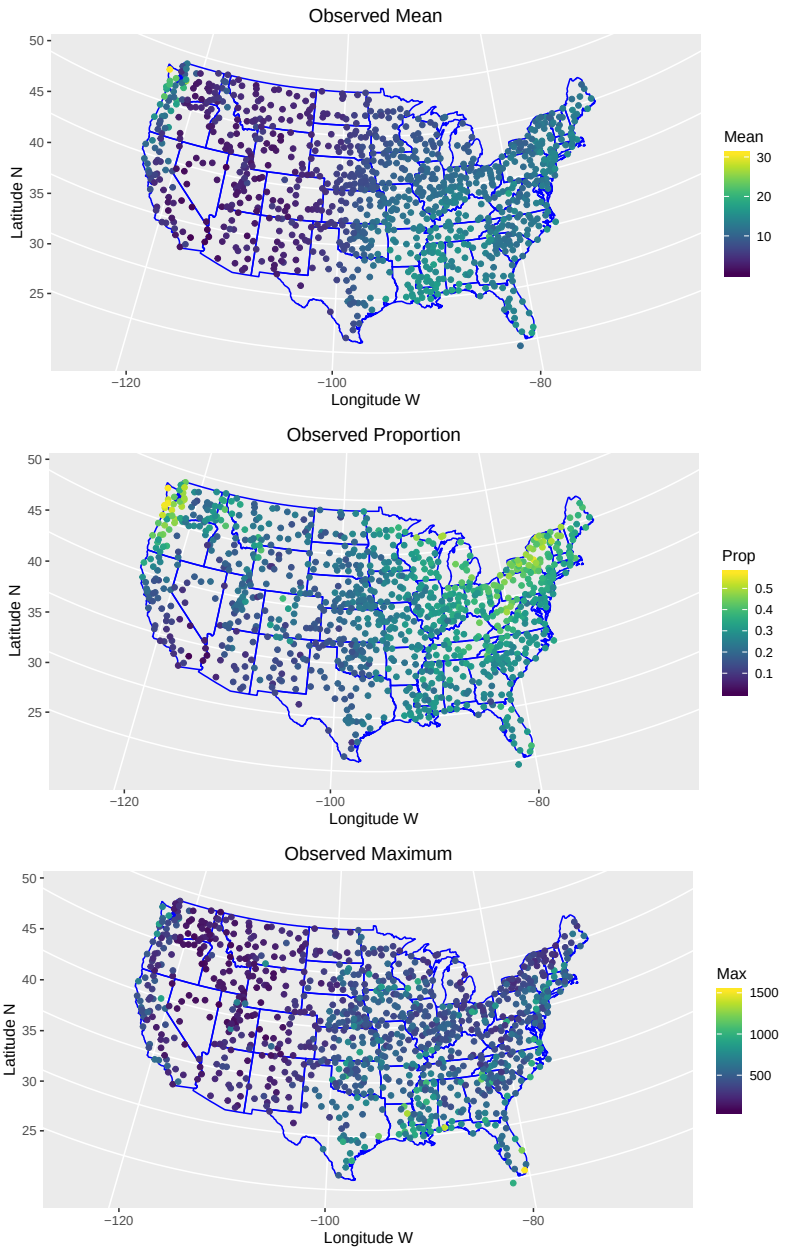
Homework 2: Solution

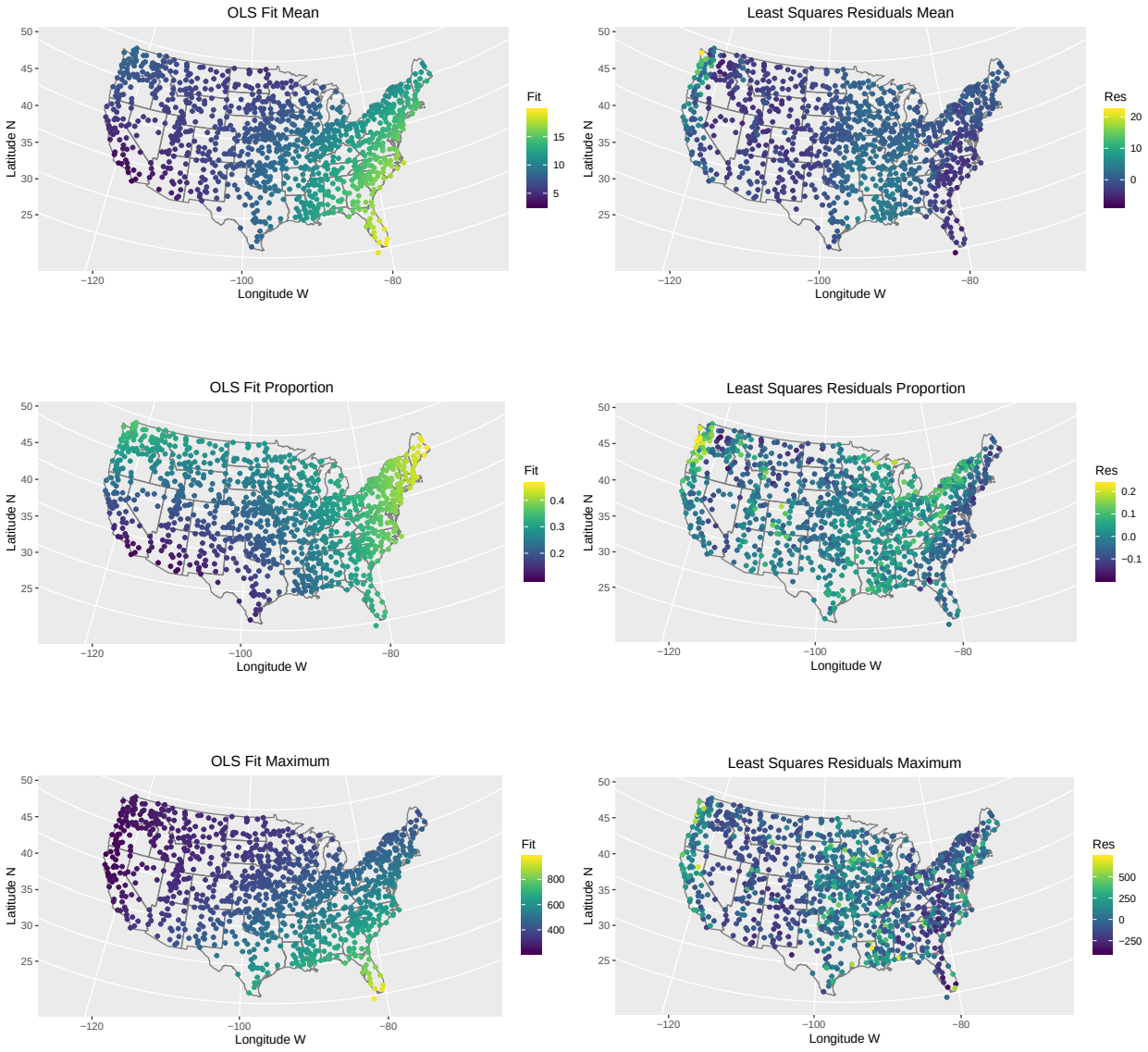
September 2, 2020

1) Mapping the summary statistics of precipitation:



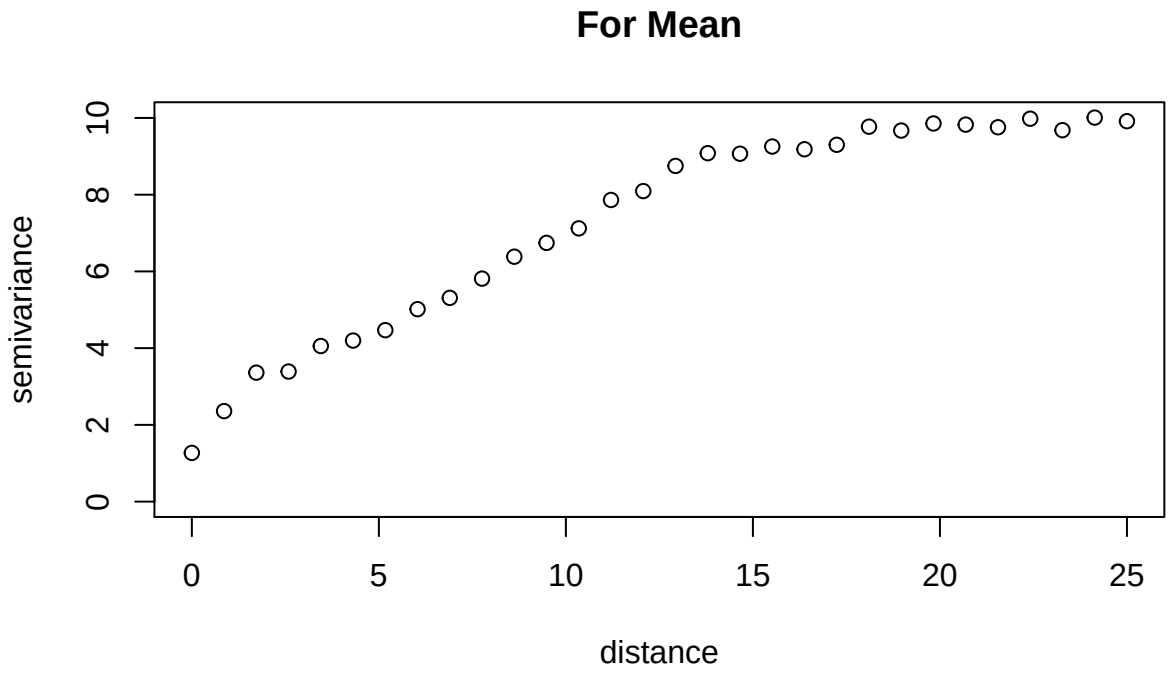
There appears to be spatial dependence in all three of the plots.

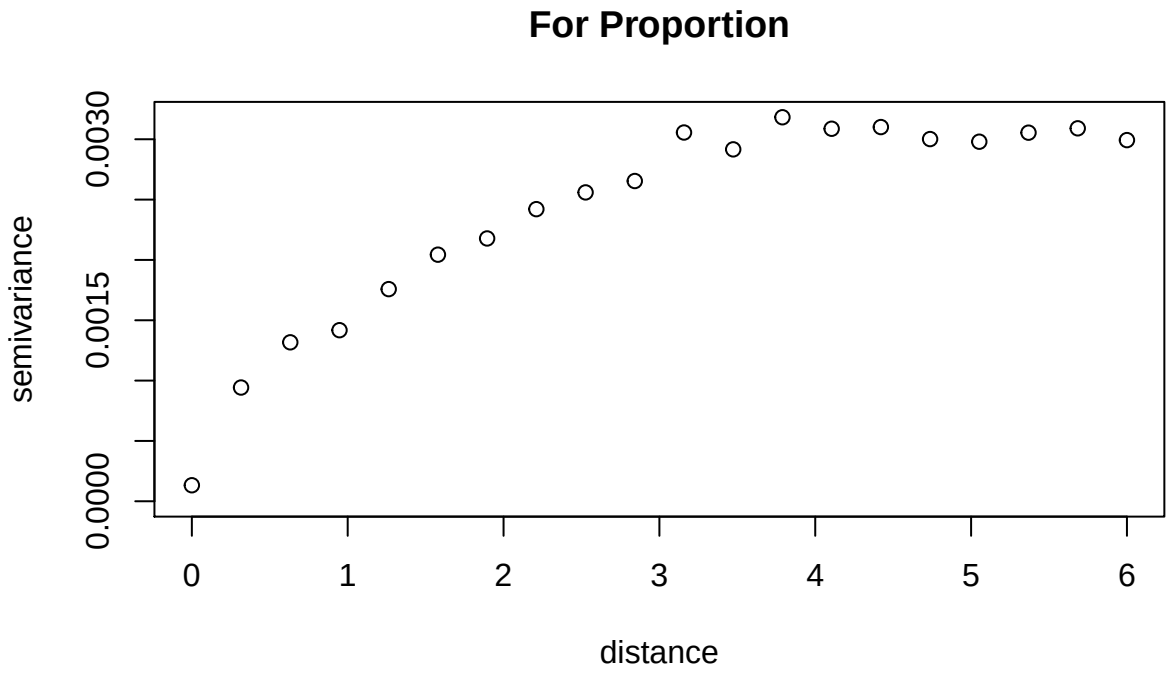
2) Fits and Residuals:

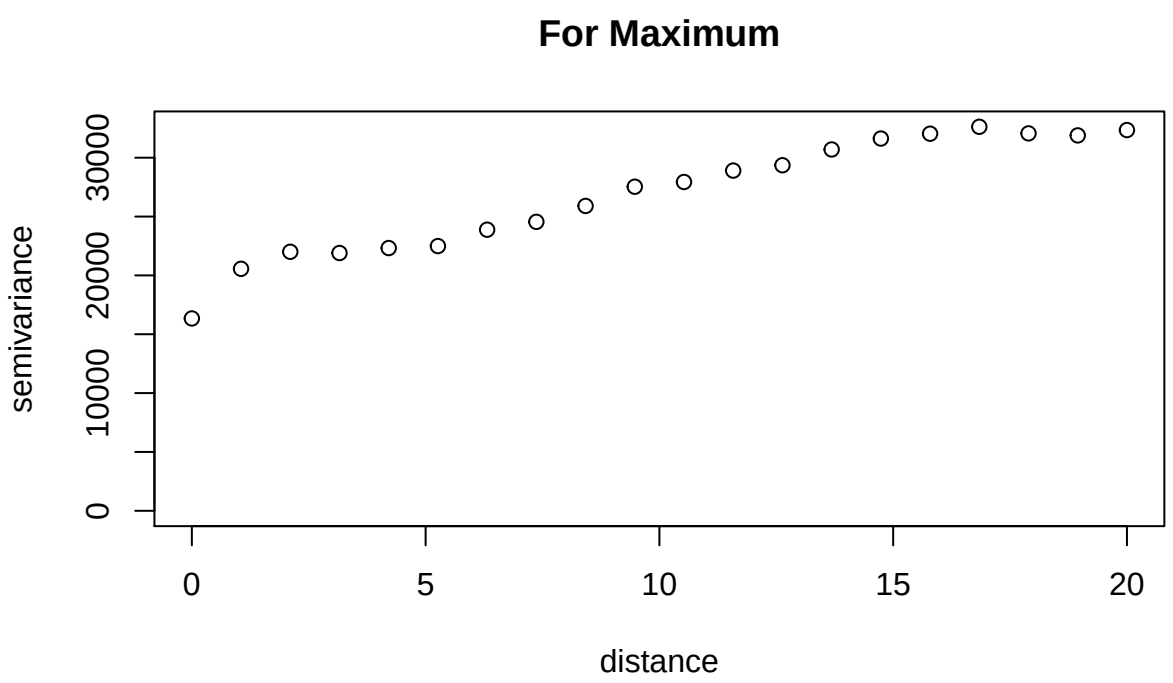


The residuals plots appear to be spatially correlated but look non-stationary and hence non-isotropic. However, the residual plot for Proportion can be argued to be somewhat stationary and isotropic if we kind of leave-off the coastal regions. This is one of the reasons why the variogram fit is better for the Proportion.

3)Emprical variograms:







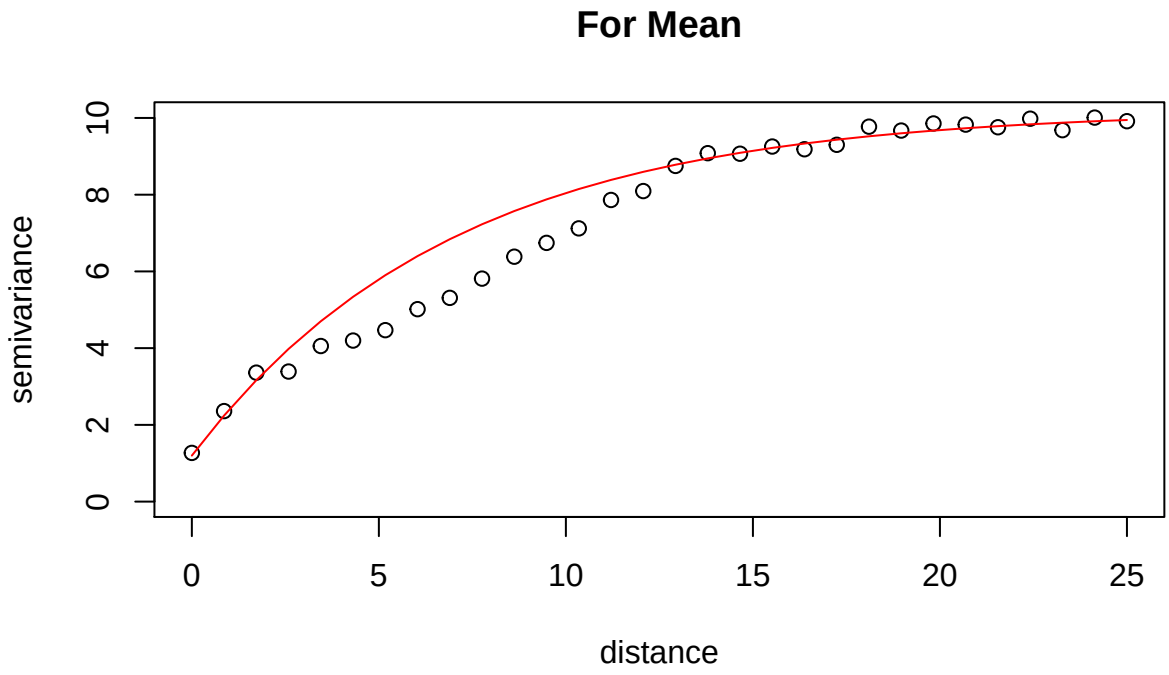
There are nugget effects. Roughly the effective spatial ranges are:

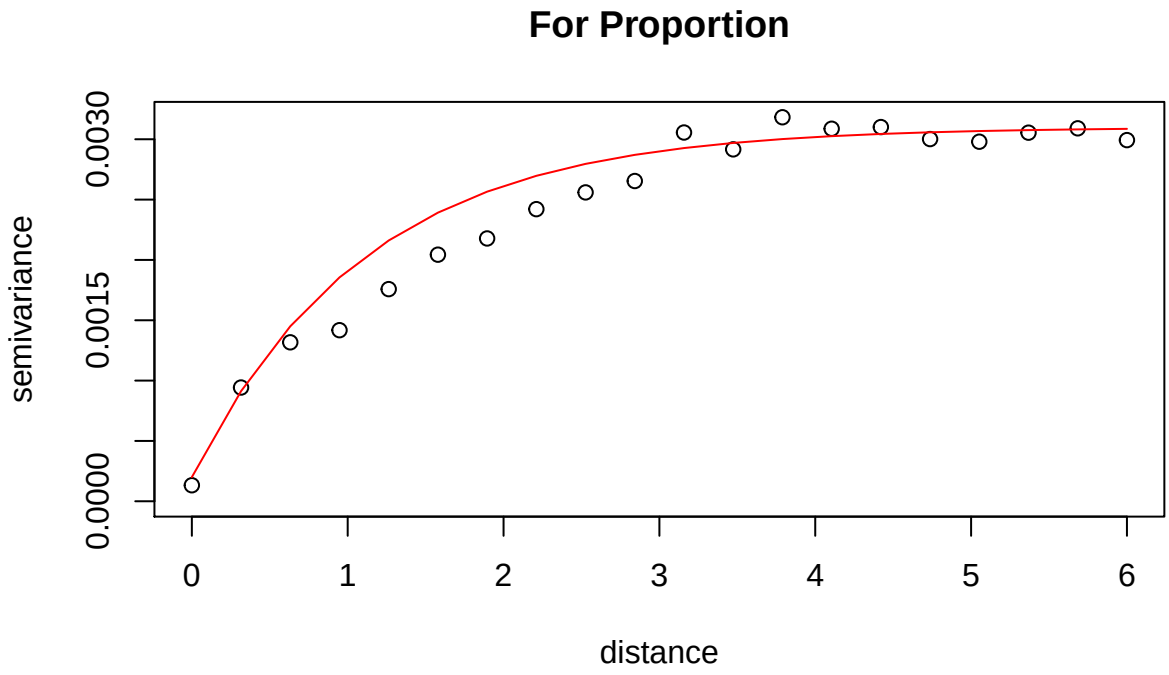
Mean ~ 18

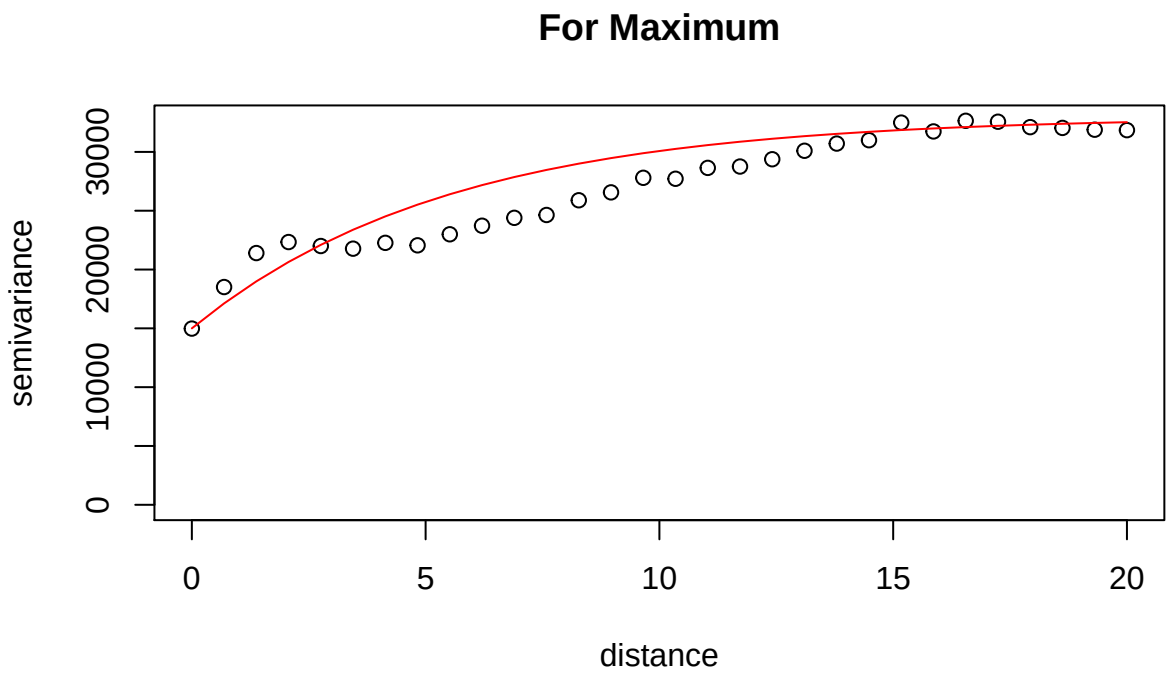
Proportion ~ 3.5

Maximum ~ 15

4)Fitted exponential variogram







Values of nugget, partial sill and range:

Mean:

Nugget: 1.2

Partial sill: 9

Range: 7

Proportion:

Nugget: 0.0002

Partial sill: 0.0029

Range: 1.12

Maximum:

Nugget: 15000

Partial sill: 18000

Range: 5.5

They fit pretty well, the fit for proportion being better.

Codes

```
library(dplyr)
library(ggplot2)
library(viridis)
library(gridExtra)
library(geoR)

load("C:/Users/acer/Downloads/USHCNprcpSetup (2).RData")

#Years:2000-2009
PRCP_req <- PRCP[((years>=2000)&(years<=2009)),]


#Checking for sites having obs>=3000
check_obs <- apply(PRCP_req,2,function(x){
  sum(!is.na(x))>=3000
})

#Cleaned data
PRCP_obs <- PRCP_req[,check_obs]


mean_dy <- apply(PRCP_obs,2,mean,na.rm=T)
prop_dy <- apply(PRCP_obs,2,function(x){
  mean(x>0,na.rm=T)
})
max_dy <- apply(PRCP_obs,2,max,na.rm=T)

#data frame for ggplots
map.dat <- data.frame(Lon=lon.lat[check_obs,1],
                      Lat=lon.lat[check_obs,2],
                      Mean=round(mean_dy,3),
                      Prop=round(prop_dy,3),
                      Max=max_dy)

#observed summary statistics maps

map_mean <- ggplot(data = map.dat, aes(Lon, Lat)) + borders("state",col="blue")+
  geom_point(aes(color = Mean))+
  scale_colour_gradientn(colours = viridis(10))+
  coord_map(projection = "albers",lat0 = 39, lat1 = 45)+
  xlab("Longitude W") + ylab("Latitude N")+
  theme(plot.title = element_text(hjust=0.5))+
  ggtitle("Observed Mean")

map_prop <- ggplot(data = map.dat, aes(Lon, Lat)) + borders("state",col="blue")+
  geom_point(aes(color = Prop))+
  scale_colour_gradientn(colours = viridis(10))+
  coord_map(projection = "albers",lat0 = 39, lat1 = 45)+
  xlab("Longitude W") + ylab("Latitude N")+
  theme(plot.title = element_text(hjust=0.5))+
  ggtitle("Observed Proportion")
```

```r
map_max <- ggplot(data = map.dat, aes(Lon, Lat)) + borders("state",col="blue")+
  geom_point(aes(color = Max))+
  scale_colour_gradientn(colours = viridis(10))+
  coord_map(projection = "albers",lat0 = 39, lat1 = 45)+
  xlab("Longitude W") + ylab("Latitude N")+
  theme(plot.title = element_text(hjust=0.5))+
  ggtitle("Observed Maximum")

map_mean
map_prop
map_max

#OLS

lon    <- map.dat$Lon
lat    <- map.dat$Lat
lon2   <- lon^2
lat2   <- lat^2
lonlat <- lon*lat

fit_mean <- lm(map.dat$Mean~lon+lat+lon2+lat2+lonlat)$fitted.values
res_mean <- map.dat$Mean-fit_mean
fit_prop <- lm(map.dat$Prop~lon+lat+lon2+lat2+lonlat)$fitted.values
res_prop <- map.dat$Prop-fit_prop
fit_max <- lm(map.dat$Max~lon+lat+lon2+lat2+lonlat)$fitted.values
res_max <- map.dat$Max-fit_max

#fitted values
df1 <- data.frame(long=lon,lat=lat,Fit=fit_mean)
mean_fit <- ggplot(df1, aes(long, lat)) +
  borders("state") +
  geom_point(aes(colour = Fit)) +
  scale_colour_gradientn(colours = viridis(10)) +
  coord_map(projection = "albers", lat0 = 39, lat1 = 45) +
  theme(plot.title = element_text(hjust=0.5))+
  xlab("Longitude W")+ylab("Latitude N")+labs(title="OLS Fit Mean")

df2 <- data.frame(long=lon,lat=lat,Fit=fit_prop)
prop_fit <- ggplot(df2, aes(long, lat)) +
  borders("state") +
  geom_point(aes(colour = Fit)) +
  scale_colour_gradientn(colours = viridis(10)) +
  coord_map(projection = "albers", lat0 = 39, lat1 = 45) +
  theme(plot.title = element_text(hjust=0.5))+
  xlab("Longitude W")+ylab("Latitude N")+labs(title="OLS Fit Proportion")

df3 <- data.frame(long=lon,lat=lat,Fit=fit_max)
max_fit <- ggplot(df3, aes(long, lat)) +
  borders("state") +
  geom_point(aes(colour = Fit)) +
  scale_colour_gradientn(colours = viridis(10)) +
  coord_map(projection = "albers", lat0 = 39, lat1 = 45) +
  theme(plot.title = element_text(hjust=0.5))+
```

```
  xlab("Longitude W")+ylab("Latitude N")+labs(title="OLS Fit Maximum")


#residuals
df4 <- data.frame(long=lon,lat=lat,Res=res_mean)
mean_res <- ggplot(df4, aes(long, lat)) +
  borders("state") +
  geom_point(aes(colour = Res)) +
  scale_colour_gradientn(colours = viridis(10)) +
  coord_map(projection = "albers", lat0 = 39, lat1 = 45) +
  theme(plot.title = element_text(hjust=0.5))+
  xlab("Longitude W")+ylab("Latitude N")+labs(title="Least Squares Residuals Mean")

df5 <- data.frame(long=lon,lat=lat,Res=res_prop)
prop_res <- ggplot(df5, aes(long, lat)) +
  borders("state") +
  geom_point(aes(colour = Res)) +
  scale_colour_gradientn(colours = viridis(10)) +
  coord_map(projection = "albers", lat0 = 39, lat1 = 45) +
  theme(plot.title = element_text(hjust=0.5))+
  xlab("Longitude W")+ylab("Latitude N")+labs(title="Least Squares Residuals Proportion")

df6 <- data.frame(long=lon,lat=lat,Res=res_max)
max_res <- ggplot(df6, aes(long, lat)) +
  borders("state") +
  geom_point(aes(colour = Res)) +
  scale_colour_gradientn(colours = viridis(10)) +
  coord_map(projection = "albers", lat0 = 39, lat1 = 45) +
  theme(plot.title = element_text(hjust=0.5))+
  xlab("Longitude W")+ylab("Latitude N")+labs(title="Least Squares Residuals Maximum")

grid.arrange(mean_fit,mean_res,nrow=1)
grid.arrange(prop_fit,prop_res,nrow=1)
grid.arrange(max_fit,max_res,nrow=1)

#Variograms
vgmean <- variog(coords = map.dat[,1:2], data = res_mean, uvec = seq(0,25,length=30),messages = F)
d=seq(0,25,length=30)
sig2 <- 9
tau2 <- 1.2
rho  <- 7
vg_fitted1 <- sig2 + tau2 - sig2*exp(-d/rho)
plot(vgmean,main="For Mean")
lines(d,vg_fitted1,col="red")



vgprop <- variog(coords = map.dat[,1:2], data = res_prop, uvec = seq(0,6,length=20),messages = F)
d=seq(0,6,length=20)
sig2 <- 0.0029
tau2 <- 0.0002
rho  <- 1.12
vg_fitted1 <- sig2 + tau2 - sig2*exp(-d/rho)
```

```r
plot(vgprop, main="For Proportion")
lines(d,vg_fitted1,col="red")


vgmax <- variog(coords = map.dat[,1:2], data = res_max, uvec = seq(0,20,length=30),messages = F)
d=seq(0,20,length=30)
sig2 <- 18000
tau2 <- 15000
rho  <- 5.5
vg_fitted1 <- sig2 + tau2 - sig2*exp(-d/rho)
plot(vgmax, main="For Maximum")
lines(d,vg_fitted1,col="red")
```