STAT 506 Homework 1

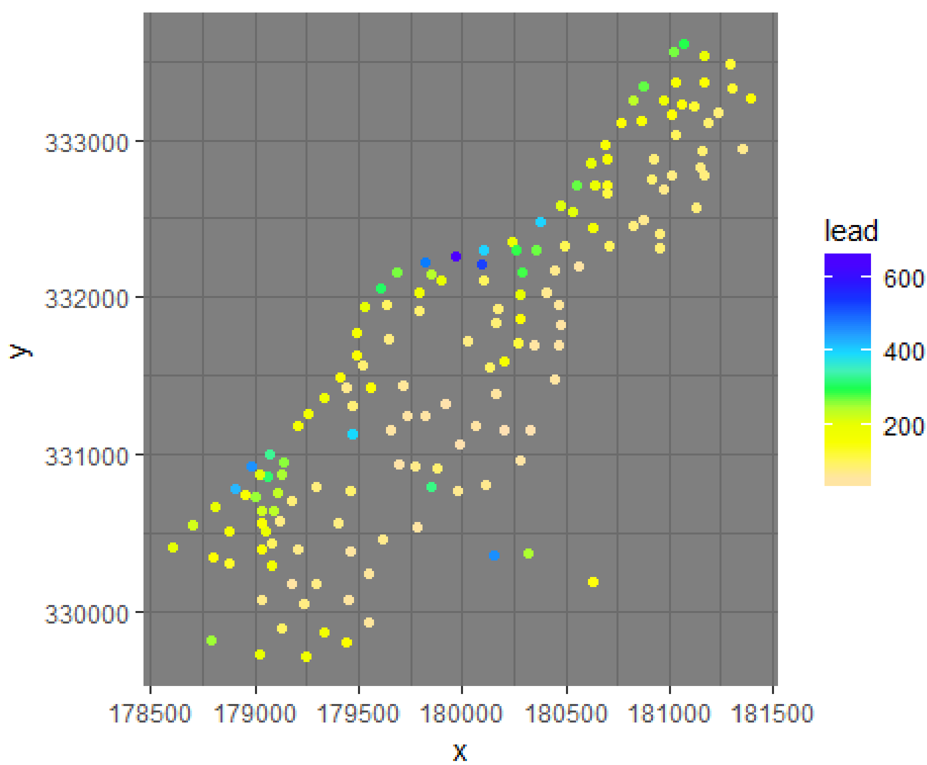
Steve Harms

January 24, 2018

4)

a)

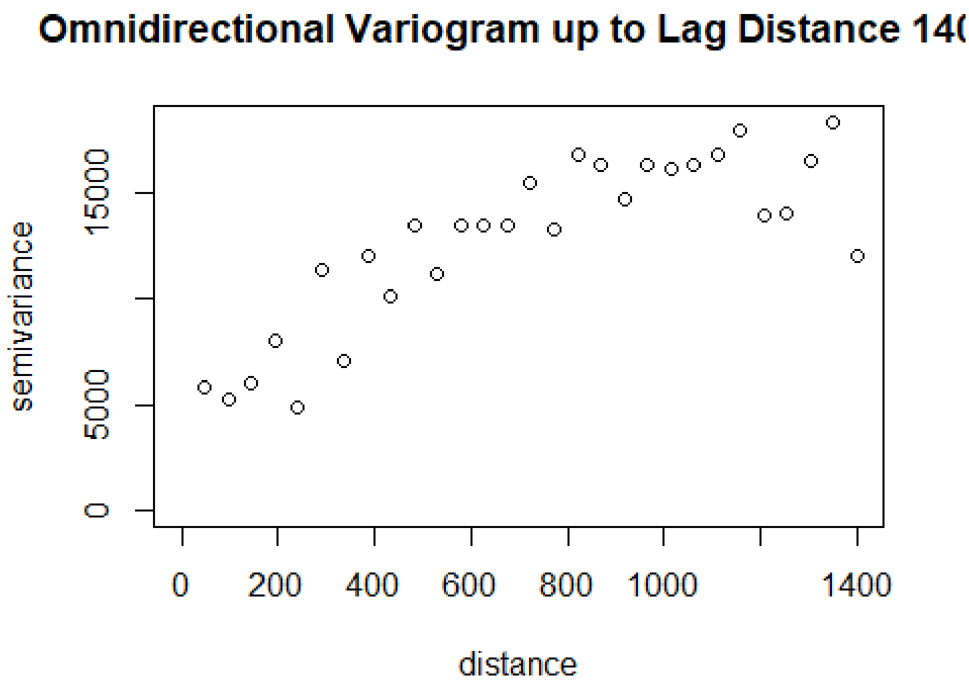
```
soil <- read.table(file = "soil.txt", header = T)
head(soil)
##        x      y lead zinc
## 1 181072 333611  299 1022
## 2 181025 333558  277 1141
## 3 181165 333537  199  640
## 4 181298 333484  116  257
## 5 181307 333330  117  269
## 6 181390 333260  137  281
#quick summary of the data
summary(soil)
##        x                y                lead            zinc       
##  Min.   :178605   Min.   :329714   Min.   : 37.0   Min.   : 113.0  
##  1st Qu.:179371   1st Qu.:330762   1st Qu.: 72.5   1st Qu.: 198.0  
##  Median :179991   Median :331633   Median :123.0   Median : 326.0  
##  Mean   :180005   Mean   :331635   Mean   :153.4   Mean   : 469.7  
##  3rd Qu.:180630   3rd Qu.:332463   3rd Qu.:207.0   3rd Qu.: 674.5  
##  Max.   :181390   Max.   :333611   Max.   :654.0   Max.   :1839.0  
#plot of the geodata
g <- ggplot(data = soil, aes(x=x, y=y))
h <- g + geom_point(aes(colour = lead)) + scale_color_gradientn(colours =rev(topo.colors(8))) +
  theme_dark()
h
```



b)

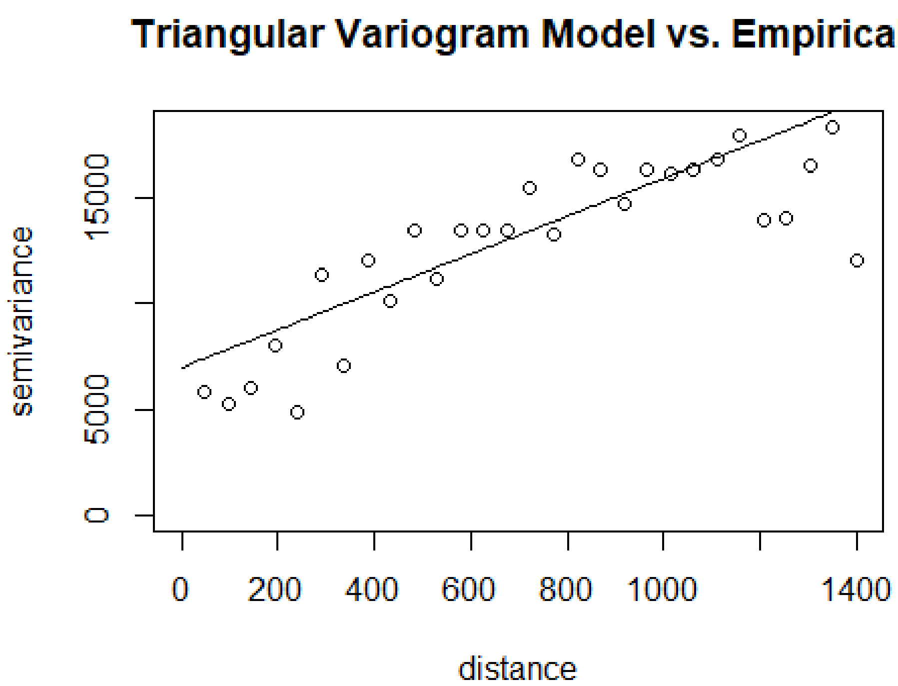
```
#convert to geodata format
soilgeo <- as.geodata(soil, coords.col = 1:2, data.col = 3, covar.col = 4)
#compute empirical variogram and plot
```

```
#visually, nugget is 5000, sill is 17500, range ~ 800, partial sill 17500-5000=12500
omndvar <- variog(soilgeo, option = "bin", uvec = seq(0, 1400, length = 30 ))
## variog: computing omnidirectional variogram
plot(omndvar, main = "Omnidirectional Variogram up to Lag Distance 1400")
```

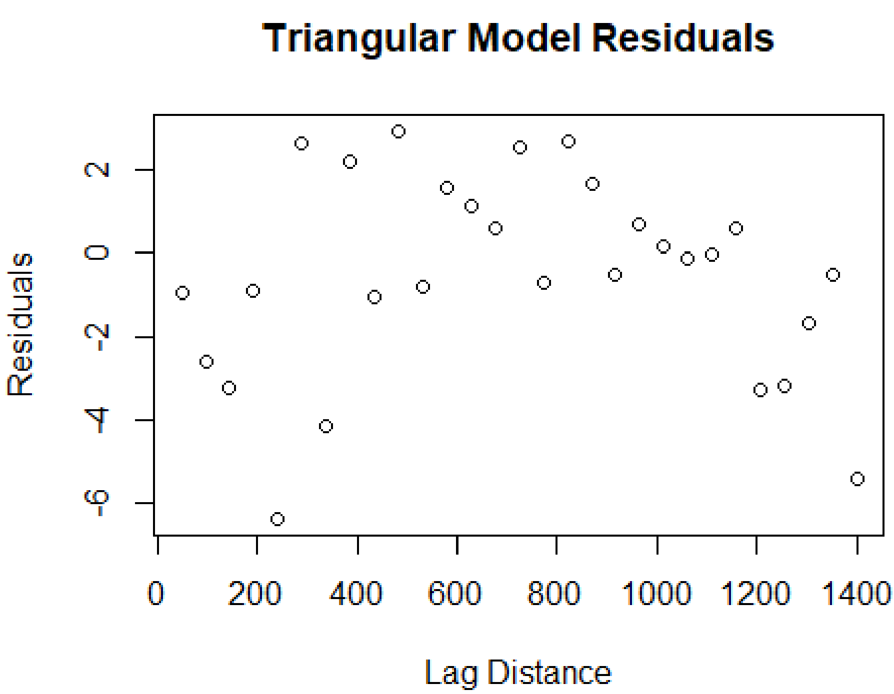


c)

```
##Triangular Model
#Compute variogram
#Need to use linear model with sigma^2/a instead of sigma^2,
#and then range becomes A/A = 1 using geoR's specification
vario.tri <- variofit(omndvar, ini.cov.pars = c(12500/800, 1), cov.model = "linear", weights =
"cressie")
## variofit: covariance model used is linear
## variofit: weights used: cressie
## variofit: minimisation function used: optim
## Warning in variofit(omndvar, ini.cov.pars = c(12500/800, 1), cov.model =
## "linear", : unreasonable initial value for sigmasq + nugget (too low)
vario.tri
## variofit: model parameters estimated by WLS (weighted least squares):
## covariance model is: linear
## parameter estimates:
##     tausq   sigmasq       phi
## 6992.4645    8.9583    1.0000
## Practical Range with cor=0.05 for asymptotic range: Inf
##
## variofit: minimised weighted sum of squares = 175.2857
#Fit values and calculate residuals
fit.tri = vario.tri$nugget + vario.tri$cov.pars[1] * (omndvar$u/vario.tri$cov.pars[2])
res.tri = sqrt(omndvar$n) * ((omndvar$v - fit.tri)/fit.tri)
#plot of Empirical Variogram vs. Fitted Model
plot(omndvar, main = "Triangular Variogram Model vs. Empirical") + lines(vario.tri)
```

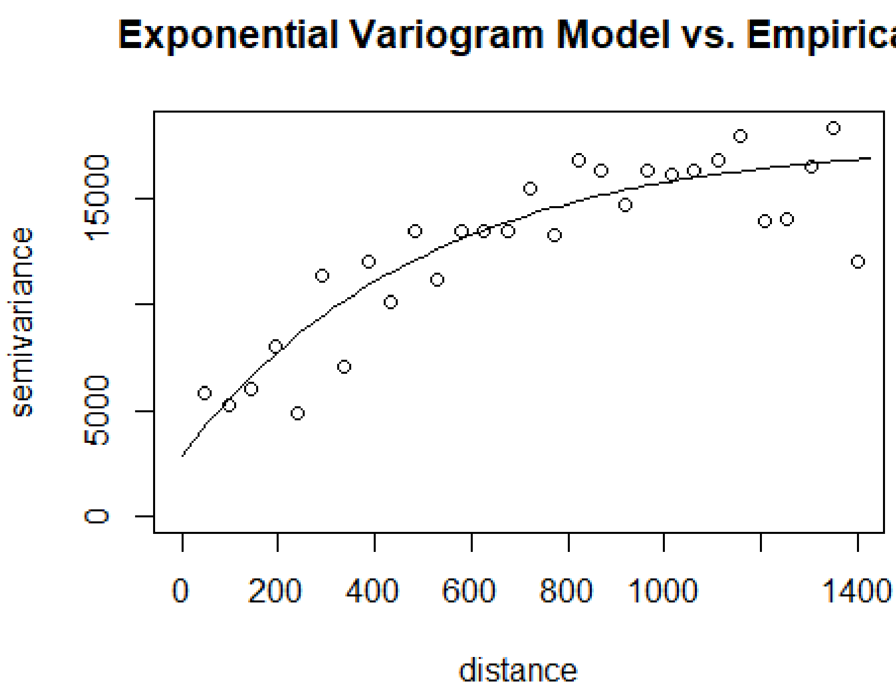



```
## integer(0)
#Plot of residuals
plot(y = res.tri, main = "Triangular Model Residuals", ylab = "Residuals", xlab= "Lag Distance",
x=omndvar$u)
```

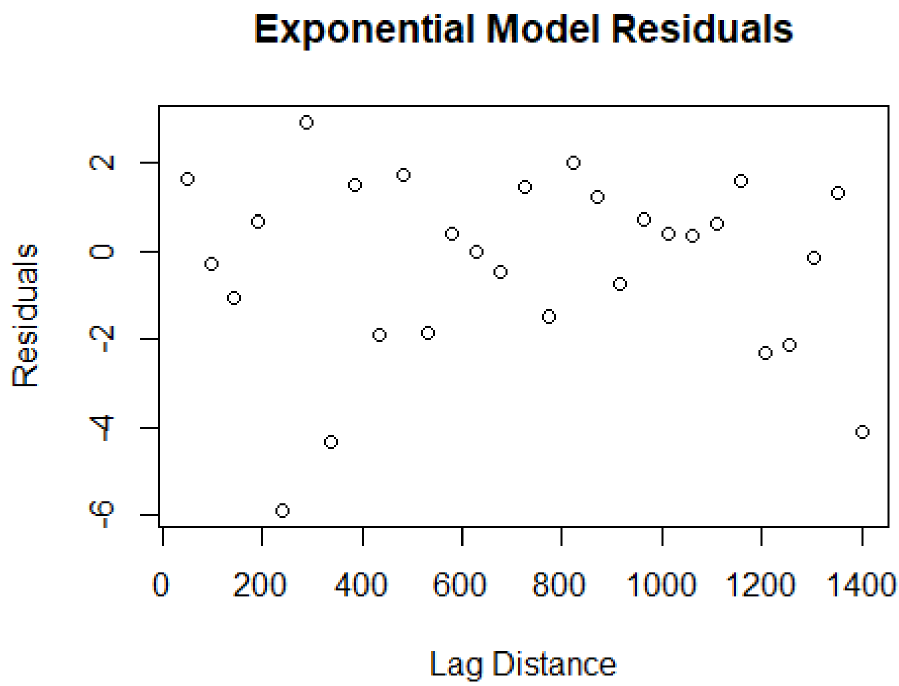



```
##Exponential Model
#Compute Variogram
vario.exp <- variofit(omndvar, ini.cov.pars = c(12500, 800), cov.model = "exp", weights =
"cressie", nugget = 5000)
## variofit: covariance model used is exponential
## variofit: weights used: cressie
## variofit: minimisation function used: optim
```

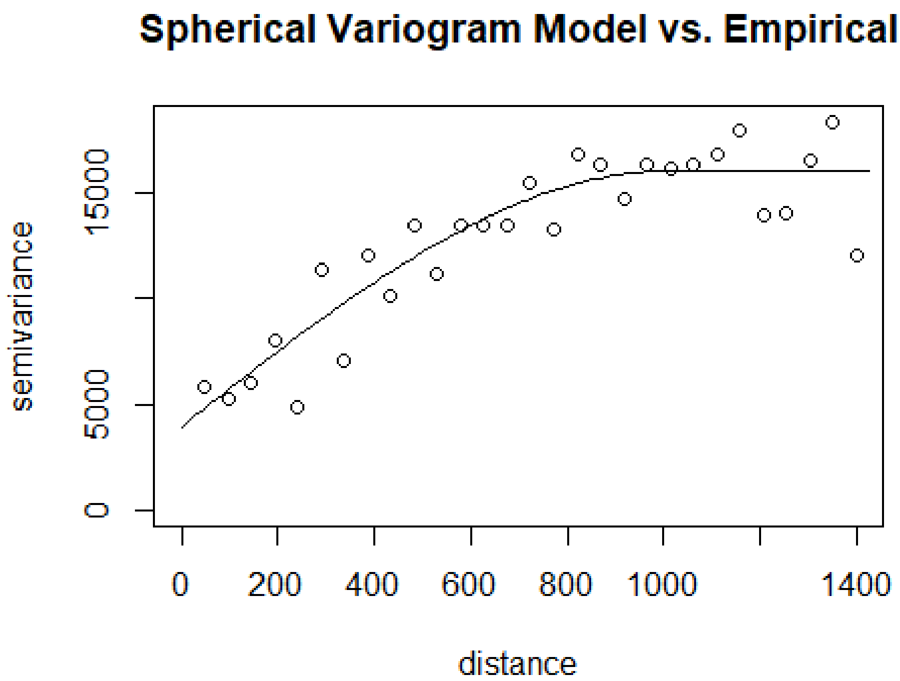
```
vario.exp
## variofit: model parameters estimated by WLS (weighted least squares):
## covariance model is: exponential
## parameter estimates:
##       tausq    sigmasq        phi
##   2837.1802 14957.5009   498.8027
## Practical Range with cor=0.05 for asymptotic range: 1494.279
##
## variofit: minimised weighted sum of squares = 121.38
#Fit Values and calculate residuals
fit.exp = vario.exp$nugget + vario.exp$cov.pars[1] * (1 - exp(-omndvar$u/vario.exp$cov.pars[2]))
res.exp = sqrt(omndvar$n) * ((omndvar$v - fit.exp)/fit.exp)
#plot of Empirical Variogram vs. Fitted Model
plot(omndvar, main = "Exponential Variogram Model vs. Empirical") +lines(vario.exp)
```



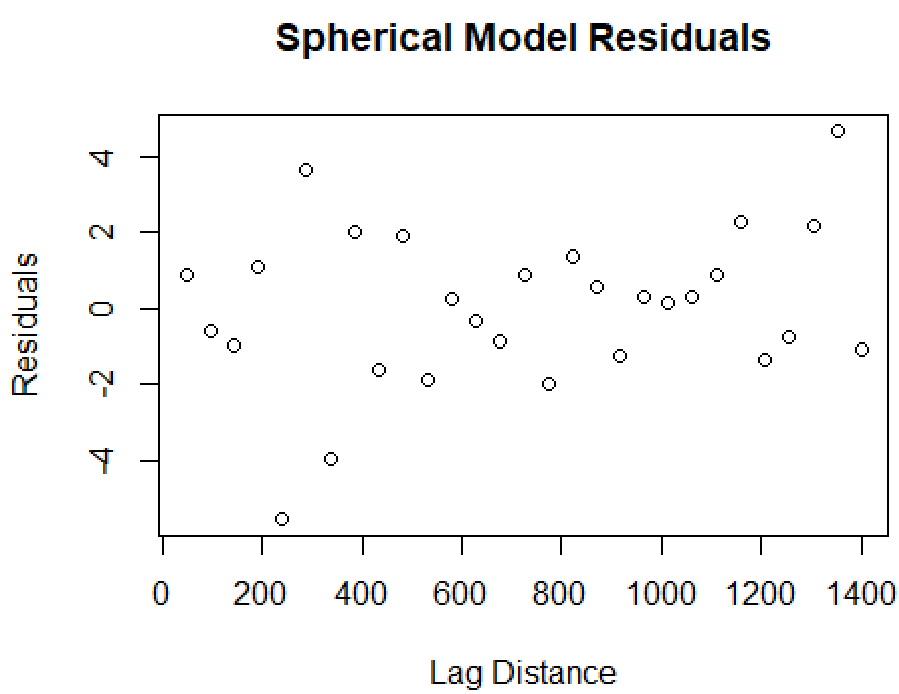

```
## integer(0)
#Plot of residuals
plot(y=res.exp, main = "Exponential Model Residuals",ylab = "Residuals", xlab= "Lag Distance",
x=omndvar$u)
```




```
##Spherical Model
#Compute Variogram
vario.sph <- variofit(omndvar, ini.cov.pars = c(12500, 800), cov.model = "spherical", weights =
"cressie", nugget = 5000)
## variofit: covariance model used is spherical
## variofit: weights used: cressie
## variofit: minimisation function used: optim
vario.sph
## variofit: model parameters estimated by WLS (weighted least squares):
## covariance model is: spherical
## parameter estimates:
##     tausq   sigmasq       phi
##  3959.946 12114.263  1010.333
## Practical Range with cor=0.05 for asymptotic range: 1010.333
##
## variofit: minimised weighted sum of squares = 114.6208
#Fit Values and calculate residuals
fit.sph = vario.sph$nugget + vario.sph$cov.pars[1] * (1.5*(omndvar$u/vario.sph$cov.pars[2])-
.5*(omndvar$u/vario.sph$cov.pars[2])^3)
res.sph = sqrt(omndvar$n) * ((omndvar$v - fit.sph)/fit.sph)
#plot of Empirical Variogram vs. Fitted Model
plot(omndvar, main = "Spherical Variogram Model vs. Empirical") + lines(vario.sph)
```

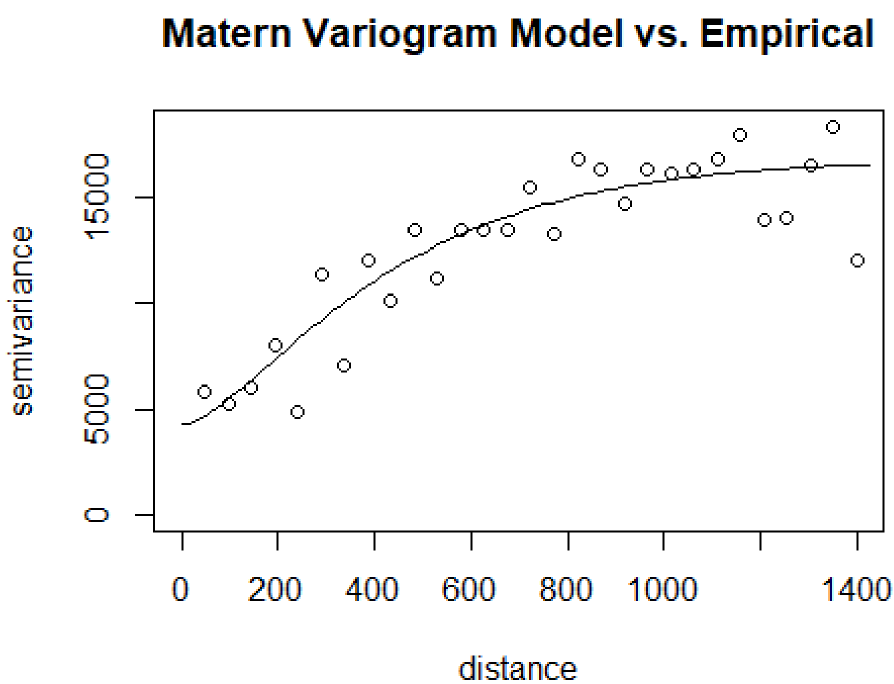



```
## integer(0)
#Plot of residuals
plot(y = res.sph, main = "Spherical Model Residuals", ylab = "Residuals", xlab= "Lag Distance",
x=omndvar$u)
```

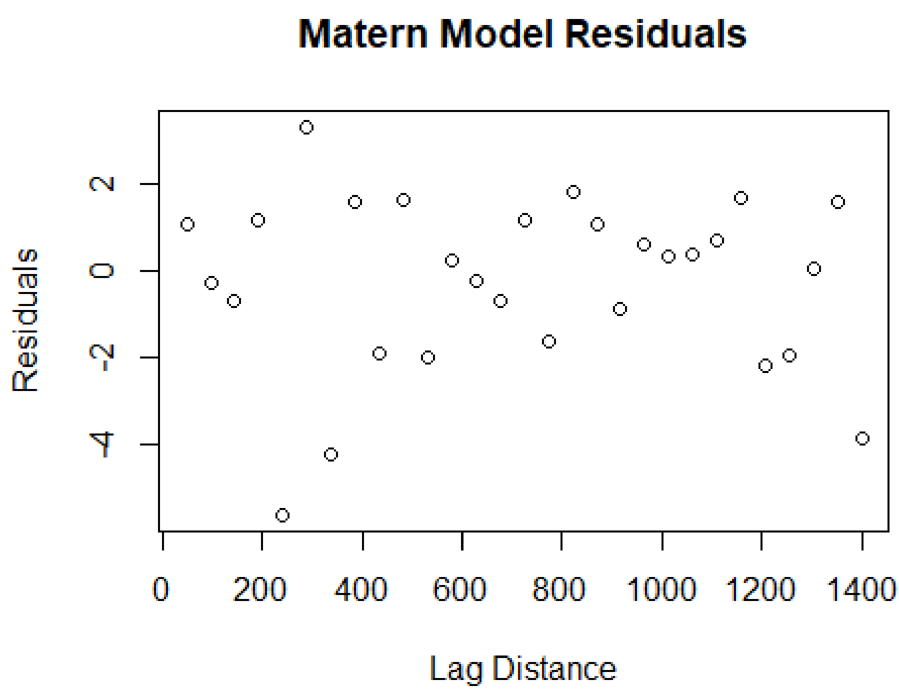



```
##Matern Model
#Compute Variogram
vario.matern <- variofit(omndvar, ini.cov.pars = c(12500, 800), cov.model = "matern", weights =
"cressie", kappa = 1)
```

```
## variofit: covariance model used is matern
## variofit: weights used: cressie
## variofit: minimisation function used: optim
vario.matern
## variofit: model parameters estimated by WLS (weighted least squares):
## covariance model is: matern with fixed kappa = 1
## parameter estimates:
##      tausq    sigmasq        phi
##  4261.6368 12630.1329   293.6961
## Practical Range with cor=0.05 for asymptotic range: 1174.35
##
## variofit: minimised weighted sum of squares = 116.6109
#Fit Values and calculate residuals
fit.mtr = vario.matern$nugget + vario.matern$cov.pars[1] * (1 - (1/gamma(1)) * (abs(omndvar$u) /
vario.matern$cov.pars[2]) * besselK((abs(omndvar$u) / vario.matern$cov.pars[2]), nu = 1))
res.mtr = sqrt(omndvar$n) * ((omndvar$v - fit.mtr)/fit.mtr)
#plot of Empirical Variogram vs. Fitted Model
plot(omndvar, main = "Matern Variogram Model vs. Empirical") + lines(vario.matern)
```



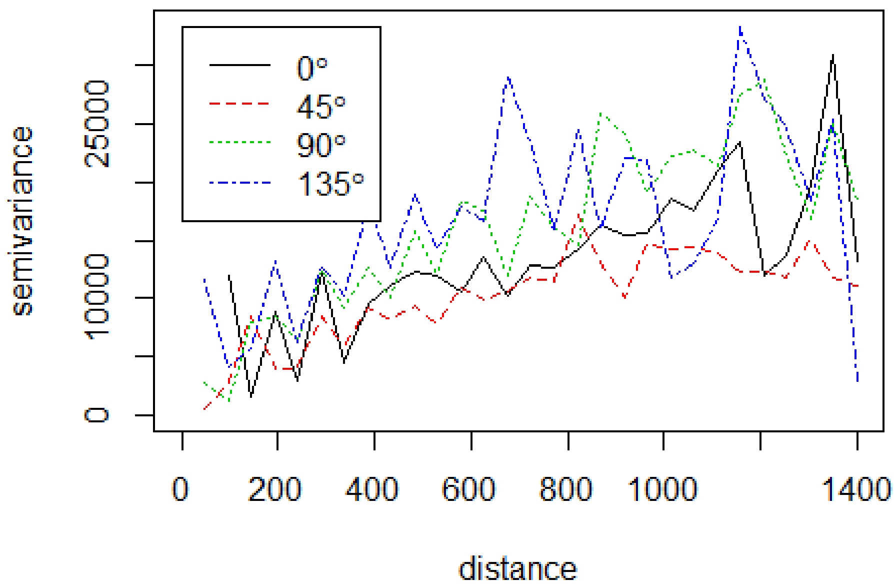

```
## integer(0)
#Plot of residuals
plot(y = res.mtr, main = "Matern Model Residuals", ylab = "Residuals", xlab= "Lag Distance",
x=omndvar$u)
```



The triangular model is clearly a poor fit, but the others (Exponential, Spherical, Matern) all appear to be reasonable fits for the data based on the fitted model and residual plots. Spherical has the smallest weighted sum of squares and also the most reasonable looking residual plot, so by that primitive reasoning I would say spherical model is the best.

d)

```
#Compute Direction Variograms
var.direct <- variog4(soilgeo, uvec = seq(0, 1400, length = 30), direction =
c(0,pi/4,pi/2,3*pi/4))
## variog: computing variogram for direction = 0 degrees (0 radians)
##         tolerance angle = 22.5 degrees (0.393 radians)
## variog: computing variogram for direction = 45 degrees (0.785 radians)
##         tolerance angle = 22.5 degrees (0.393 radians)
## variog: computing variogram for direction = 90 degrees (1.571 radians)
##         tolerance angle = 22.5 degrees (0.393 radians)
## variog: computing variogram for direction = 135 degrees (2.356 radians)
##         tolerance angle = 22.5 degrees (0.393 radians)
## variog: computing omnidirectional variogram
#Plot Empirical Variograms
plot(var.direct)
```



There is no convincing visual evidence for anisotropy in any direction. All of the directional variograms appear to follow similar patterns, especially at short lag distances, which suggests that the variogram only depends on lag distance and not direction. Most of the difference is likely due to the nugget effect. IF there is anisotropy, it would likely be in the 3pi/4 direction as that one is visually the most different.