
## hu8_zoneidHU8_436 -0.16016    0.29807  -0.537  0.59304
## hu8_zoneidHU8_437 -0.12289    0.29031  -0.423  0.67361
## hu8_zoneidHU8_440 -0.25422    0.29411  -0.864  0.39086
## hu8_zoneidHU8_449 -0.22700    0.25417  -0.893  0.37539
## hu8_zoneidHU8_452 -0.21039    0.27126  -0.776  0.44104
## hu8_zoneidHU8_453 -0.01747    0.22225  -0.079  0.93762
## hu8_zoneidHU8_456 -0.27408    0.30541  -0.897  0.37311
## hu8_zoneidHU8_481 -0.12556    0.28008  -0.448  0.65557
## hu8_zoneidHU8_484 -0.15192    0.28386  -0.535  0.59451
## hu8_zoneidHU8_49  -0.18768    0.21553  -0.871  0.38736
## hu8_zoneidHU8_491 -0.09295    0.34902  -0.266  0.79091
## hu8_zoneidHU8_492  0.10630    0.23600   0.450  0.65403
## hu8_zoneidHU8_493 -0.22129    0.26424  -0.837  0.40568
## hu8_zoneidHU8_494  0.16612    0.24442   0.680  0.49936
## hu8_zoneidHU8_495 -0.10559    0.28001  -0.377  0.70745
## hu8_zoneidHU8_501  0.24662    0.28517   0.865  0.39062
## hu8_zoneidHU8_507 -0.27632    0.25440  -1.086  0.28178
## hu8_zoneidHU8_51  -0.02442    0.29340  -0.083  0.93394
## hu8_zoneidHU8_59  -0.21609    0.21188  -1.020  0.31191
## hu8_zoneidHU8_61  -0.22569    0.24927  -0.905  0.36891
## hu8_zoneidHU8_73  -0.17293    0.25320  -0.683  0.49728
## hu8_zoneidHU8_74  -0.20872    0.26599  -0.785  0.43574
## hu8_zoneidHU8_75  -0.08026    0.27814  -0.289  0.77393
## hu8_zoneidHU8_76  -0.12414    0.34001  -0.365  0.71633
## hu8_zoneidHU8_81  -0.24739    0.25407  -0.974  0.33414
## hu8_zoneidHU8_83  -0.29364    0.29416  -0.998  0.32221
## hu8_zoneidHU8_84  -0.18936    0.27444  -0.690  0.49290
## hu8_zoneidHU8_96   0.07335    0.26107   0.281  0.77971
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##                                                edf Ref.df     F
## s(cv.accndvi)                                1.000  1.000 0.410
```

```
## s(hu12_dep_totaln_tavg_mean)                   1.000  1.000 1.505
## s(hu12_nlcd2011_pct_90)                        1.000  1.000 0.646
## s(wlconnections_scrubshrubwetlands_shoreline_km) 1.447  1.702 0.525
##                                                p-value
## s(cv.accndvi)                                    0.525
## s(hu12_dep_totaln_tavg_mean)                     0.225
## s(hu12_nlcd2011_pct_90)                          0.425
## s(wlconnections_scrubshrubwetlands_shoreline_km)   0.456
##
## R-sq.(adj) =  0.0965   Deviance explained =   58%
## GCV = 0.056948  Scale est. = 0.02629   n = 129
```

```
plot(rfdat.cohst$accndvicoh.ts1, predict(gam.cohst, rfdat.cohst))
```




```
rfdat.cohlt<-left_join(coh.chlaXaccndvi[,c(10,6)], preds)
```

```
## Joining, by = "lagoslakeid"
```

```
rfdat.cohlt<-rfdat.cohlt[,!colnames(rfdat.cohlt) %in% c("lagoslakeid","start","end","lakes_nhdid","hu12_
rfdat.cohlt<-rfdat.cohlt[,!grepl("borderhu12s",colnames(rfdat.cohlt))]

for(nn in 1:ncol(rfdat.cohlt)){
  if(is.character(rfdat.cohlt[,nn])){
    rfdat.cohlt[,nn]<-as.factor(rfdat.cohlt[,nn])
  }
}

cf.cohlt<-party::cforest(accndvicoh.ts2 ~ ., data=rfdat.cohlt, controls=cforest_control(ntree=50000,min

varimp.coh.lt<-varimp(cf.cohlt)
print(varimp.coh.lt[order(varimp.coh.lt, decreasing=T)][1:10])
```

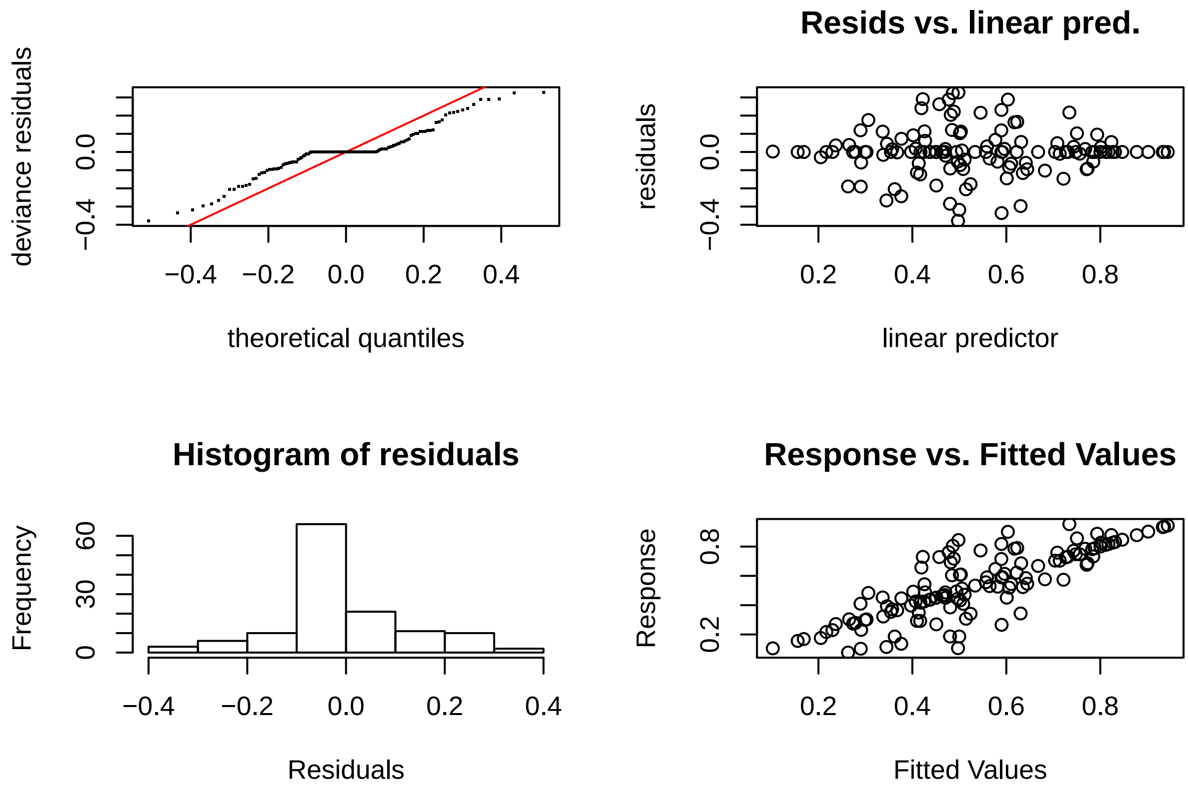
```
##                                    hu8_zoneid
##                                  1.203560e-03
```

```
##                            hu12_nlcd2011_pct_52
##                                    9.442723e-04
##                                      cv.accndvi
##                                    4.297765e-04
##    wlconnections_forestedwetlands_shoreline_km
##                                    1.483532e-04
##  hu12_prism_ppt_30yr_normal_800mm2_annual_mean
##                                    1.395582e-04
## wlconnections_allwetlands_contributing_area_ha
##                                    1.272004e-04
##                            hu12_nlcd2011_pct_95
##                                    1.247903e-04
##                          hu12_dep_so4_tavg_mean
##                                    8.831864e-05
##         wlconnections_allwetlands_shoreline_km
##                                    8.766973e-05
##                  hu12_groundwaterrecharge_mean
##                                    5.901385e-05
```

```r
#hist(predcoh.st)
#hist(modvars.accndvi$accndvicoh.ts1)
predcoh.lt<-predict(cf.cohlt, newdata=rfdat.cohlt,type="response")
# plot(predcoh.lt, rfdat.cohlt$accndvicoh.ts2, xlab="predicted", ylab="empirical", main="Coherence, lon
#      xlim=c(0,1), ylim=c(0,1))
# abline(a=0,b=1)
cor.test(predcoh.lt,rfdat.cohlt$accndvicoh.ts2)
```

```
##
##  Pearson's product-moment correlation
##
## data:  predcoh.lt and rfdat.cohlt$accndvicoh.ts2
## t = 16.476, df = 129, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.7589904 0.8717455
## sample estimates:
##       cor
## 0.8233297
```

```r
lwgt<-preds$tslength/mean(preds$tslength)

gam.cohlt<-gam(accndvicoh.ts2 ~ hu8_zoneid + s(hu12_nlcd2011_pct_52) + s(cv.accndvi) +
                 s(wlconnections_allwetlands_contributing_area_ha) +
                 s(wlconnections_forestedwetlands_shoreline_km), data=rfdat.cohlt, gamma=1, weights=lwg
gam.check(gam.cohlt)
```




```
##
## Method: GCV   Optimizer: magic
## Smoothing parameter selection converged after 14 iterations.
## The RMS GCV score gradient at convergence was 3.147617e-08 .
## The Hessian was positive definite.
## Model rank =  101 / 101
##
## Basis dimension (k) checking results. Low p-value (k-index<1) may
## indicate that k is too low, especially if edf is close to k'.
##
##                                                k'  edf k-index
## s(hu12_nlcd2011_pct_52)                      9.00 1.00    0.98
## s(cv.accndvi)                                9.00 1.00    0.97
## s(wlconnections_allwetlands_contributing_area_ha) 9.00 2.94    0.88
## s(wlconnections_forestedwetlands_shoreline_km)    9.00 1.00    0.87
##                                              p-value
## s(hu12_nlcd2011_pct_52)                        0.330
## s(cv.accndvi)                                  0.315
## s(wlconnections_allwetlands_contributing_area_ha)  0.115
## s(wlconnections_forestedwetlands_shoreline_km)     0.045 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
concurvity(gam.cohlt)
```

```
##          para s(hu12_nlcd2011_pct_52) s(cv.accndvi)
## worst       1               0.9965257     0.9732582
## observed    1               0.8294950     0.7268313
## estimate    1               0.7960808     0.6203958
##          s(wlconnections_allwetlands_contributing_area_ha)
```

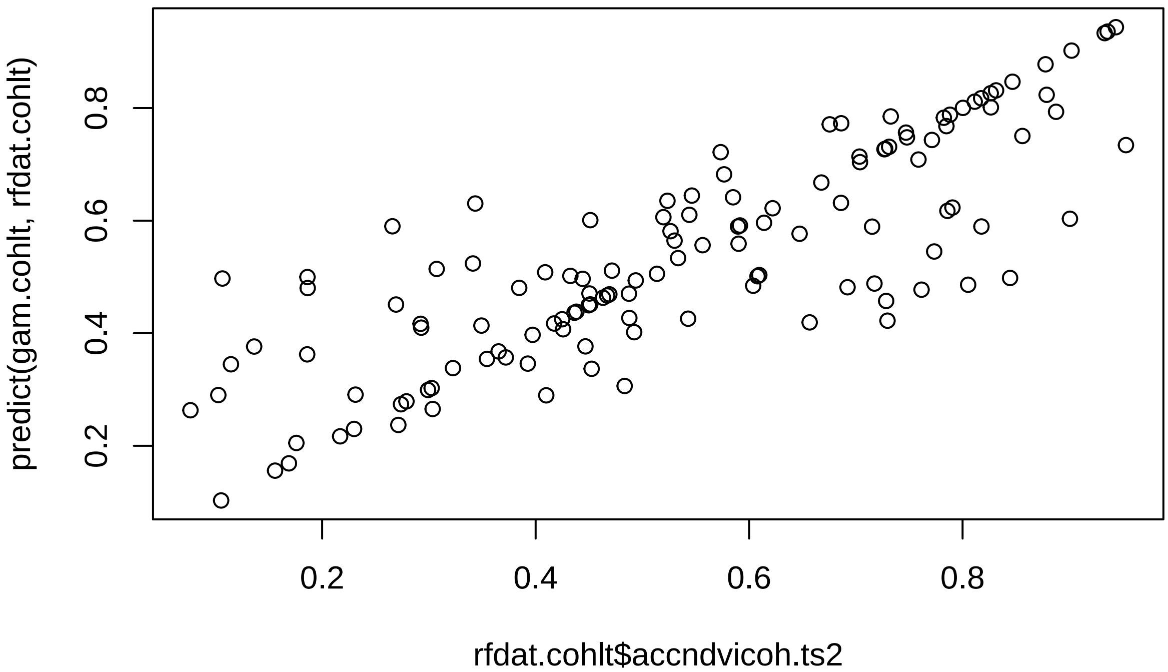
```
## worst                                                 1.0000000
## observed                                              0.9993585
## estimate                                              0.9973873
##          s(wlconnections_forestedwetlands_shoreline_km)
## worst                                              1.0000000
## observed                                           0.9991374
## estimate                                           0.9989994
```

```
summary(gam.cohlt)
```

```
##
## Family: gaussian
## Link function: identity
##
## Formula:
## accndvicoh.ts2 ~ hu8_zoneid + s(hu12_nlcd2011_pct_52) + s(cv.accndvi) +
##     s(wlconnections_allwetlands_contributing_area_ha) + s(wlconnections_forestedwetlands_shoreline_km
##
## Parametric coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.459483   0.199384   2.305  0.02479 *
## hu8_zoneidHU8_13    0.154205   0.216467   0.712  0.47909
## hu8_zoneidHU8_141 -0.108910   0.221421  -0.492  0.62467
## hu8_zoneidHU8_142  0.263643   0.279154   0.944  0.34886
## hu8_zoneidHU8_143 -0.090400   0.221025  -0.409  0.68404
## hu8_zoneidHU8_160  0.235402   0.279808   0.841  0.40363
## hu8_zoneidHU8_161 -0.040268   0.204648  -0.197  0.84470
## hu8_zoneidHU8_163  0.029075   0.224731   0.129  0.89751
## hu8_zoneidHU8_164 -0.288622   0.278342  -1.037  0.30407
## hu8_zoneidHU8_165  0.065517   0.302669   0.216  0.82939
## hu8_zoneidHU8_167 -0.165512   0.238266  -0.695  0.49004
## hu8_zoneidHU8_168 -0.061557   0.274895  -0.224  0.82360
## hu8_zoneidHU8_169  0.127621   0.239817   0.532  0.59665
## hu8_zoneidHU8_170  0.019899   0.213293   0.093  0.92599
## hu8_zoneidHU8_173  0.243513   0.277382   0.878  0.38362
## hu8_zoneidHU8_177 -0.256173   0.281080  -0.911  0.36586
## hu8_zoneidHU8_181 -0.239423   0.308515  -0.776  0.44087
## hu8_zoneidHU8_186 -0.318734   0.277026  -1.151  0.25463
## hu8_zoneidHU8_204  0.193240   0.251171   0.769  0.44480
## hu8_zoneidHU8_23   0.206390   0.282204   0.731  0.46751
## hu8_zoneidHU8_24   0.300875   0.280716   1.072  0.28824
## hu8_zoneidHU8_256 -0.488772   0.309339  -1.580  0.11953
## hu8_zoneidHU8_325 -0.122465   0.208448  -0.588  0.55914
## hu8_zoneidHU8_331  0.123107   0.293160   0.420  0.67609
## hu8_zoneidHU8_335 -0.028978   0.208582  -0.139  0.88999
## hu8_zoneidHU8_343 -0.281800   0.207320  -1.359  0.17932
## hu8_zoneidHU8_346 -0.148742   0.272133  -0.547  0.58676
## hu8_zoneidHU8_350  0.023331   0.271195   0.086  0.93174
## hu8_zoneidHU8_370 -0.686644   0.471030  -1.458  0.15030
## hu8_zoneidHU8_382  0.223720   0.279294   0.801  0.42639
## hu8_zoneidHU8_39   0.248929   0.273765   0.909  0.36696
## hu8_zoneidHU8_40  -0.001363   0.273519  -0.005  0.99604
## hu8_zoneidHU8_428 -0.008345   0.235863  -0.035  0.97190
## hu8_zoneidHU8_429 -0.007849   0.266903  -0.029  0.97664
## hu8_zoneidHU8_431  0.368631   0.281212   1.311  0.19506
```

```
## hu8_zoneidHU8_434 -0.163704   0.276096  -0.593  0.55553
## hu8_zoneidHU8_435 -0.470122   0.266090  -1.767  0.08252 .
## hu8_zoneidHU8_436  0.249541   0.273622   0.912  0.36554
## hu8_zoneidHU8_437 -0.058549   0.277020  -0.211  0.83335
## hu8_zoneidHU8_440  0.274181   0.282282   0.971  0.33543
## hu8_zoneidHU8_449 -0.347526   0.326131  -1.066  0.29101
## hu8_zoneidHU8_452 -0.315904   0.266983  -1.183  0.24154
## hu8_zoneidHU8_453 -0.101221   0.221395  -0.457  0.64924
## hu8_zoneidHU8_456  5.599102   1.740011   3.218  0.00211 **
## hu8_zoneidHU8_481 -0.294839   0.267284  -1.103  0.27454
## hu8_zoneidHU8_484  1.780125   0.552400   3.223  0.00209 **
## hu8_zoneidHU8_49   0.044751   0.245059   0.183  0.85574
## hu8_zoneidHU8_491  2.414507   1.052197   2.295  0.02538 *
## hu8_zoneidHU8_492  0.013972   0.241681   0.058  0.95410
## hu8_zoneidHU8_493  0.151808   0.317070   0.479  0.63389
## hu8_zoneidHU8_494  0.169272   0.234965   0.720  0.47416
## hu8_zoneidHU8_495  4.843348   1.513325   3.200  0.00223 **
## hu8_zoneidHU8_501 -0.211336   0.270369  -0.782  0.43759
## hu8_zoneidHU8_507  0.129229   0.267792   0.483  0.63121
## hu8_zoneidHU8_51  -0.224977   0.273152  -0.824  0.41352
## hu8_zoneidHU8_59  -0.109449   0.228821  -0.478  0.63422
## hu8_zoneidHU8_61   0.173809   0.223491   0.778  0.43990
## hu8_zoneidHU8_73   0.016568   0.277245   0.060  0.95255
## hu8_zoneidHU8_74  -0.204799   0.270273  -0.758  0.45166
## hu8_zoneidHU8_75   0.138522   0.258448   0.536  0.59402
## hu8_zoneidHU8_76   0.131873   0.310906   0.424  0.67302
## hu8_zoneidHU8_81  -0.331164   0.291764  -1.135  0.26102
## hu8_zoneidHU8_83  -0.326157   0.273137  -1.194  0.23729
## hu8_zoneidHU8_84  -0.138984   0.270699  -0.513  0.60960
## hu8_zoneidHU8_96   0.255229   0.275005   0.928  0.35720
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##                                                   edf Ref.df      F
## s(hu12_nlcd2011_pct_52)                          1.00  1.000  3.040
## s(cv.accndvi)                                    1.00  1.000  1.386
## s(wlconnections_allwetlands_contributing_area_ha) 2.94  3.421  4.536
## s(wlconnections_forestedwetlands_shoreline_km)    1.00  1.000 10.106
##                                                  p-value
## s(hu12_nlcd2011_pct_52)                          0.08643 .
## s(cv.accndvi)                                    0.24391
## s(wlconnections_allwetlands_contributing_area_ha) 0.00763 **
## s(wlconnections_forestedwetlands_shoreline_km)    0.00234 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.305   Deviance explained = 68.5%
## GCV = 0.081122  Scale est. = 0.036511  n = 129
```

```
plot(rfdat.cohlt$accndvicoh.ts2, predict(gam.cohlt, rfdat.cohlt))
```




```
rfdat.phist<-left_join(coh.chlaXaccndvi[,c(10,2)], preds)
```

```
## Joining, by = "lagoslakeid"
```

```
rfdat.phist<-rfdat.phist[,!colnames(rfdat.phist) %in%
                          c("lagoslakeid","start","end","lakes_nhdid","hu12_zoneid","tslength","county_
rfdat.phist<-rfdat.phist[,!grepl("borderhu12s",colnames(rfdat.phist))]

rfdat.phist<-rfdat.phist[coh.chlaXaccndvi$accndvip.ts1<0.3,]

for(nn in 1:ncol(rfdat.phist)){
  if(is.character(rfdat.phist[,nn])){
    rfdat.phist[,nn]<-as.factor(rfdat.phist[,nn])
  }
}

cf.phist<-party::cforest(cos(accndviphi.ts1) ~ ., data=rfdat.phist,
                         controls=cforest_control(ntree=50000,mincriterion = 0.9,mtry=3))

varimp.phi.st<-varimp(cf.phist)
print(varimp.phi.st[order(varimp.phi.st, decreasing=T)][1:10])
```

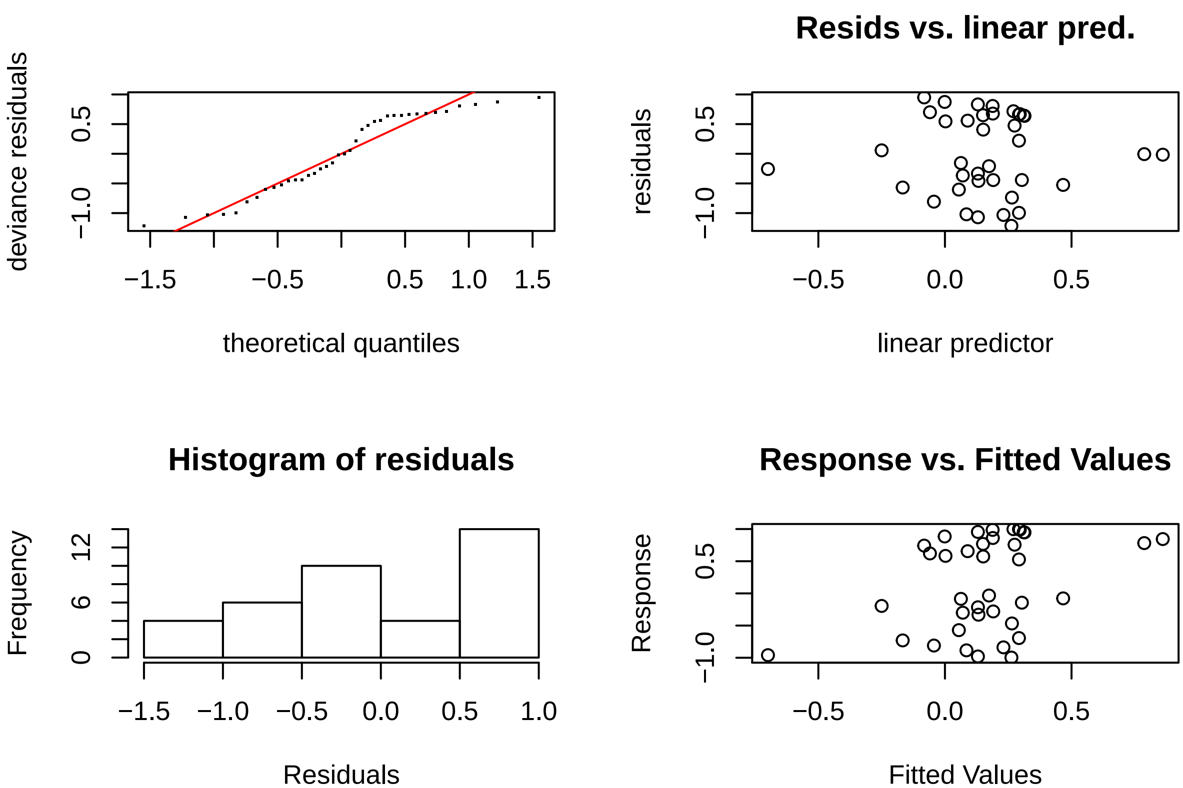
```
##       wlconnections_openwaterwetlands_shoreline_km
##                                       0.0033487759
## buffer500m_streamdensity_headwaters_density_mperha
##                                       0.0032202533
##                                               chla
##                                       0.0020374995
##                               hu12_nlcd2011_pct_21
##                                       0.0018605382
## wlconnections_openwaterwetlands_contributing_area_
##                                       0.0018041944
##    buffer500m_streamdensity_streams_density_mperha
```

```
##                                      0.0016664426
##                             hu12_dep_no3_tavg_mean
##                                      0.0003453840
##                               hu12_nlcd2011_pct_22
##                                      0.0003389173
##                    hu12_roaddensity_density_mperha
##                                      0.0003335029
##     wlconnections_allwetlands_contributing_area_ha
##                                      0.0001936666
```

```
predphi.st<-predict(cf.phist, newdata=rfdat.phist,type="response")
cor.test(predphi.st,cos(rfdat.phist$accndviphi.ts1))
```

```
##
##  Pearson's product-moment correlation
##
## data:  predphi.st and cos(rfdat.phist$accndviphi.ts1)
## t = 10.234, df = 37, p-value = 2.431e-12
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.7466154 0.9244186
## sample estimates:
##       cor
## 0.8596296
```

```
lwgt<-preds$tslength[coh.chlaXaccndvi$accndvip.ts1<0.3]/mean(preds$tslength[coh.chlaXaccndvi$accndvip.ts

gam.phist<-gam(cos(accndviphi.ts1) ~ s(wlconnections_openwaterwetlands_shoreline_km) +
                 s(buffer500m_streamdensity_headwaters_density_mperha) +
                 s(buffer500m_streamdensity_streams_density_mperha),
               data=rfdat.phist, gamma=1, weights=lwgt)
gam.check(gam.phist)
```




```
##
## Method: GCV   Optimizer: magic
## Smoothing parameter selection converged after 19 iterations.
## The RMS GCV score gradient at convergence was 3.737094e-08 .
## The Hessian was positive definite.
## Model rank =  28 / 28
##
## Basis dimension (k) checking results. Low p-value (k-index<1) may
## indicate that k is too low, especially if edf is close to k'.
##
##                                                        k' edf k-index
## s(wlconnections_openwaterwetlands_shoreline_km)        9   1    1.08
## s(buffer500m_streamdensity_headwaters_density_mperha)  9   1    1.03
## s(buffer500m_streamdensity_streams_density_mperha)     9   1    1.08
##                                                       p-value
## s(wlconnections_openwaterwetlands_shoreline_km)          0.62
## s(buffer500m_streamdensity_headwaters_density_mperha)    0.50
## s(buffer500m_streamdensity_streams_density_mperha)       0.59
```

```
concurvity(gam.phist)
```

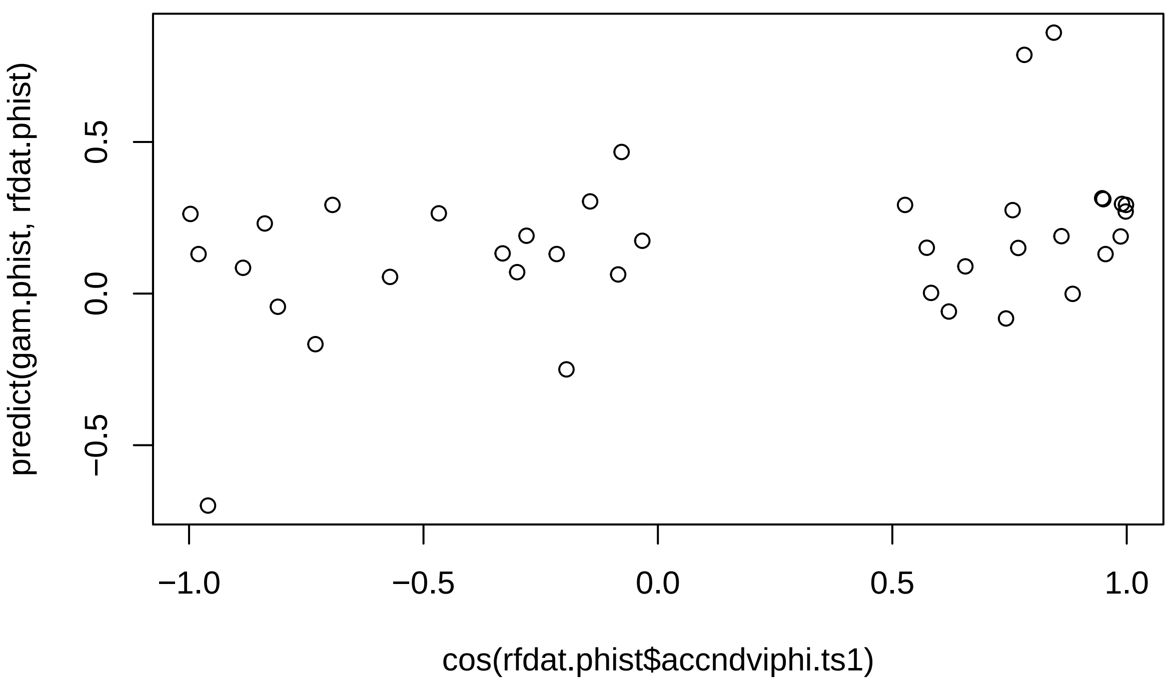
```
##                  para s(wlconnections_openwaterwetlands_shoreline_km)
## worst    1.188255e-15                                       0.9998076
## observed 1.188255e-15                                       0.4282308
## estimate 1.188255e-15                                       0.3699462
##          s(buffer500m_streamdensity_headwaters_density_mperha)
## worst                                                1.0000000
## observed                                             0.9976951
## estimate                                             0.9953973
##          s(buffer500m_streamdensity_streams_density_mperha)
```

```
## worst                                                        1.000000
## observed                                                     0.944058
## estimate                                                     0.905371
```

```r
summary(gam.phist)
```

```
##
## Family: gaussian
## Link function: identity
##
## Formula:
## cos(accndviphi.ts1) ~ s(wlconnections_openwaterwetlands_shoreline_km) +
##     s(buffer500m_streamdensity_headwaters_density_mperha) + s(buffer500m_streamdensity_streams_densi
##
## Parametric coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.1619     0.1133   1.429    0.162
##
## Approximate significance of smooth terms:
##                                                        edf Ref.df     F
## s(wlconnections_openwaterwetlands_shoreline_km)          1      1 1.611
## s(buffer500m_streamdensity_headwaters_density_mperha)    1      1 0.034
## s(buffer500m_streamdensity_streams_density_mperha)       1      1 0.100
##                                                        p-value
## s(wlconnections_openwaterwetlands_shoreline_km)          0.213
## s(buffer500m_streamdensity_headwaters_density_mperha)    0.855
## s(buffer500m_streamdensity_streams_density_mperha)       0.754
##
## R-sq.(adj) =  0.0512   Deviance explained = 12.8%
## GCV = 0.5431  Scale est. = 0.48593   n = 38
```

```r
plot(cos(rfdat.phist$accndviphi.ts1), predict(gam.phist, rfdat.phist))
```



```
rfdat.philt<-left_join(coh.chlaXaccndvi[,c(10,5)], preds)
```

```
## Joining, by = "lagoslakeid"
```

```
rfdat.philt<-rfdat.philt[,!colnames(rfdat.philt) %in%
                           c("lagoslakeid","start","end","lakes_nhdid","hu12_zoneid","tslength","county_
rfdat.philt<-rfdat.philt[,!grepl("borderhu12s",colnames(rfdat.philt))]

rfdat.philt<-rfdat.philt[coh.chlaXaccndvi$accndvip.ts2<0.3,]

for(nn in 1:ncol(rfdat.philt)){
  if(is.character(rfdat.philt[,nn])){
    rfdat.philt[,nn]<-as.factor(rfdat.philt[,nn])
  }
}

cf.philt<-party::cforest(cos(accndviphi.ts2) ~ ., data=rfdat.philt,
                         controls=cforest_control(ntree=50000,mincriterion = 0.9,mtry=3))

varimp.phi.lt<-varimp(cf.philt)
print(varimp.phi.lt[order(varimp.phi.lt, decreasing=T)][1:10])
```

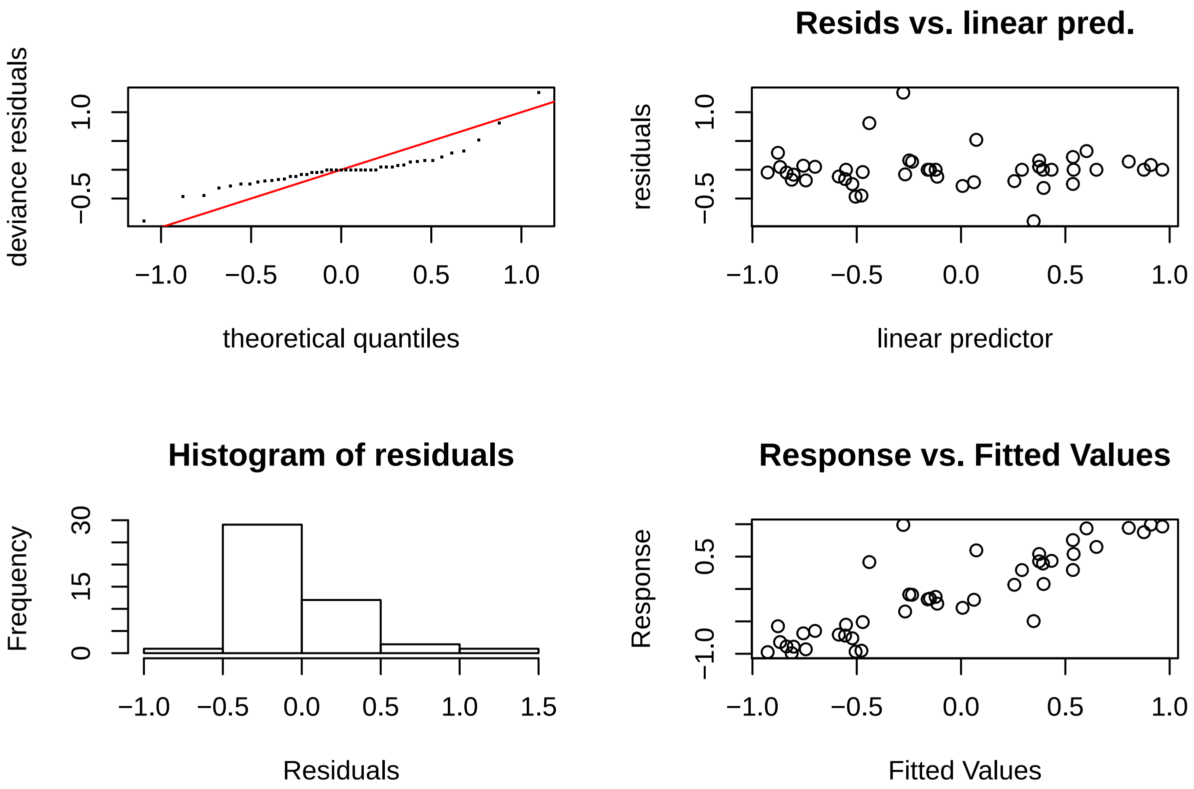
```
##                                          hu4_zoneid
##                                         0.011753488
##                                          hu6_zoneid
##                                         0.009020676
##                                     hu12_slope_mean
##                                         0.007476076
##                                       hu12_tri_mean
##                                         0.007014192
##                                 hu12_nlcd2011_pct_90
##                                         0.004850795
##                                 hu12_nlcd2011_pct_41
##                                         0.004511979
##                            hu12_dep_totaln_tavg_mean
##                                         0.004094007
## buffer500m_streamdensity_headwaters_density_mperha
##                                         0.004016839
##                                          hu8_zoneid
##                                         0.003994655
##    buffer500m_streamdensity_streams_density_mperha
##                                         0.002823475
```

```
predphi.lt<-predict(cf.philt, newdata=rfdat.philt,type="response")
cor.test(predphi.lt,cos(rfdat.philt$accndviphi.ts2))
```

```
##
##  Pearson's product-moment correlation
##
## data:  predphi.lt and cos(rfdat.philt$accndviphi.ts2)
## t = 10.401, df = 43, p-value = 2.579e-13
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.7348649 0.9128035
## sample estimates:
```

```
##      cor
## 0.845926
```

```
lwgt<-preds$tslength[coh.chlaXaccndvi$accndvip.ts2<0.3]/mean(preds$tslength[coh.chlaXaccndvi$accndvip.t

gam.philt<-gam(cos(accndviphi.ts2) ~ hu4_zoneid + s(hu12_slope_mean) +
                 s(hu12_nlcd2011_pct_41),
               data=rfdat.philt, gamma=1, weights=lwgt)
gam.check(gam.philt)
```




```
##
## Method: GCV   Optimizer: magic
## Smoothing parameter selection converged after 15 iterations.
## The RMS GCV score gradient at convergence was 6.449737e-08 .
## The Hessian was positive definite.
## Model rank =  40 / 40
##
## Basis dimension (k) checking results. Low p-value (k-index<1) may
## indicate that k is too low, especially if edf is close to k'.
##
##                           k' edf k-index p-value
## s(hu12_slope_mean)        9   1    1.21    0.90
## s(hu12_nlcd2011_pct_41)  9   1    1.16    0.84
```

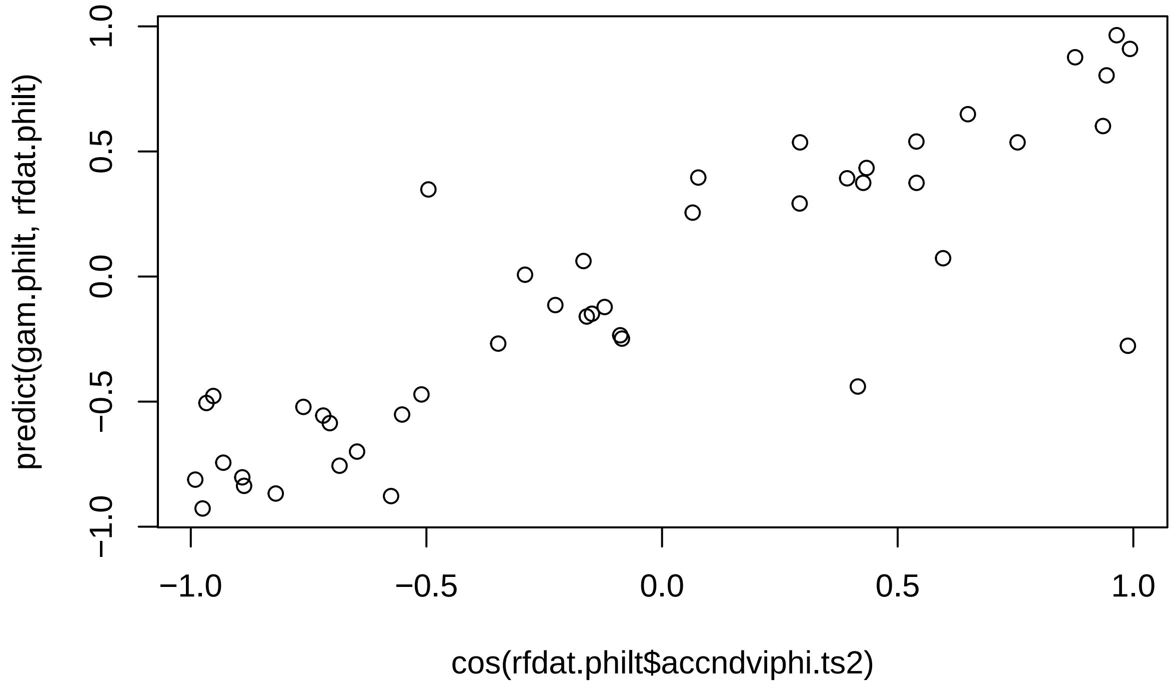
```
concurvity(gam.philt)
```

```
##              para s(hu12_slope_mean) s(hu12_nlcd2011_pct_41)
## worst    0.984192          0.9819486               0.9818135
## observed 0.984192          0.8987787               0.8624219
## estimate 0.984192          0.8822897               0.8418742
```

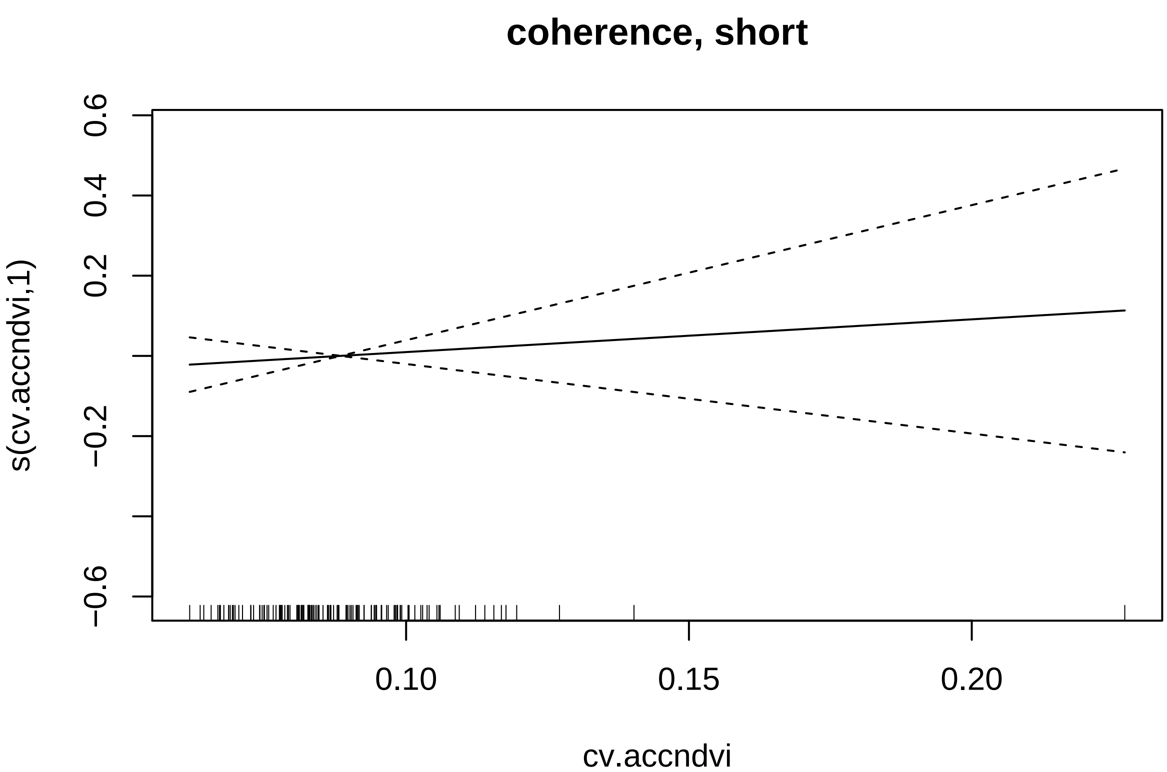
```r
summary(gam.philt)
```

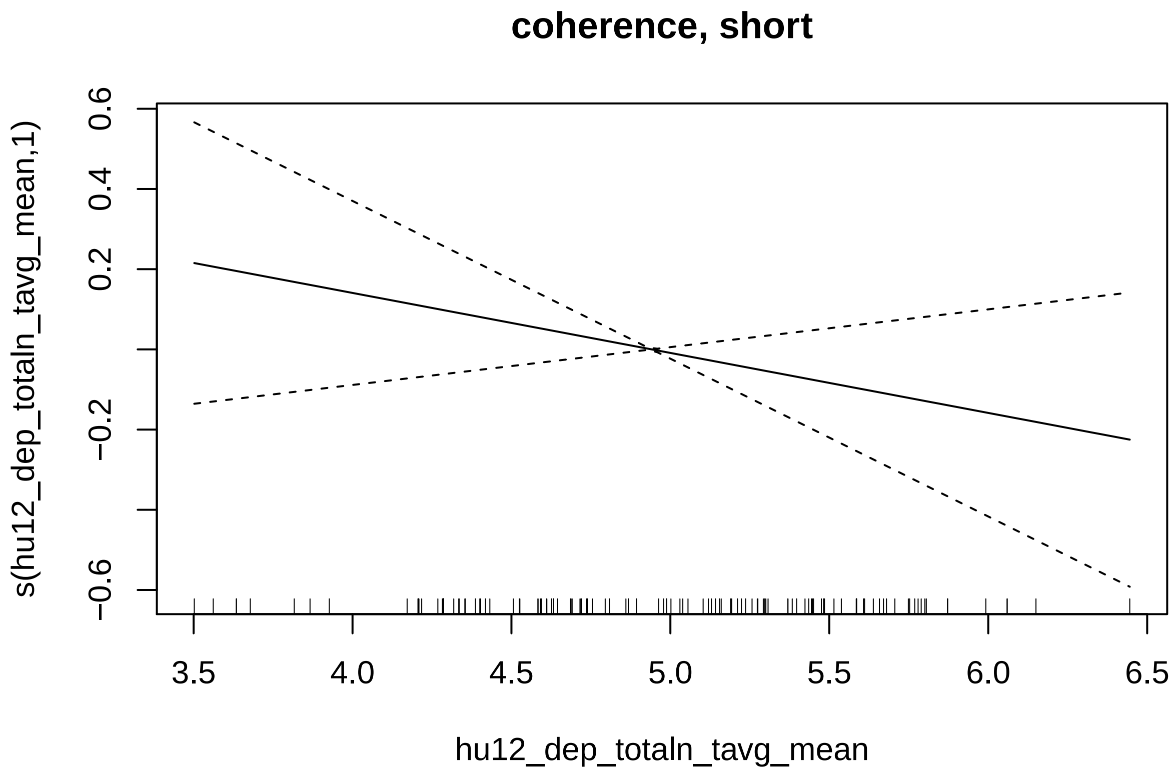
```
##
## Family: gaussian
## Link function: identity
##
## Formula:
## cos(accndviphi.ts2) ~ hu4_zoneid + s(hu12_slope_mean) + s(hu12_nlcd2011_pct_41)
##
## Parametric coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)      -0.651351   0.349285  -1.865  0.07624 .
## hu4_zoneidHU4_12 -0.028146   0.564751  -0.050  0.96072
## hu4_zoneidHU4_16  1.417506   0.614320   2.307  0.03131 *
## hu4_zoneidHU4_18  1.312060   0.624254   2.102  0.04781 *
## hu4_zoneidHU4_25  0.516108   0.426450   1.210  0.23963
## hu4_zoneidHU4_27  0.464429   0.605579   0.767  0.45167
## hu4_zoneidHU4_29  1.870058   0.500044   3.740  0.00121 **
## hu4_zoneidHU4_30  1.295962   0.493991   2.623  0.01588 *
## hu4_zoneidHU4_32  0.277325   0.432565   0.641  0.52838
## hu4_zoneidHU4_33  0.340073   0.425823   0.799  0.43345
## hu4_zoneidHU4_35  1.438875   0.606217   2.374  0.02724 *
## hu4_zoneidHU4_36  1.232482   0.611728   2.015  0.05691 .
## hu4_zoneidHU4_4  -0.003587   0.462749  -0.008  0.99389
## hu4_zoneidHU4_5   0.650996   0.500521   1.301  0.20748
## hu4_zoneidHU4_51  1.252175   0.607479   2.061  0.05188 .
## hu4_zoneidHU4_54  1.844177   0.588426   3.134  0.00501 **
## hu4_zoneidHU4_60  0.783300   0.423913   1.848  0.07877 .
## hu4_zoneidHU4_63 -0.142882   0.582825  -0.245  0.80872
## hu4_zoneidHU4_64  0.100078   0.586574   0.171  0.86616
## hu4_zoneidHU4_65  0.804693   0.571655   1.408  0.17386
## hu4_zoneidHU4_67  0.441239   0.737271   0.598  0.55592
## hu4_zoneidHU4_7  -0.739634   0.584041  -1.266  0.21923
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##                         edf Ref.df      F p-value
## s(hu12_slope_mean)        1      1 12.442 0.00191 **
## s(hu12_nlcd2011_pct_41)   1      1  7.458 0.01235 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.465   Deviance explained = 74.5%
## GCV = 0.49152  Scale est. = 0.22938   n = 45
```

```r
plot(cos(rfdat.philt$accndviphi.ts2), predict(gam.philt, rfdat.philt))
```

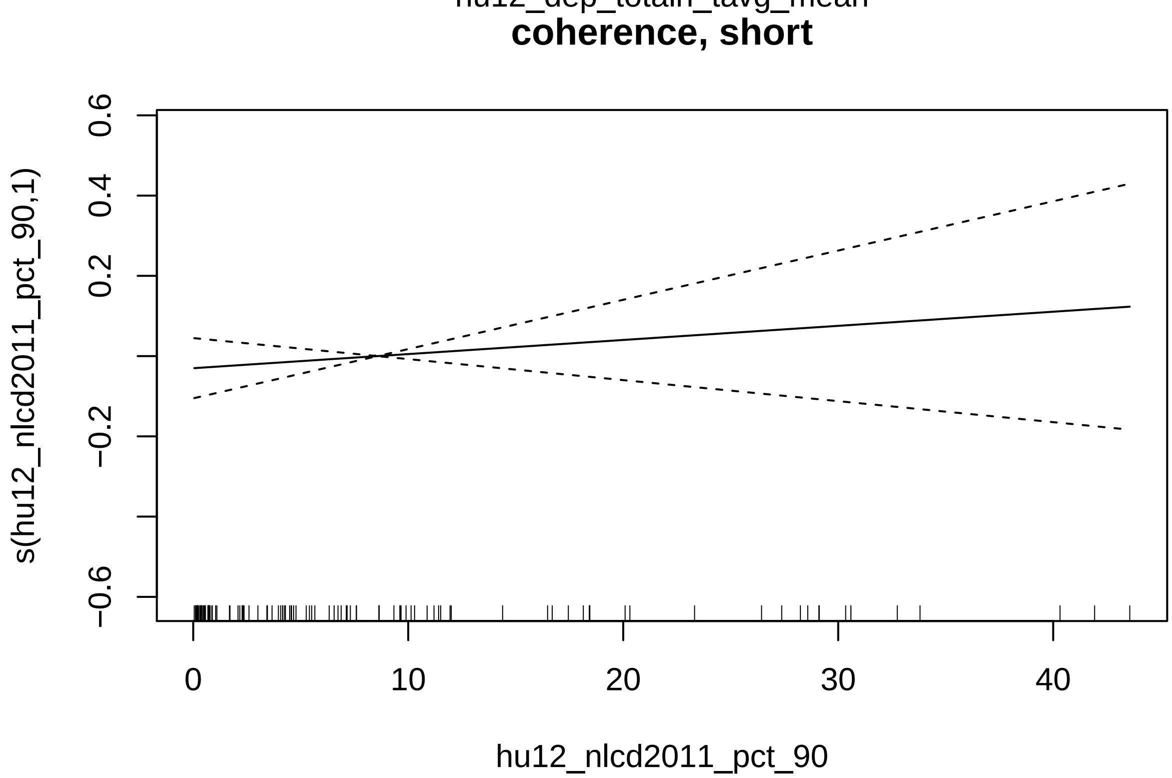



```
plot(gam.cohst, main="coherence, short")
```

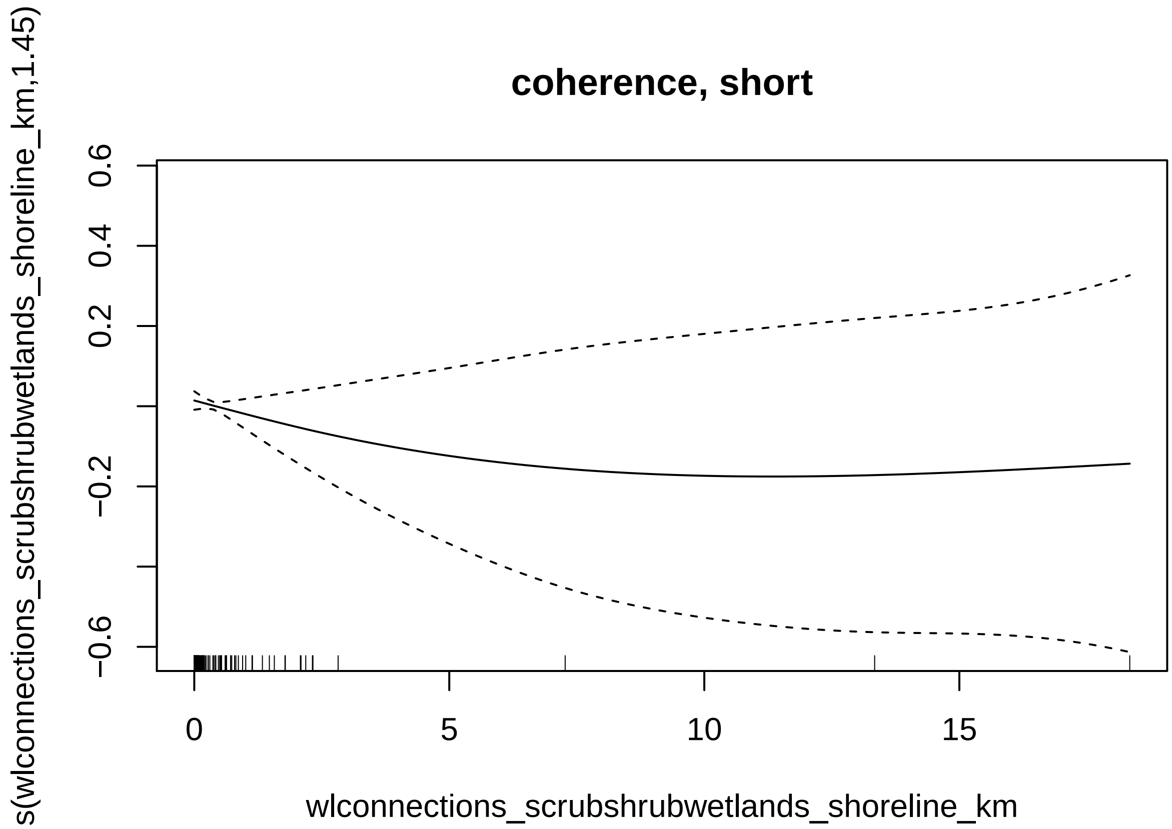




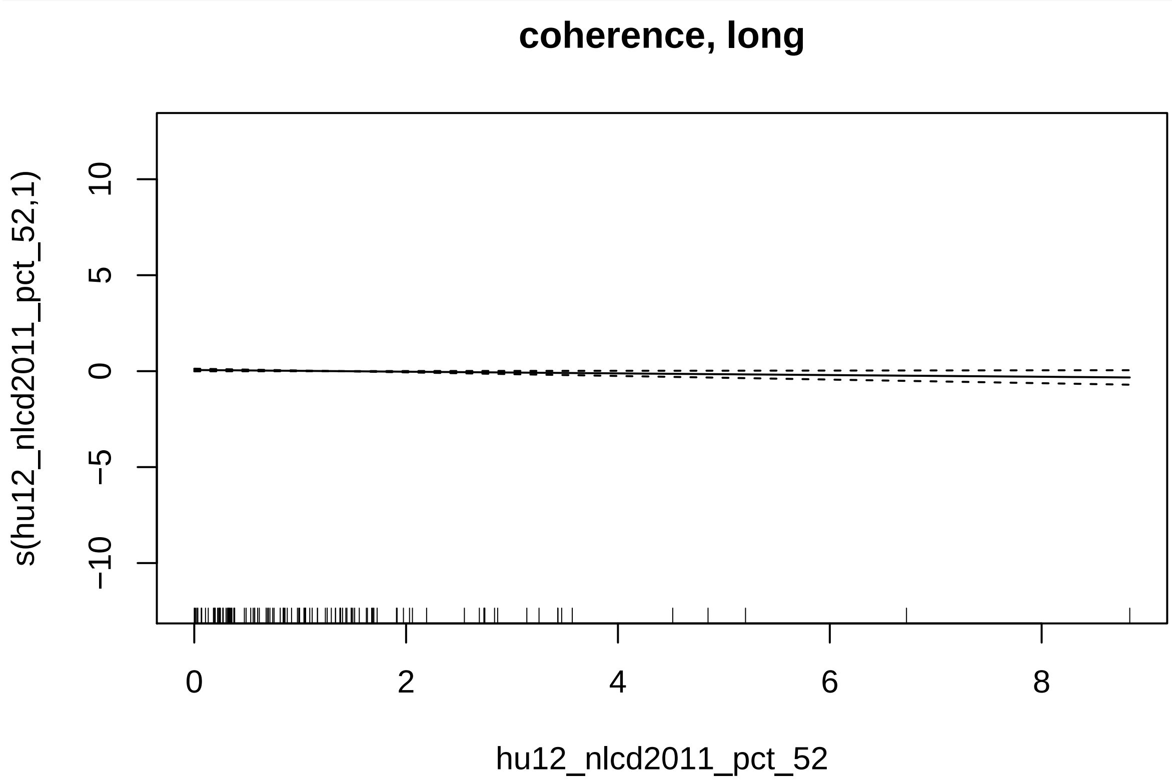
coherence, short
s(hu12_dep_totaln_tavg_mean,1)
0.6
0.4
0.2
−0.2
−0.6
3.5
4.0
4.5
5.0
5.5
6.0
6.5
hu12_dep_totaln_tavg_mean

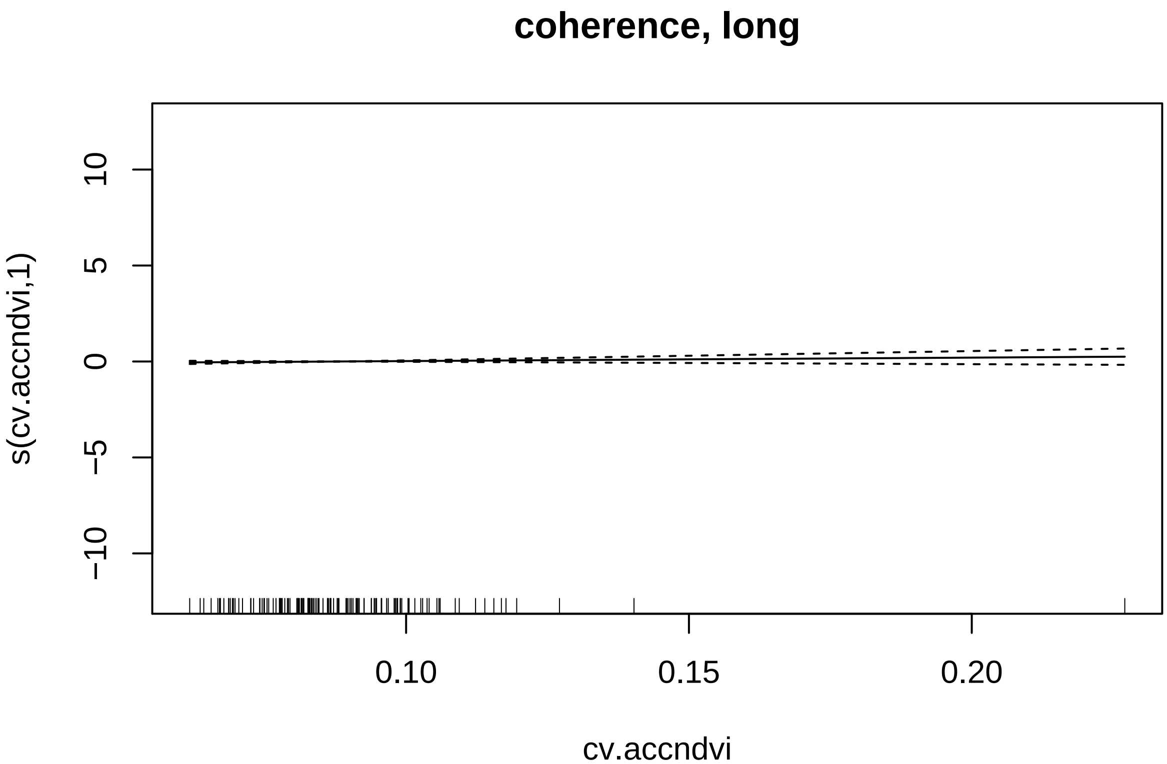


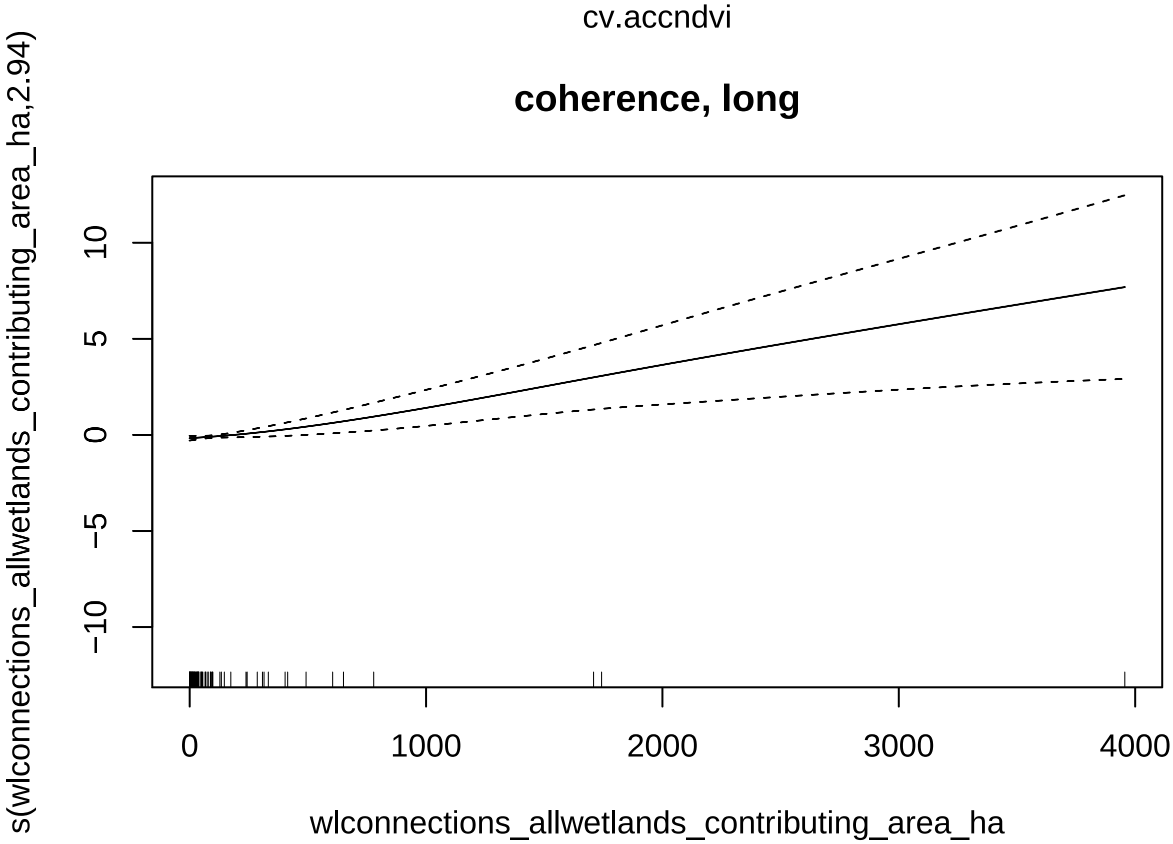
coherence, short
s(hu12_nlcd2011_pct_90,1)
0.6
0.4
0.2
−0.2
−0.6
0
10
20
30
40
hu12_nlcd2011_pct_90

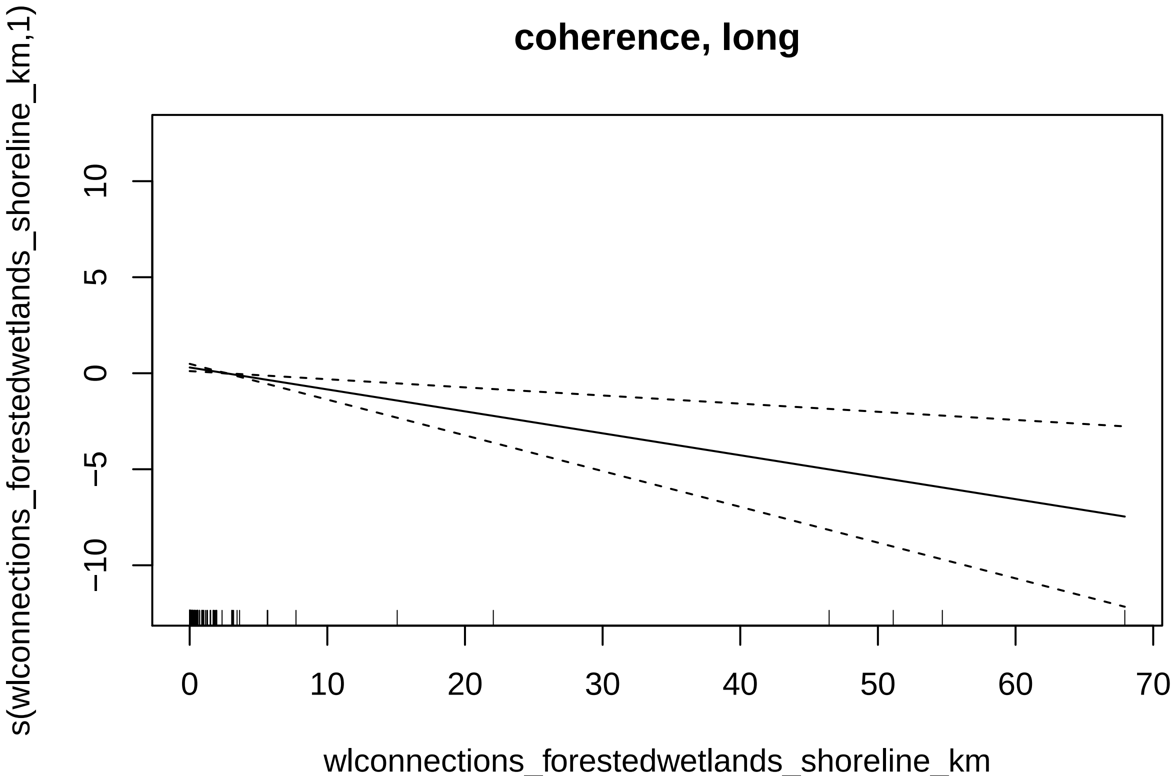



```
plot(gam.cohlt, main="coherence, long")
```

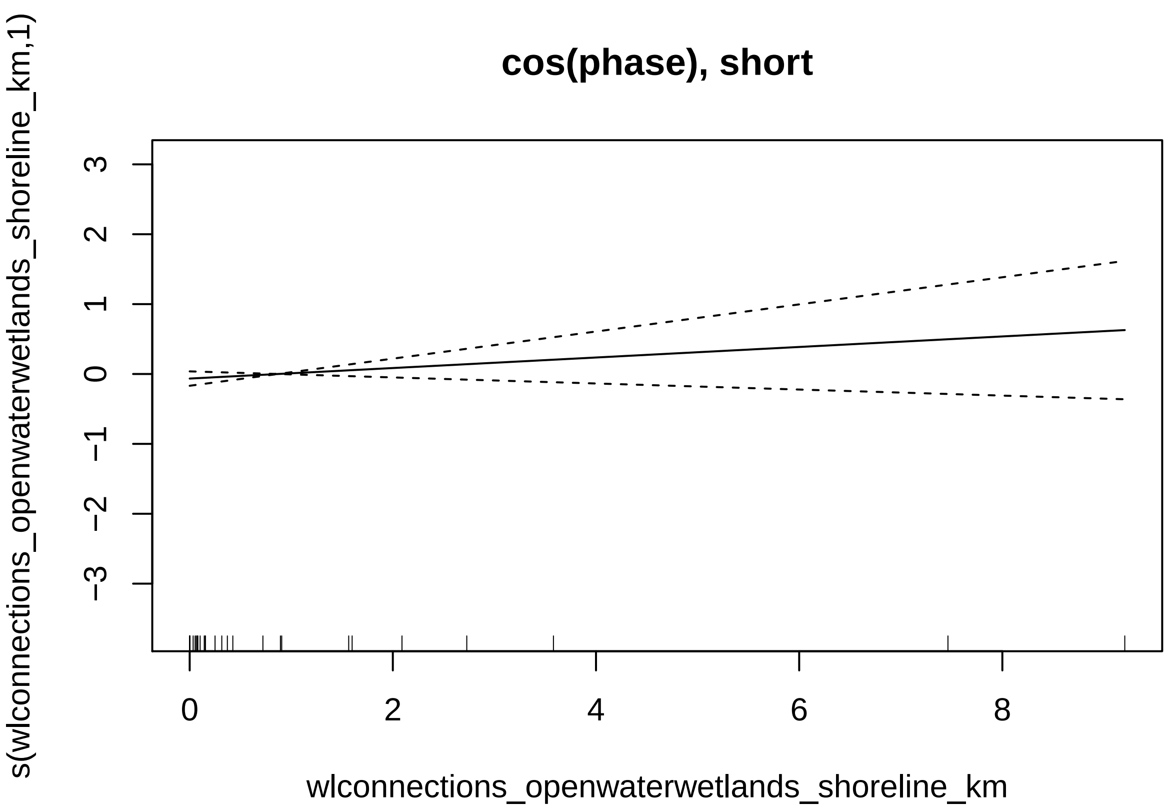


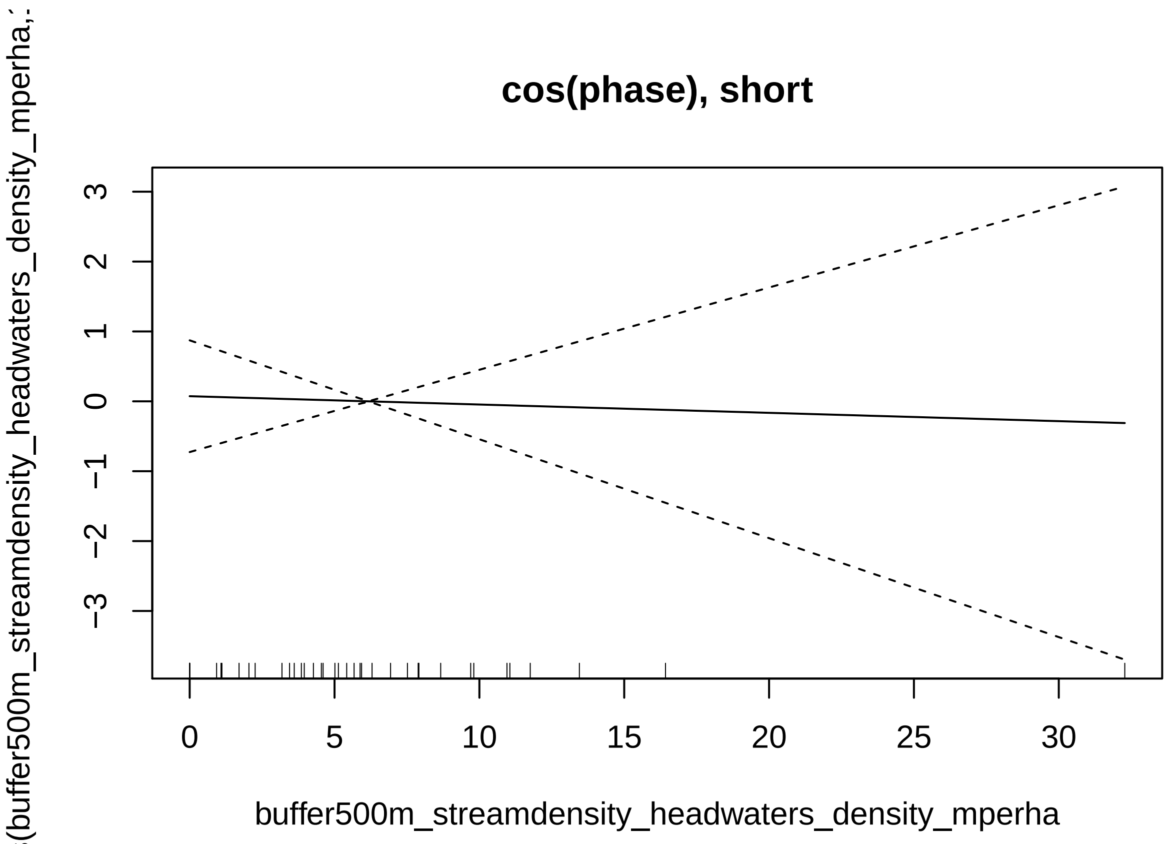

coherence, long
s(cv.accndvi,1)
10
5
0
−5
−10
0.10
0.15
0.20
cv.accndvi

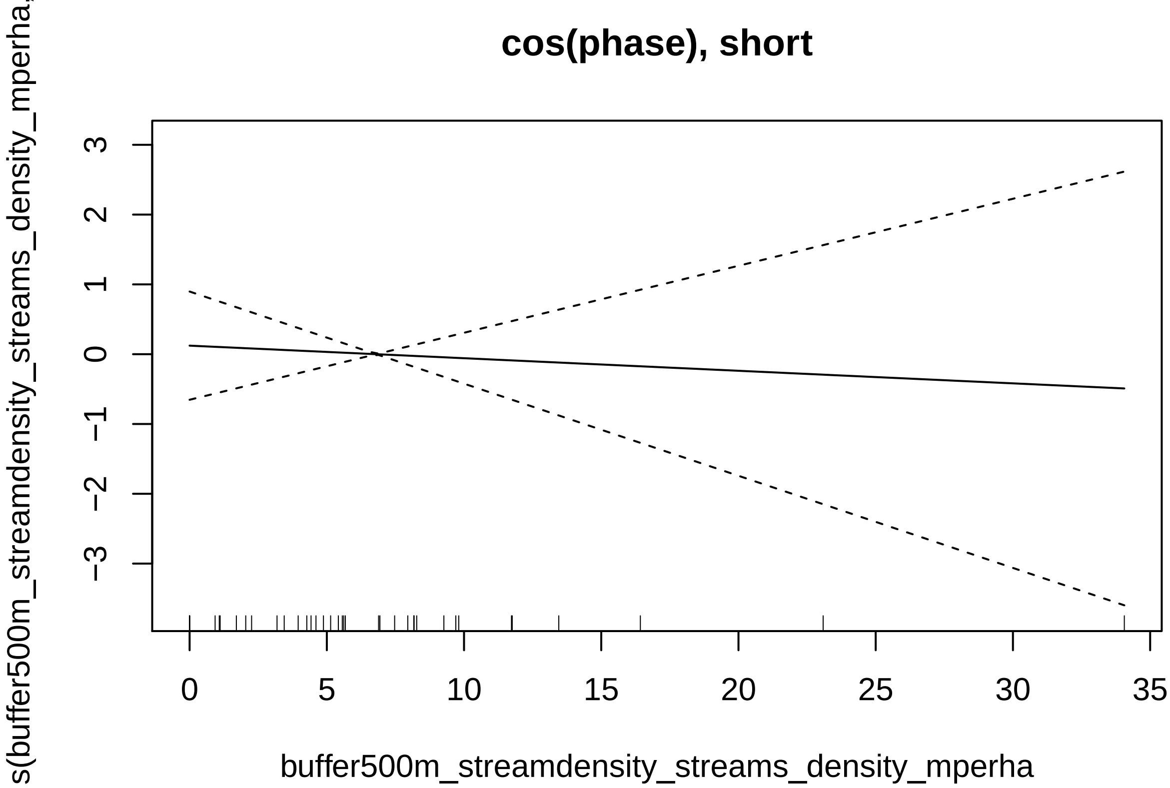

coherence, long
s(wlconnections_allwetlands_contributing_area_ha,2.94)
10
5
0
−5
−10
0
1000
2000
3000
4000
wlconnections_allwetlands_contributing_area_ha



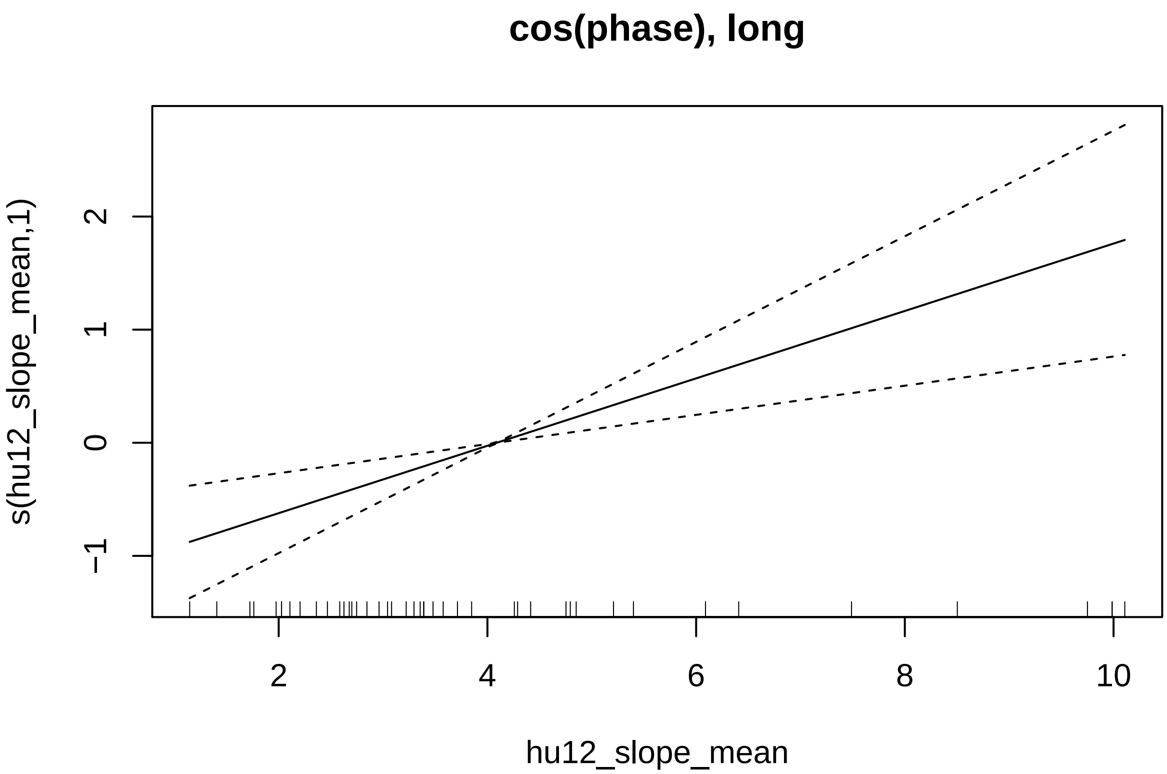

```
plot(gam.phist, main="cos(phase), short")
```

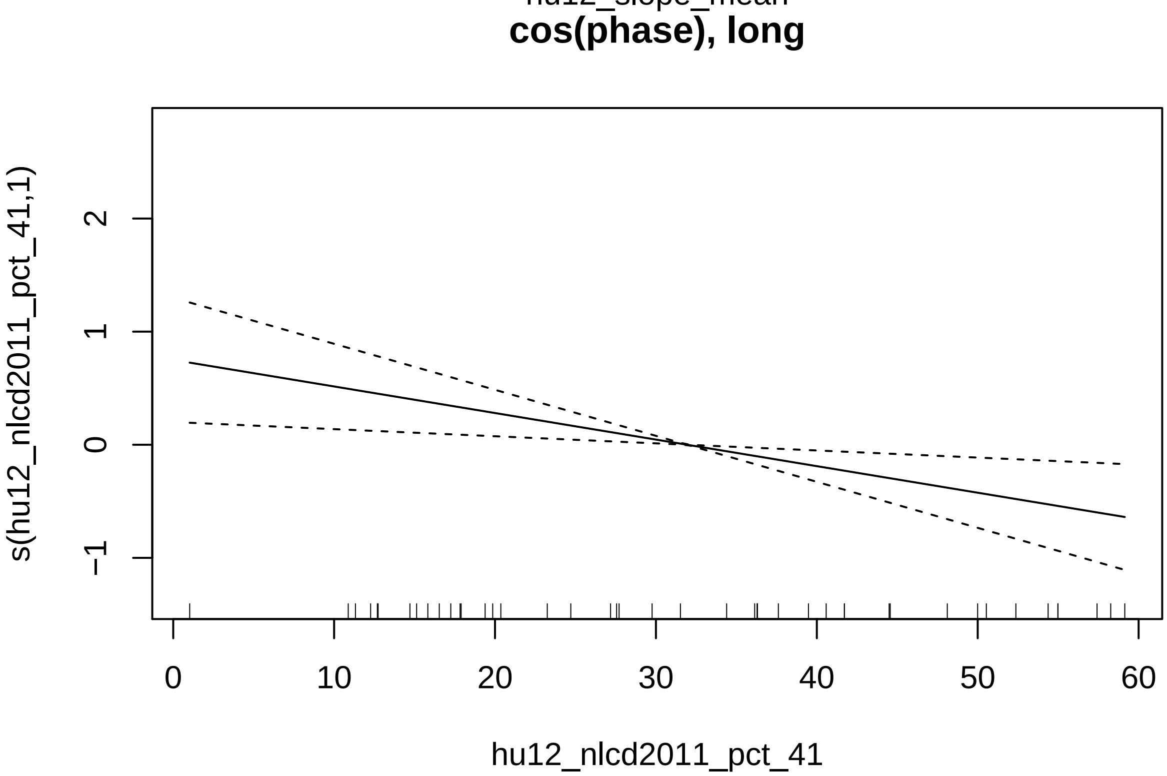



cos(phase), short
s(buffer500m_streamdensity_headwaters_density_mperha,1
3
2
1
0
−1
−2
−3
0
5
10
15
20
25
30
buffer500m_streamdensity_headwaters_density_mperha


cos(phase), short
s(buffer500m_streamdensity_streams_density_mperha,1)
3
2
1
0
−1
−2
−3
0
5
10
15
20
25
30
35
buffer500m_streamdensity_streams_density_mperha

```r
plot(gam.philt, main="cos(phase), long")
```






```r
print(varimp.coh.st[order(varimp.coh.st, decreasing=T)][1:10])
```

```
##                                        cv.accndvi
##                                      0.0005710355
```

```
##                                    hu8_zoneid
##                                  0.0005568739
##                     hu12_dep_totaln_tavg_mean
##                                  0.0004949402
##                          hu12_nlcd2011_pct_90
##                                  0.0001939783
##                          hu12_nlcd2011_pct_95
##                                  0.0001609831
## wlconnections_scrubshrubwetlands_shoreline_km
##                                  0.0001473959
##        wlconnections_allwetlands_shoreline_km
##                                  0.0001400305
##                               hu12_slope_mean
##                                  0.0001274735
##                                 hu12_tri_mean
##                                  0.0001274102
##   wlconnections_forestedwetlands_shoreline_km
##                                  0.0000901705
```

```
ltxt.st<-c("HUC-8 sub-basin","cv(NDVI)","total N deposition","% glaciofluvial outwash","% woody wetlands
           "shrub wetlands shoreline","% herbaceous wetlands","topographic roughness","slope","wetlands

print(varimp.coh.lt[order(varimp.coh.lt, decreasing=T)][1:10])
```

```
##                                    hu8_zoneid
##                                  1.203560e-03
##                          hu12_nlcd2011_pct_52
##                                  9.442723e-04
##                                    cv.accndvi
##                                  4.297765e-04
##   wlconnections_forestedwetlands_shoreline_km
##                                  1.483532e-04
##  hu12_prism_ppt_30yr_normal_800mm2_annual_mean
##                                  1.395582e-04
## wlconnections_allwetlands_contributing_area_ha
##                                  1.272004e-04
##                          hu12_nlcd2011_pct_95
##                                  1.247903e-04
##                        hu12_dep_so4_tavg_mean
##                                  8.831864e-05
##        wlconnections_allwetlands_shoreline_km
##                                  8.766973e-05
##                   hu12_groundwaterrecharge_mean
##                                  5.901385e-05
```

```
ltxt.lt<-c("HUC-8 sub-basin","% shrub/scrub","cv(NDVI)","% alluvial soils","% solution residuum",
           "all wetlands contrib. area","% till & loam","annual precipitation","% ice-contact depositi
           "groundwater recharge")

print(varimp.phi.st[order(varimp.phi.st, decreasing = T)][1:10])
```

```
##       wlconnections_openwaterwetlands_shoreline_km
##                                       0.0033487759
## buffer500m_streamdensity_headwaters_density_mperha
##                                       0.0032202533
```

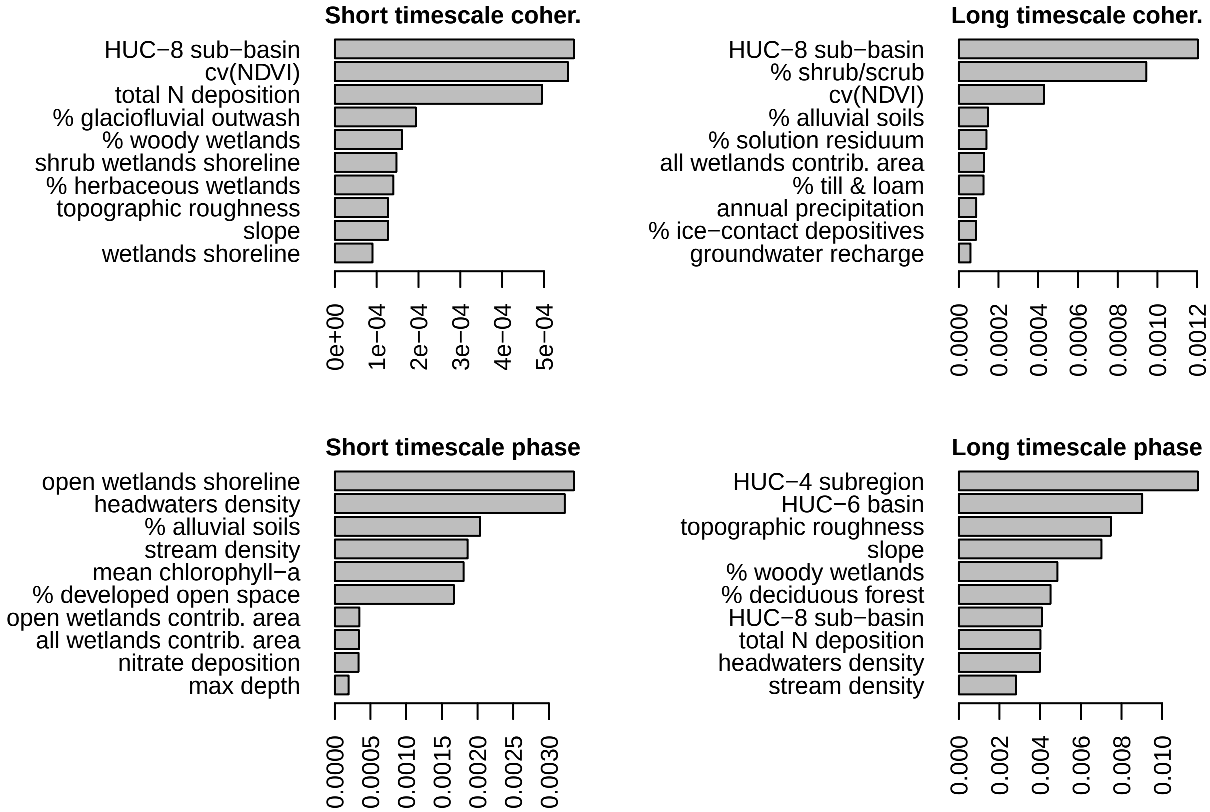
```
##                                               chla
##                                       0.0020374995
##                               hu12_nlcd2011_pct_21
##                                       0.0018605382
## wlconnections_openwaterwetlands_contributing_area_
##                                       0.0018041944
##    buffer500m_streamdensity_streams_density_mperha
##                                       0.0016664426
##                             hu12_dep_no3_tavg_mean
##                                       0.0003453840
##                               hu12_nlcd2011_pct_22
##                                       0.0003389173
##                     hu12_roaddensity_density_mperha
##                                       0.0003335029
##     wlconnections_allwetlands_contributing_area_ha
##                                       0.0001936666
```

```
ltxt.phist<-c("open wetlands shoreline","headwaters density","% alluvial soils","stream density","mean 
              "% developed open space","open wetlands contrib. area", "all wetlands contrib. area",
              "nitrate deposition","max depth")

print(varimp.phi.lt[order(varimp.phi.lt, decreasing = T)][1:10])
```

```
##                                        hu4_zoneid
##                                       0.011753488
##                                        hu6_zoneid
##                                       0.009020676
##                                   hu12_slope_mean
##                                       0.007476076
##                                     hu12_tri_mean
##                                       0.007014192
##                              hu12_nlcd2011_pct_90
##                                       0.004850795
##                              hu12_nlcd2011_pct_41
##                                       0.004511979
##                         hu12_dep_totaln_tavg_mean
##                                       0.004094007
## buffer500m_streamdensity_headwaters_density_mperha
##                                       0.004016839
##                                        hu8_zoneid
##                                       0.003994655
##    buffer500m_streamdensity_streams_density_mperha
##                                       0.002823475
```

```
ltxt.philt<-c("HUC-4 subregion","HUC-6 basin","topographic roughness","slope","% woody wetlands","% dec
               "HUC-8 sub-basin","total N deposition","headwaters density","stream density")

#tiff("~/Box Sync/NSF EAGER Synchrony/Manuscripts/1_CoherenceSpatialVariation/fig4_varimp_top10.tif",un

par(mfrow=c(2,2), mar=c(5,11,1,1), cex.main=0.9, cex.axis=0.9)

barplot(rev(varimp.coh.st[order(varimp.coh.st, decreasing=T)][1:10]),names.arg=rev(ltxt.st),las=2,main=

barplot(rev(varimp.coh.lt[order(varimp.coh.lt, decreasing=T)][1:10]),names.arg=rev(ltxt.lt),las=2,main=
```

```
barplot(rev(varimp.phi.st[order(varimp.phi.st, decreasing=T)][1:10]),names.arg=rev(ltxt.phist),las=2,ma

barplot(rev(varimp.phi.lt[order(varimp.phi.lt, decreasing=T)][1:10]),names.arg=rev(ltxt.philt),las=2,ma
```




```
#dev.off()
```

```
mar1<-c(3,1.5,0.5,1)
#mar2<-c(2.5,1.5,1,1)
fudge=1/40

#tiff("~/Box Sync/NSF EAGER Synchrony/Manuscripts/1_CoherenceSpatialVariation/fig5_gamfits.tif",units="

par(mfrow=c(4,5),mgp=c(1.5,0.5,0),oma=c(0,2.5,0,0),mar=mar1)

plot(NA,NA,xlab="HUC-8 sub-basin",xlim=c(0,1),ylim=c(0,1))
plot(gam.cohst,select=1,residuals=T,rug=FALSE,shade=T,cex=2,xlab="cv(NDVI)",ylab="short coherence")
plot(gam.cohst,select=2,residuals=T,rug=FALSE,shade=T,cex=2,xlab="total N deposition",ylab="")
plot(gam.cohst,select=3,residuals=T,rug=FALSE,shade=T,cex=2,xlab="% glaciofluvial outwash",ylab="")
plot(gam.cohst,select=4,residuals=T,rug=FALSE,shade=T,cex=2,xlab="% woody wetlands",ylab="")

plot(gam.phist,select=1,residuals=T,rug=FALSE,shade=T,cex=2,xlab="open wetlands shoreline",ylab="short p
plot(gam.phist,select=2,residuals=T,rug=FALSE,shade=T,cex=2,xlab="headwaters density",ylab="")
plot(gam.phist,select=3,residuals=T,rug=FALSE,shade=T,cex=2,xlab="% alluvial soils",ylab="")
plot.new()
plot.new()

plot(NA,NA,xlab="HUC-8 sub-basin",xlim=c(0,1),ylim=c(0,1))
plot(gam.cohlt,select=1,residuals=T,rug=FALSE,shade=T,cex=2,xlab="% shrub/scrub",ylab="long coherence")
```

```
plot(gam.cohlt,select=2,residuals=T,rug=FALSE,shade=T,cex=2,xlab="cv(NDVI)",ylab="")
plot(gam.cohlt,select=3,residuals=T,rug=FALSE,shade=T,cex=2,xlab="% alluvial soils",ylab="")
plot(gam.cohlt,select=4,residuals=T,rug=FALSE,shade=T,cex=2,xlab="% solution residuum",ylab="")

plot(NA,NA,xlab="HUC-4 subregion",xlim=c(0,1),ylim=c(0,1),ylab="long phase")
plot(gam.philt,select=1,residuals=T,rug=FALSE,shade=T,cex=2,xlab="topographic roughness",ylab="")
plot(gam.philt,select=2,residuals=T,rug=FALSE,shade=T,cex=2,xlab="% woody wetlands",ylab="")

mtext("Partial residuals",2,outer=T,line=1.2,cex=0.8)
mtext("cos(long phase)",2,at=1/8+fudge,outer=T,line=0,cex=0.7)
mtext("long coherence)",2,at=1/8+1/4+fudge,outer=T,line=0,cex=0.7)
mtext("cos(short phase)",2,at=1/8+2/4+fudge,outer=T,line=0,cex=0.7)
mtext("short coherence",2,at=1/8+3/4+fudge,outer=T,line=0,cex=0.7)

#dev.off()
```