Case Study 2, Part 2

Peter Claussen

10/3/2017

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.3.2
```

```
library(rsm)
library(ncf)
library(ape)
```

```
## Warning: package 'ape' was built under R version 3.3.2
```

```
##
## Attaching package: 'ape'
```

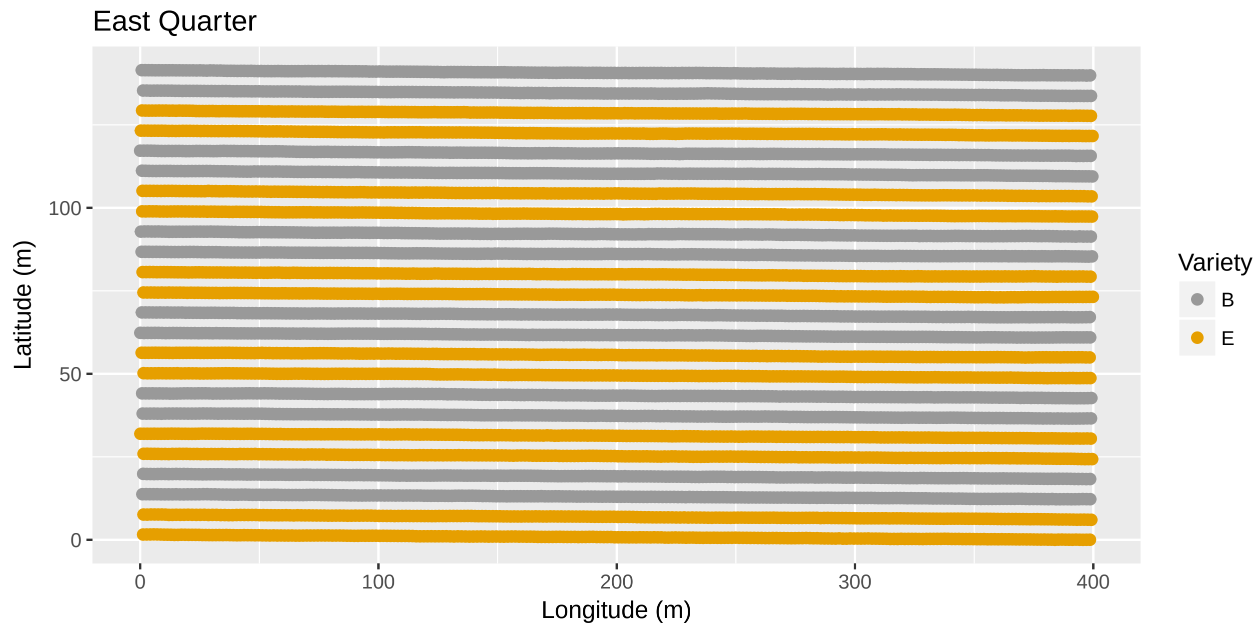
```
## The following object is masked from 'package:ncf':
##
##     mantel.test
```

```
cbPalette <- c("#999999", "#E69F00", "#56B4E9", "#009E73", "#0072B2", "#D55E00", "#F0E442","#CC79A7","#(
```

East Quarter

```
load(file="Strips.Rda")
```

```
ggplot(EastQuarter.dat, aes(Easting,Northing)) +
geom_point(aes(colour = Product),size=2) +
scale_colour_manual(values=cbPalette) +
labs(colour = "Variety", x="Longitude (m)", y="Latitude (m)", title = "East Quarter")
```



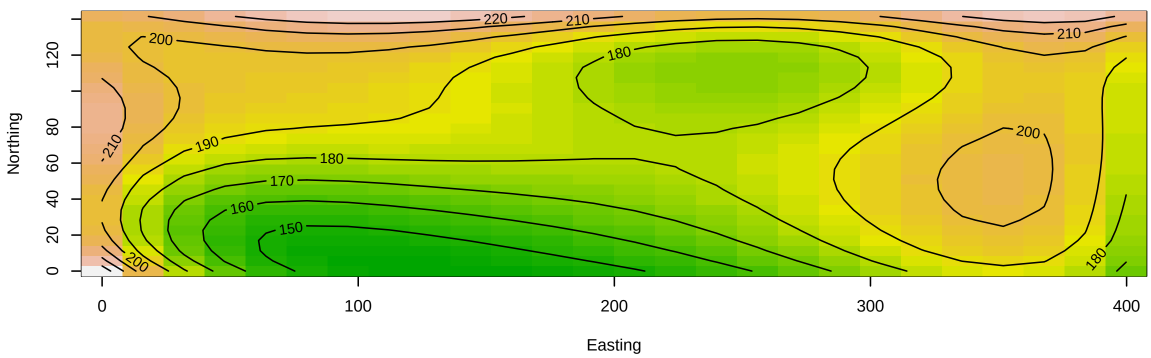
We'll want to look at measures of spatial correlation, but first let's try to determine a trend model for each. We aren't too interested in kriging, since we won't be trying to map yields to a uniform set of points.

Trend Surface

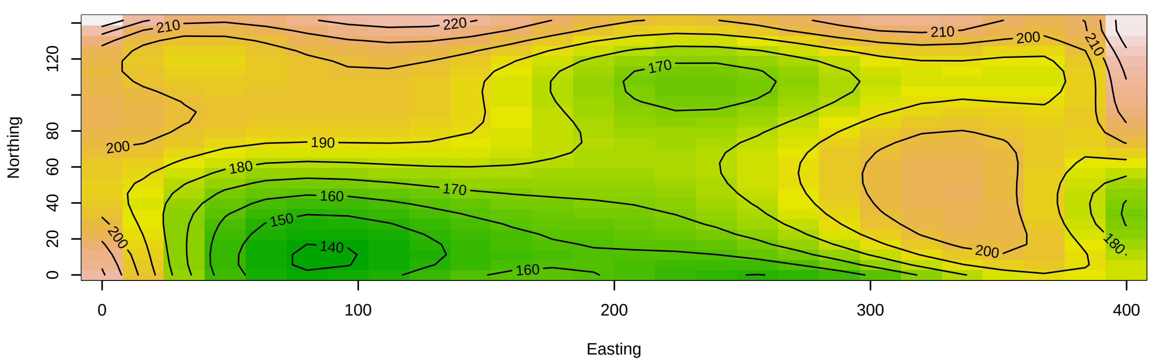
```
Yield5.lm <- lm(Yield ~ poly(Easting, Northing, degree=5), data=EastQuarter.dat)
Yield7.lm <- lm(Yield ~ poly(Easting, Northing, degree=7), data=EastQuarter.dat)
Yield9.lm <- lm(Yield ~ poly(Easting, Northing, degree=9), data=EastQuarter.dat)
Yield11.lm <- lm(Yield ~ poly(Easting, Northing, degree=11), data=EastQuarter.dat)
```

```
par(mfrow=c(4,1))
contour(Yield5.lm, Northing ~ Easting, image = TRUE,main="Poly 5")
contour(Yield7.lm, Northing ~ Easting, image = TRUE,main="Poly 7")
contour(Yield9.lm, Northing ~ Easting, image = TRUE,main="Poly 9")
contour(Yield11.lm, Northing ~ Easting, image = TRUE,main="Poly 11")
```

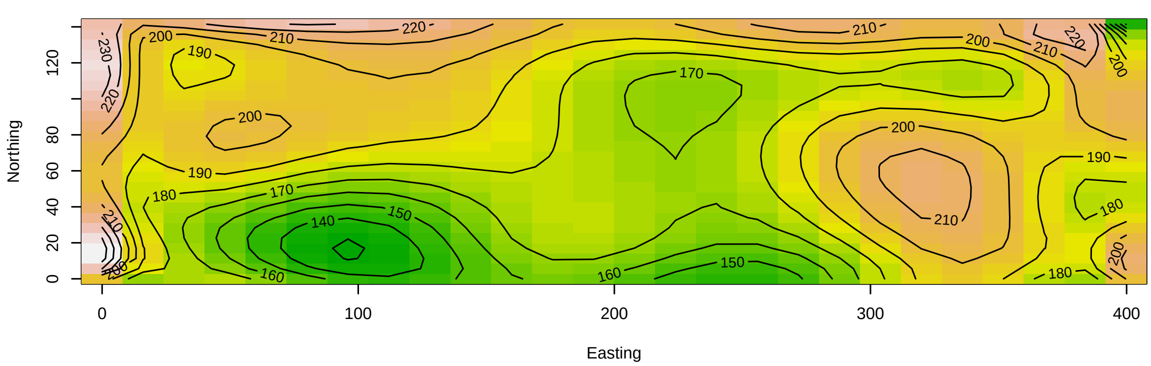
Poly 5



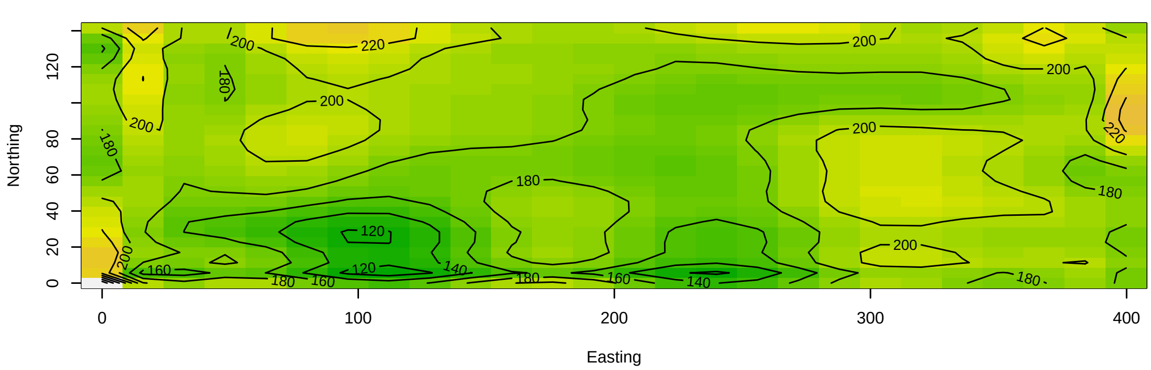
Poly 7



Poly 9



Poly 11



```r
par(mfrow=c(1,1))
```

I'm concerned that Poly 11 might be overfitting in the lower left corner, so we'll choose Poly 9 (note - we could continue with diagnostics of Local I to check that assertion).

```r
EastQuarter.dat$Yield9.resid <- Yield9.lm$residuals
```

Check for white noise

```r
Distance.mat <- as.matrix(dist(cbind(EastQuarter.dat$Easting, EastQuarter.dat$Northing)))
Distance.mat <- 1/Distance.mat
diag(Distance.mat) <- 0
```

```r
print(Moran9Yield <-Moran.I(EastQuarter.dat$Yield, Distance.mat))
```

```
## $observed
## [1] 0.1223093
##
## $expected
## [1] -0.0001570352
##
## $sd
## [1] 0.0003680123
##
## $p.value
## [1] 0
```

```r
print(Moran9Resid <-Moran.I(EastQuarter.dat$Yield9.resid, Distance.mat))
```

```
## $observed
## [1] 0.03937349
##
## $expected
## [1] -0.0001570352
##
## $sd
## [1] 0.0003680108
##
## $p.value
## [1] 0
```

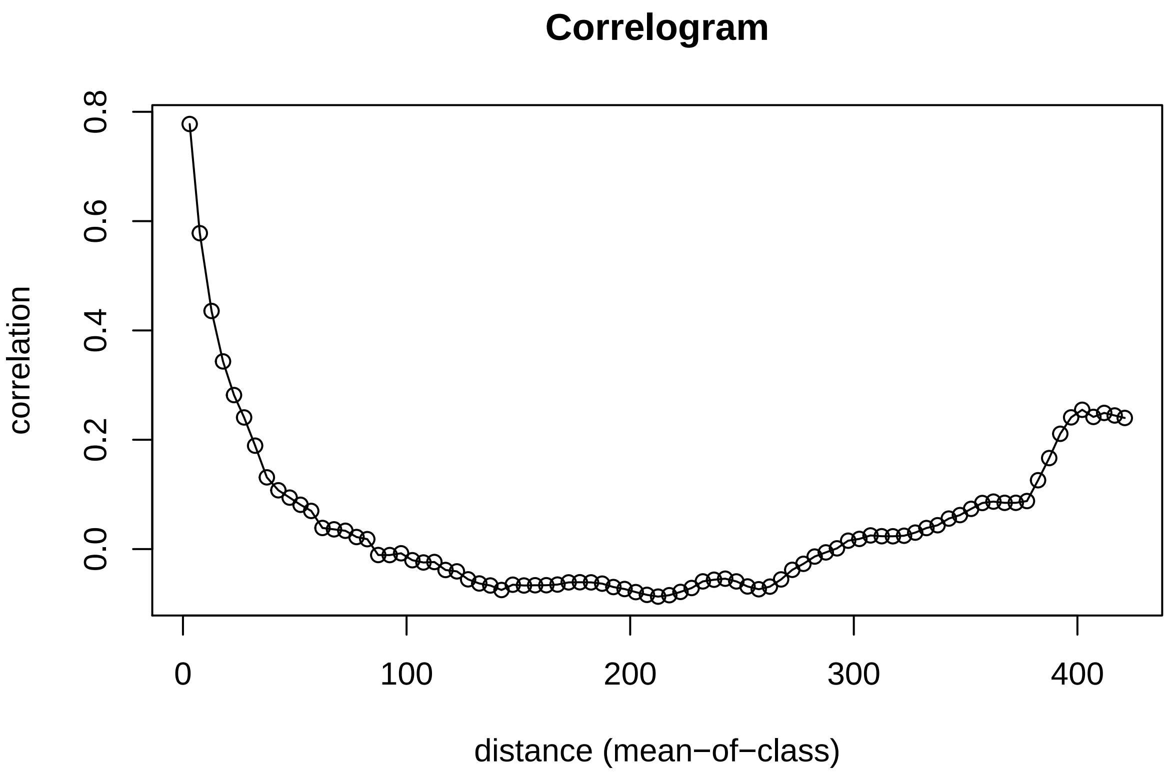
Local Correlation

Resampling can take a very long time, so I'll use a flag to control whether we resample or simply plot local measures. Resampling is needed for p-values; I'm not concerned about p-values for the points in the correlogram.

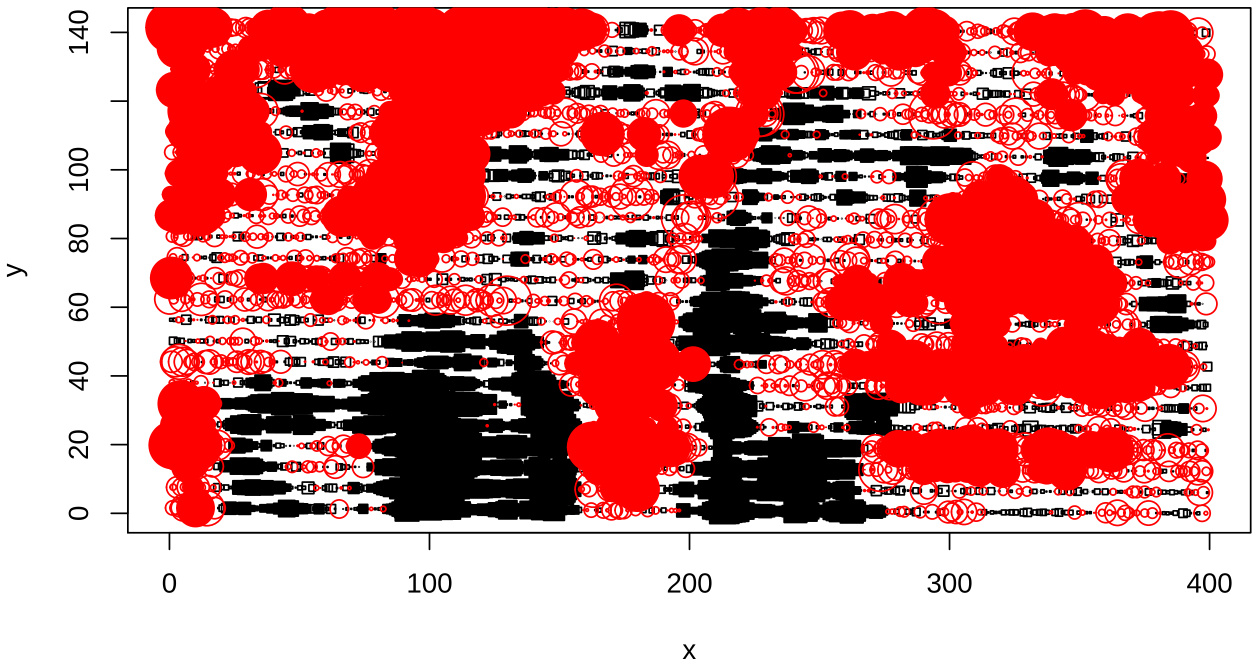
```r
resample = 100
```

Correlation in Yield

```
Yield.clg <- correlog(EastQuarter.dat$Easting, EastQuarter.dat$Northing, EastQuarter.dat$Yield,
                      increment=5, resamp=0, quiet=TRUE)
plot(Yield.clg)
```

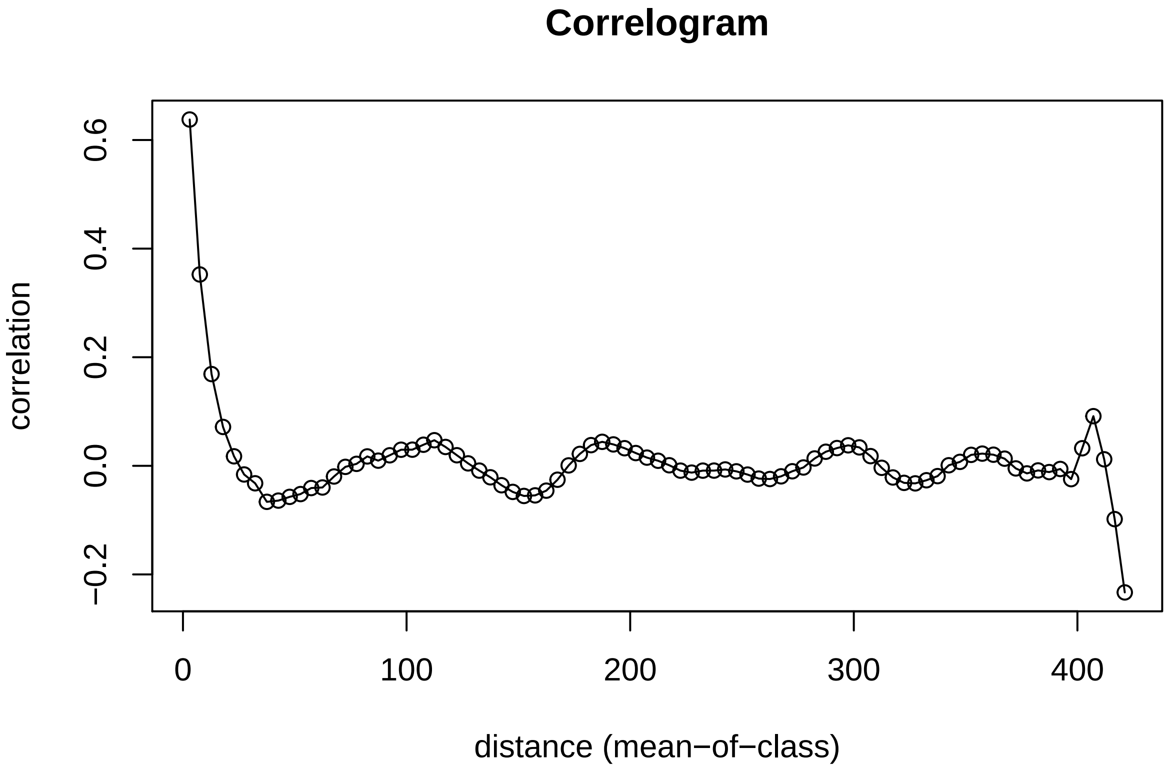



```
Yield9.lisa <- lisa(EastQuarter.dat$Easting, EastQuarter.dat$Northing, EastQuarter.dat$Yield,
                    neigh=10, resamp=resample, quiet=TRUE)
plot.lisa(Yield9.lisa, negh.mean=FALSE)
```

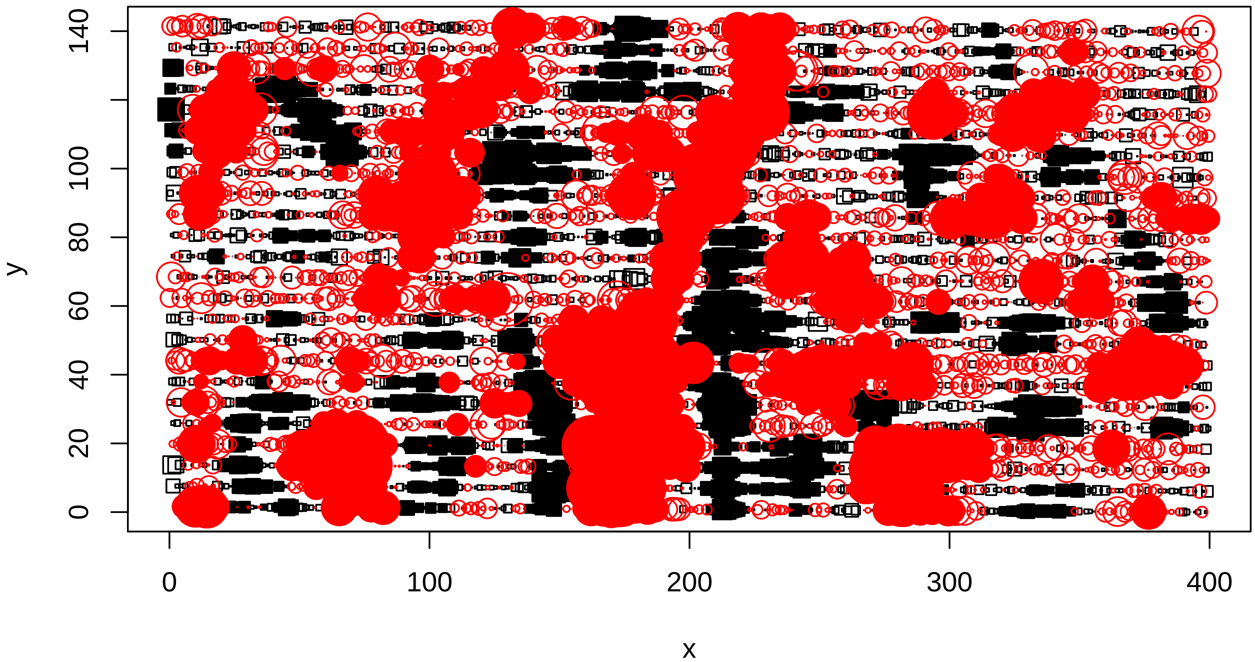


Correlation in Yield Residuals

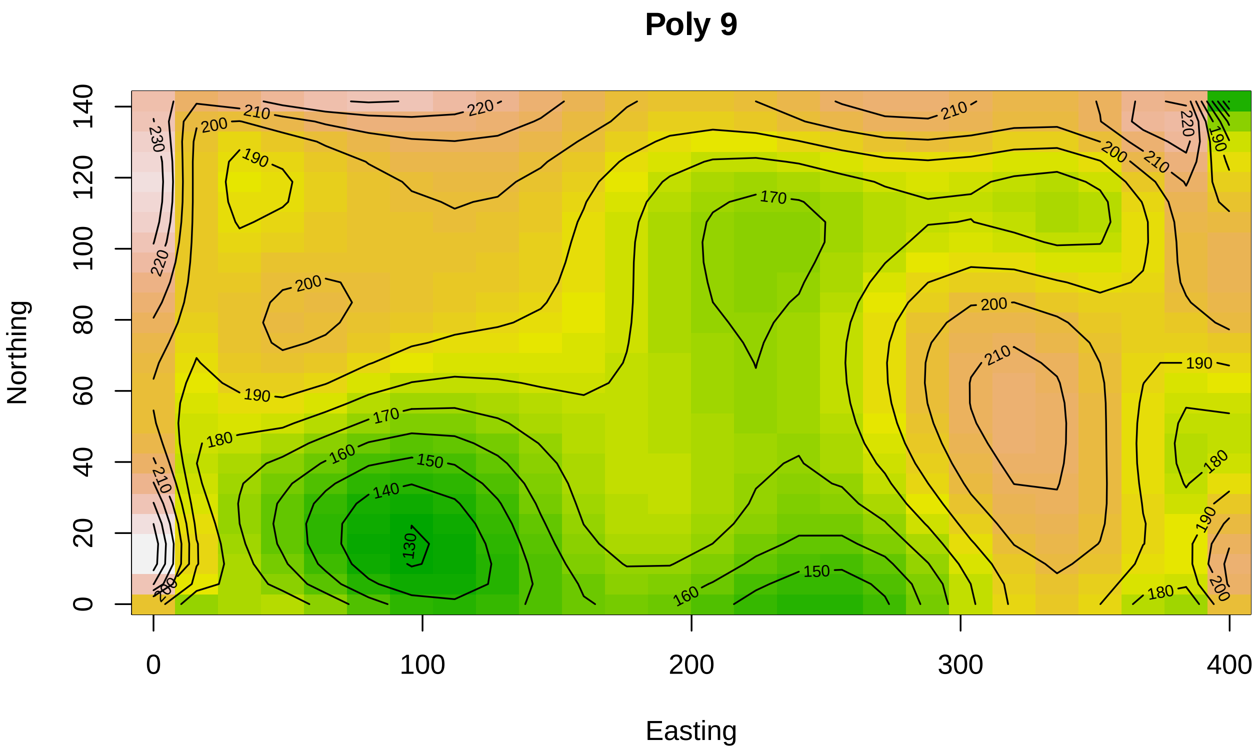
```
Yield9.resid.clg <- correlog(EastQuarter.dat$Easting, EastQuarter.dat$Northing, EastQuarter.dat$Yield9.r
                             increment=5, resamp=0, quiet=TRUE)
plot(Yield9.resid.clg)
```



```
Yield9.resid.lisa <- lisa(EastQuarter.dat$Easting, EastQuarter.dat$Northing, EastQuarter.dat$Yield9.resi
                          neigh=10, resamp=resample, quiet=TRUE)
plot.lisa(Yield9.resid.lisa, negh.mean=FALSE)
```



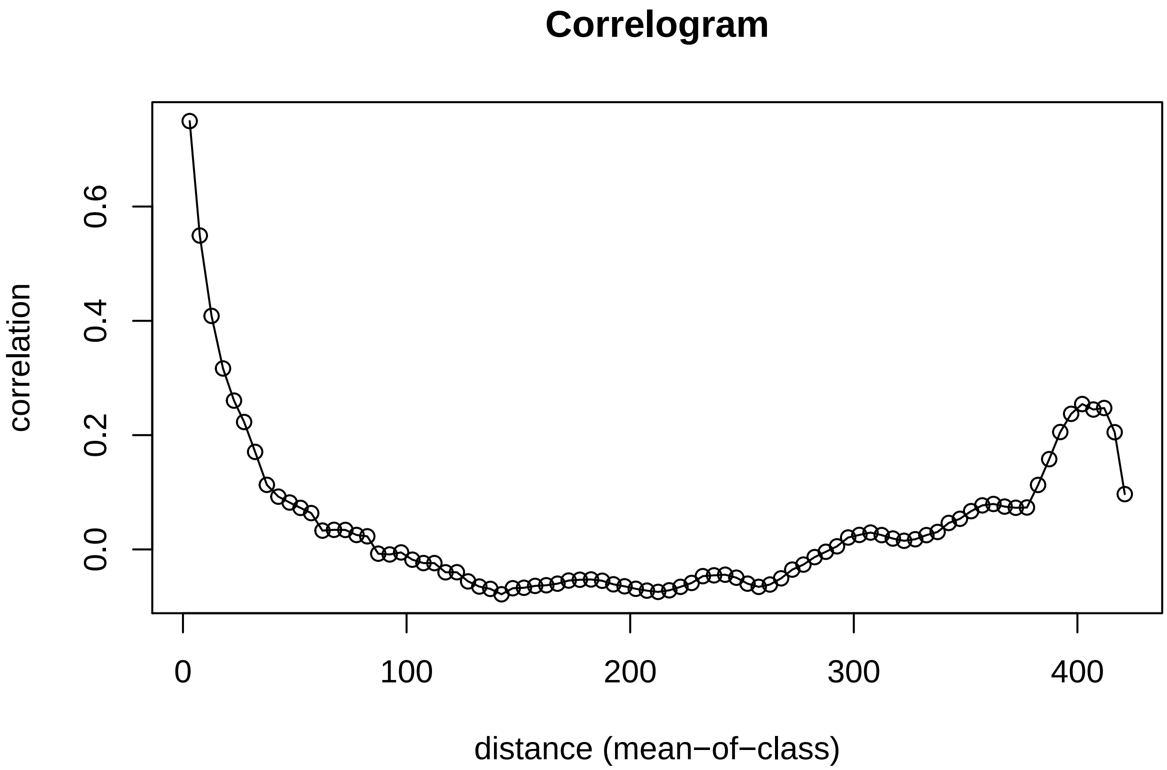

```
contour(Yield9.lm, Northing ~ Easting, image = TRUE,main="Poly 9")
```



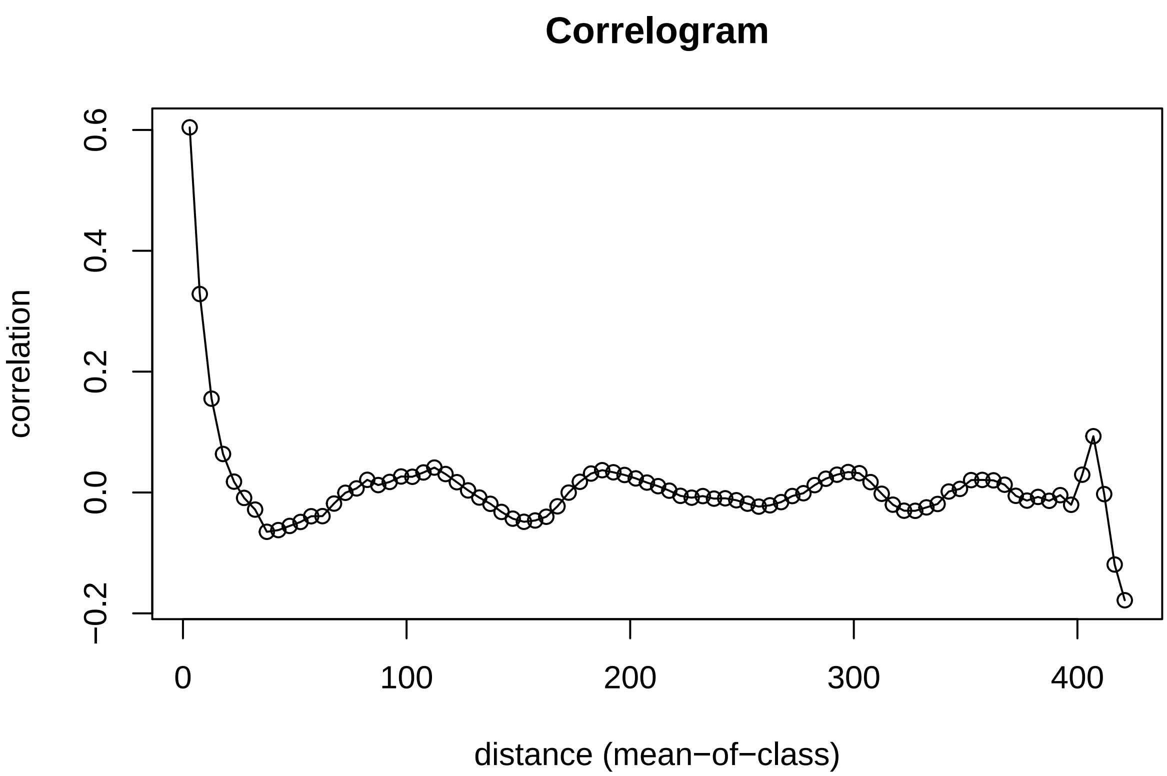
Correlation in Quantiles

```
Quantile9.lm <- lm(Quantile ~ poly(Easting, Northing, degree=9), data=EastQuarter.dat)
EastQuarter.dat$Quantile9.resid <- Quantile9.lm$residuals
```

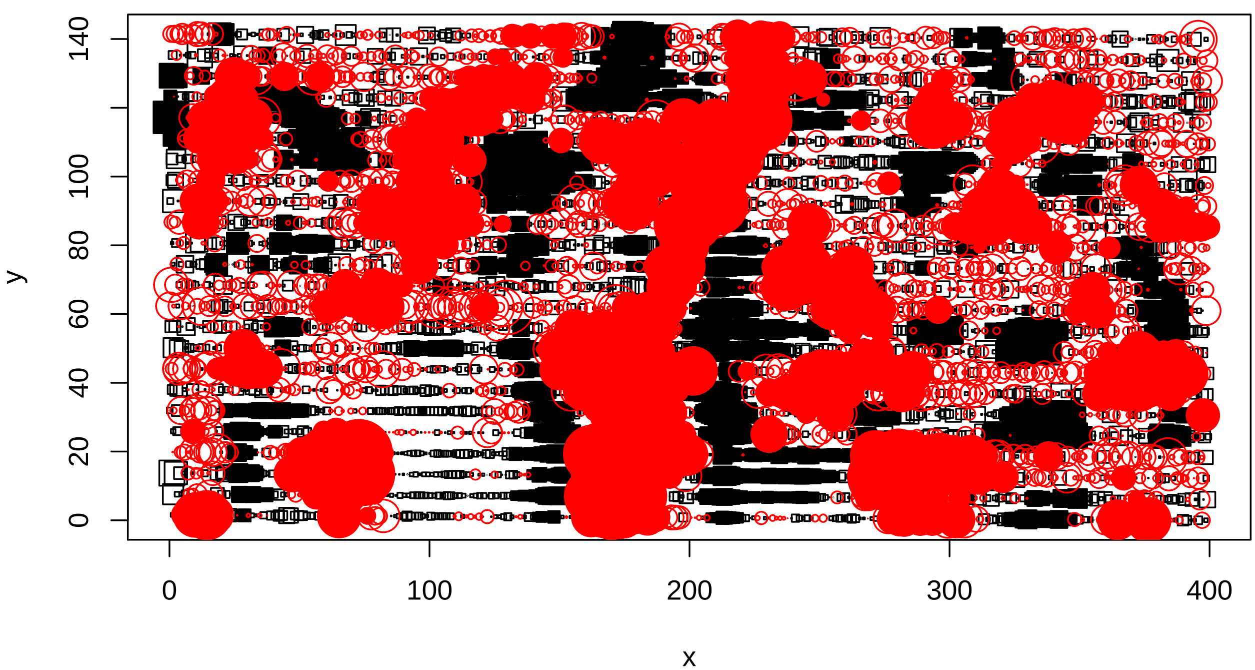
```
Quantile.clg <- correlog(EastQuarter.dat$Easting, EastQuarter.dat$Northing, EastQuarter.dat$Quantile,
                         increment=5, resamp=0, quiet=TRUE)
plot(Quantile.clg)
```



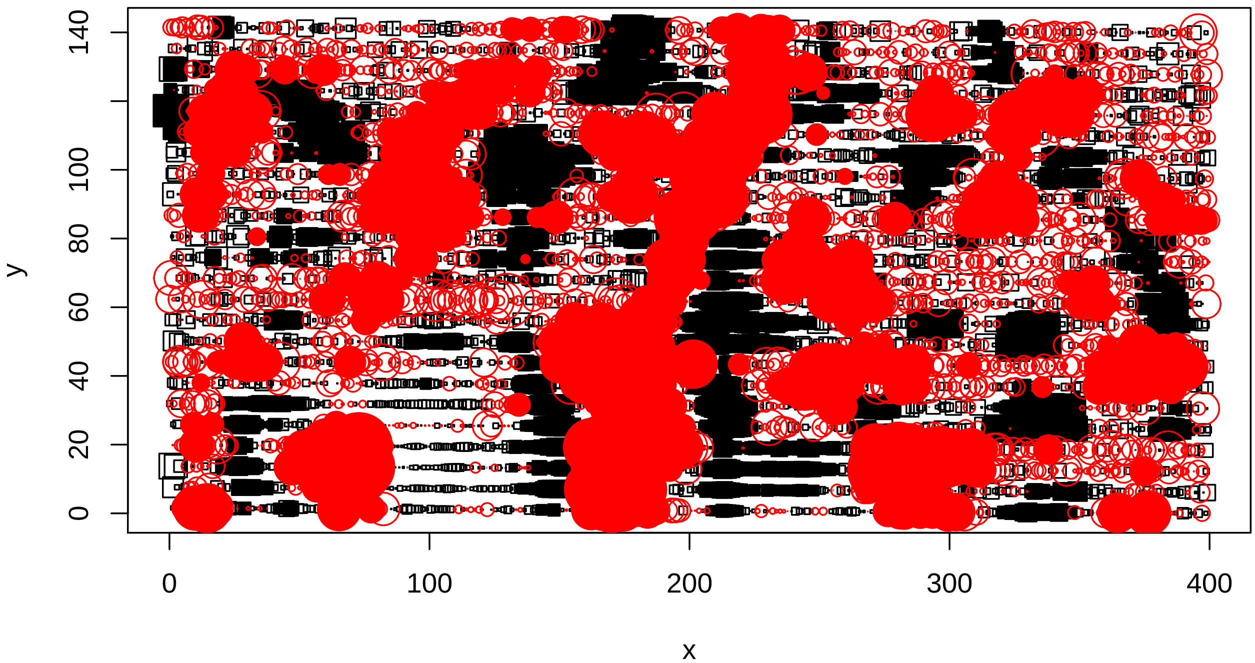

```
Quantile.resid.clg <- correlog(EastQuarter.dat$Easting, EastQuarter.dat$Northing, EastQuarter.dat$Quanti
                         increment=5, resamp=0, quiet=TRUE)
plot(Quantile.resid.clg)
```




```
Quantile.lisa <- lisa(EastQuarter.dat$Easting, EastQuarter.dat$Northing, EastQuarter.dat$Quantile9,
                         neigh=10, resamp=resample, quiet=TRUE)
plot.lisa(Quantile.lisa, negh.mean=FALSE)
```




```
Quantile.resid.lisa <- lisa(EastQuarter.dat$Easting, EastQuarter.dat$Northing, EastQuarter.dat$Quantile
                         neigh=10, resamp=resample, quiet=TRUE)
plot.lisa(Quantile.resid.lisa, negh.mean=FALSE)
```



Trend + Variety AOV

Does including variety (`Product`) in the model improve spatial correlation?

```
YieldVariety5.lm <- lm(Yield ~ poly(Easting, Northing, degree=5) + Product, data=EastQuarter.dat)
summary(aov(YieldVariety5.lm))
```

```
##                                     Df  Sum Sq Mean Sq F value Pr(>F)
## poly(Easting, Northing, degree = 5)   20 1790930   89547   144.6 <2e-16
## Product                               1  208996  208996   337.6 <2e-16
## Residuals                          6347 3929539     619
##
## poly(Easting, Northing, degree = 5) ***
## Product                             ***
## Residuals
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

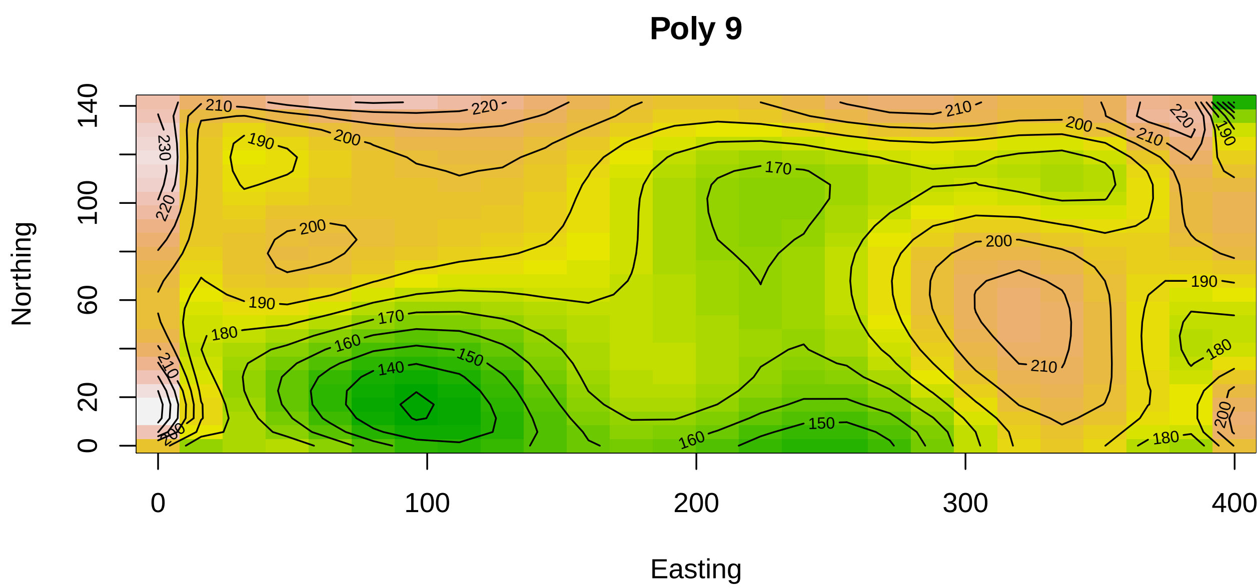
```
YieldVariety9.lm <- lm(Yield ~ poly(Easting, Northing, degree=9) + Product, data=EastQuarter.dat)
summary(aov(YieldVariety9.lm))
```

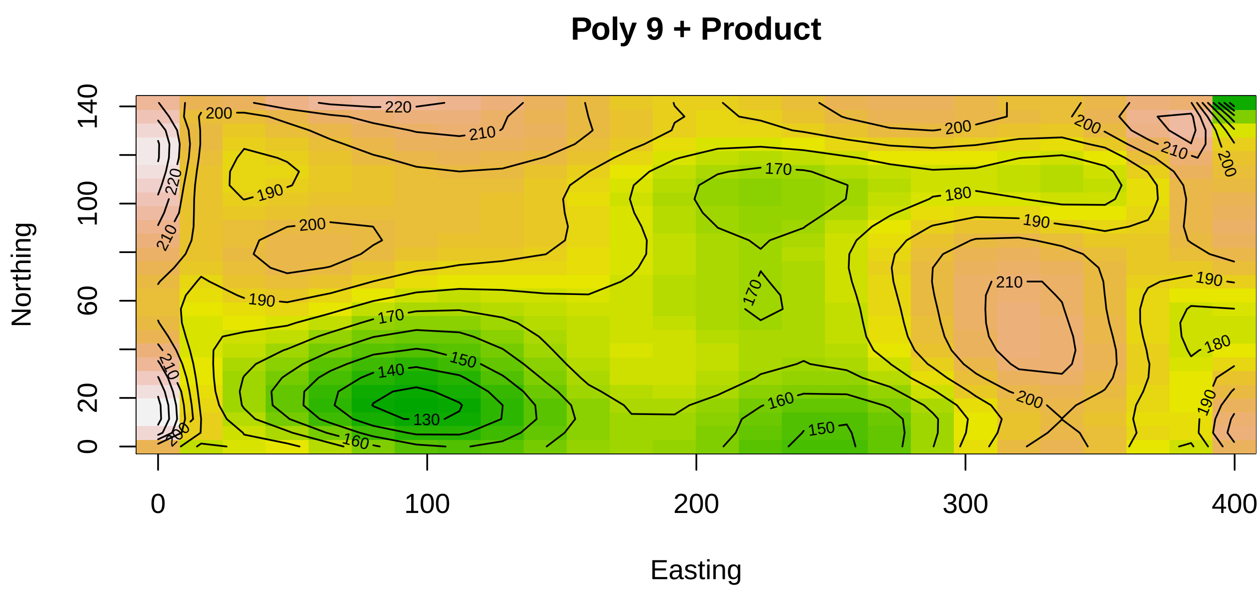
```
##                                     Df  Sum Sq Mean Sq F value Pr(>F)
## poly(Easting, Northing, degree = 9)   54 2284150   42299   78.09 <2e-16
## Product                               1  225691  225691  416.65 <2e-16
## Residuals                          6313 3419624     542
##
## poly(Easting, Northing, degree = 9) ***
## Product                             ***
## Residuals
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
YieldVariety11.lm <- lm(Yield ~ poly(Easting, Northing, degree=11) + Product, data=EastQuarter.dat)
summary(aov(YieldVariety11.lm))
```

```
##                                         Df  Sum Sq Mean Sq F value Pr(>F)
## poly(Easting, Northing, degree = 11)   77 2765104   35910   75.84 <2e-16
## Product                                 1  185973  185973  392.75 <2e-16
## Residuals                             6290 2978388     474
##
## poly(Easting, Northing, degree = 11) ***
## Product                               ***
## Residuals
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
par(mfrow=c(2,1))
contour(Yield9.lm, Northing ~ Easting, image = TRUE,main="Poly 9")
contour(YieldVariety9.lm, Northing ~ Easting, image = TRUE,main="Poly 9 + Product")
```





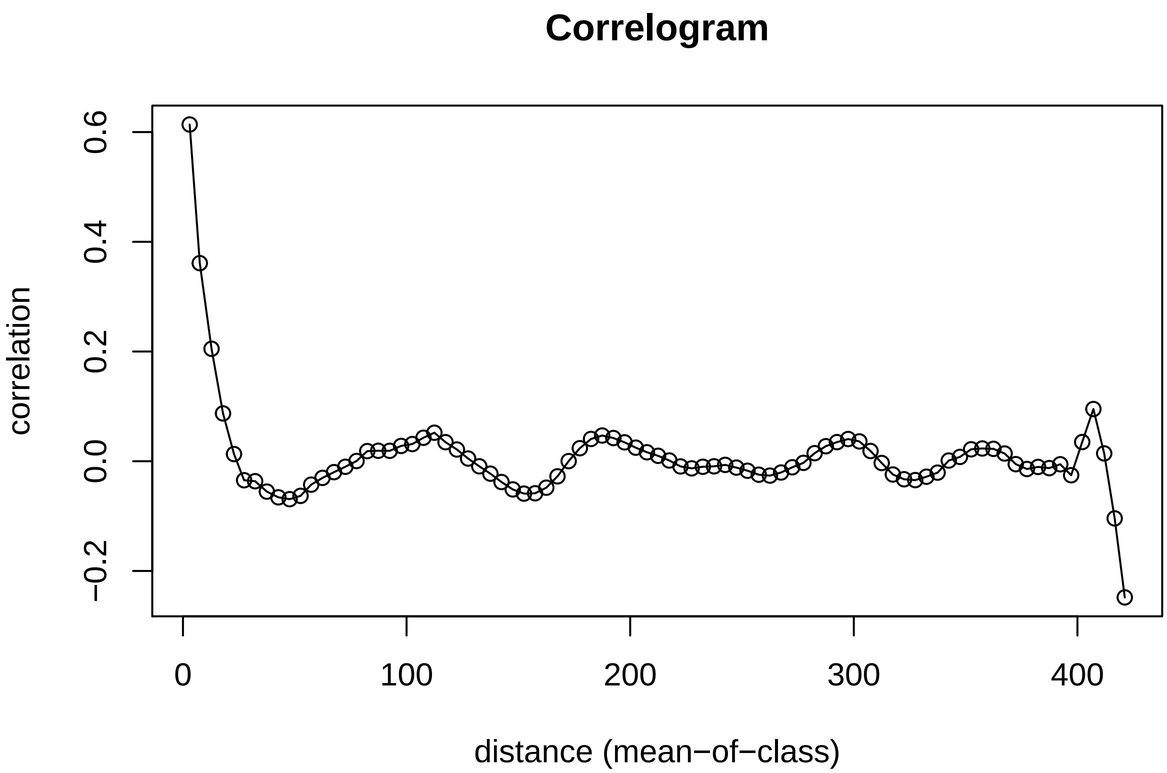

```
par(mfrow=c(1,1))
```

```
EastQuarter.dat$YieldVariety9.resid <- YieldVariety9.lm$residuals
print(Moran9Variety <-Moran.I(EastQuarter.dat$YieldVariety9.resid, Distance.mat))
```

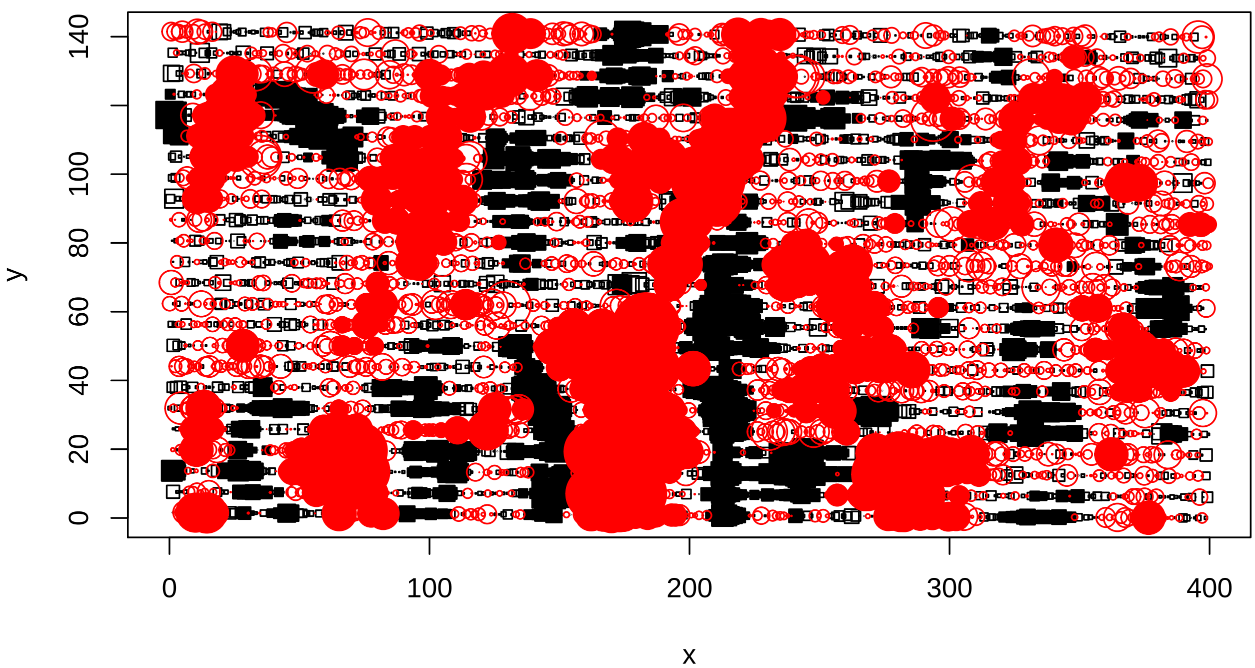
```
## $observed
## [1] 0.04017208
##
## $expected
## [1] -0.0001570352
##
## $sd
## [1] 0.0003680105
##
## $p.value
```

```
## [1] 0
```

```
YieldVariety9.resid.clg <- correlog(EastQuarter.dat$Easting, EastQuarter.dat$Northing, EastQuarter.dat$Y
                           increment=5, resamp=0, quiet=TRUE)
plot(YieldVariety9.resid.clg)
```




```
YieldVariety9.resid.lisa <- lisa(EastQuarter.dat$Easting, EastQuarter.dat$Northing, EastQuarter.dat$Yie
                          neigh=10, resamp=resample, quiet=TRUE)
plot.lisa(YieldVariety9.resid.lisa, negh.mean=FALSE)
```



```r
library(lsmeans)
```

```
## Warning: package 'lsmeans' was built under R version 3.3.2
## Loading required package: estimability
## Warning: package 'estimability' was built under R version 3.3.2
```

```r
Yield.lm <- lm(Yield ~ Product, data=EastQuarter.dat)
lsmeans(Yield.lm, cld ~ Product)
```

```
##  Product   lsmean        SE   df lower.CL upper.CL .group
##  E       177.0918 0.5237125 6367 176.0652 178.1185  1
##  B       192.3915 0.5233015 6367 191.3657 193.4174   2
##
## Confidence level used: 0.95
## significance level used: alpha = 0.05
```

```r
lsmeans(YieldVariety5.lm, cld ~ Product)
```

```
##  Product   lsmean        SE   df lower.CL upper.CL .group
##  E       174.8325 0.9233101 6347 173.0225 176.6425  1
##  B       187.0460 0.9233202 6347 185.2360 188.8560   2
##
## Confidence level used: 0.95
## significance level used: alpha = 0.05
```

```r
lsmeans(YieldVariety9.lm, cld ~ Product)
```

```
##  Product   lsmean       SE   df lower.CL upper.CL .group
##  E       166.7770 1.371456 6313 164.0885 169.4655  1
##  B       179.7785 1.375834 6313 177.0814 182.4756   2
##
## Confidence level used: 0.95
## significance level used: alpha = 0.05
```

```r
lsmeans(YieldVariety11.lm, cld ~ Product)
```

```
##  Product   lsmean       SE   df lower.CL upper.CL .group
##  E       160.8354 1.542851 6290 157.8109 163.8599  1
##  B       174.4552 1.538617 6290 171.4390 177.4714   2
##
## Confidence level used: 0.95
## significance level used: alpha = 0.05
```