

# Zonal Statistics for DMSP OLS Inter-Calibrated



Zonal statistics refers to the calculation of statistics on values of a raster within the zones of another dataset.

This project is a collaborative work of UXO India and IDFC.

In the following **example** the statistics i.e mean, median, mode, standard deviation, minimum, maximum values are calculated for DMSP OLS Inter-Calibrated data.

## load required libraries and packages

```
library(rgdal) # To import raster data

library(maptools) # To plot the data

library(proj4) # To reproject raster

library(xtable) # To export data to html tables

library(raster) # Required for rgdal

library(rgeos) # Required for maptools

library(spatstat) # Analysing spatial point patterns

library(tiff) # Read TIFF images and required for rgdal

library(sp) # Required for maptools

library(data.table) # Modifying columns

library(modeest) # To calculate mode value for the zone

library(foreign) # Required for maptools
```

## Setting working directory

```
setwd("D:/IDFC work/Bulk Zonal Stat Calculation/INPUT/R_Script_Directory") #  
To set directory
```

## Starting time calculation

```
ptm <- proc.time()
```

## Defining function for zonal statistics

```
Zonal_Stat_NTL <- function(x,y) # Define function  
{
```

```

A<-extract(x,y) # Extract raster data zone-wise

R<-array(0,dim=c(length(A),11)) # Create an empty array
for (i in 1:length(A)) # Create for loop
{
  temp=A[[i]] #Get temporary memory
  NTLLumin1<-names(NTLLumin)
  NTLLumin2<- gsub(".tif","",NTLLumin1)
  B<-mlv(temp, method = "mfv") # Find mode of that zone
  B
  R[i,1]=NTLLumin2 # To mention Luminosity Data_Source & year
  SR1<-length(temp[temp>0])
  SR2<-length(temp)
  SR3<-SR1/SR2
  SR<-paste(round(SR3*100,digits=2),"%",sep="")
  R[i,2]= SR2 # Find count for that zone
  R[i,3]= SR1 # Find Lit up pixel count for that zone
  R[i,4]= SR # Percentage area cover by light
  R[i,5]= mean(temp) # Find mean of that zone
  R[i,6]= min(temp) # Find minimum of that zone
  R[i,7]= max(temp) # Find maximum of that zone
  R[i,8]= median(temp) # Find median of that zone
  R[i,9]= sd(temp) # Find Std_Dev of that zone
  R[i,10]= sum(temp) # Find sum of that zone
  R[i,11]=B$M

  rm(temp)
}

BQ=paste(substr(names(x),(stri_length(names(x))-4),(stri_length(names(x))))
, y[[6]], sep="_")
CQ=paste(substr(names(x),(stri_length(names(x))-3),(stri_length(names(x))))
)
colnames(R) <- c("NTL_Data_Source","NTL_Count","NTL_Lit_up_Pixel_count","NTL_Percentage_light_cover_Area","NTL_Mean","NTL_Min","NTL_Max","NTL_Median","NTL_Std_Dev","NTL_Sum","NTL_Mode") # Change column header

z<-cbind(y[[1]],y[[2]],y[[3]],y[[4]],y[[5]],y[[6]],CQ,R,BQ) # Bind data with shape file
return(z)
}

```

## Main function to read files and call the zonal statistics function

```

NTLfileR <- list.files(getwd(), pattern="NTL.*.tif$", full.names=FALSE) # Read list of Raster
for(m in 1:length(NTLfileR)) # for loop to read raster
{
  NTLLumin <-raster(NTLfileR[m]) #

```

```

Move raster to variable Lumin
Sfile<- list.files(getwd(),pattern="*.shp$", full.name=FALSE)    #Read list of Zone
x = NTLLumin@crs          #Assign the projection of raster to a variable
for(q in 1:length(Sfile))  # Loop for reading the shapefiles
{
  Zone <-shapefile(Sfile[q])  # Assigning each shapefile to a variable zone
  Zone <- spTransform(Zone,x)  # reproject shapefile to raster projection system
  M<-Zonal_Stat_NTL(NTLLumin,Zone) # call the zonal statistics function
  if (q<2) { temp_shape_NTL<-M } else { temp_shape_NTL<-rbind(temp_shape_NTL,M) } # Write the output to temp_shape_NTL
}
if (m<2) { temp_raster_NTL<-temp_shape_NTL } else { temp_raster_NTL<-rbind(temp_raster_NTL,temp_shape_NTL) } #Write the output to temp_raster_NTL
rm(temp_shape_NTL) # remove the data
}

NTL<-temp_raster_NTL
proc.time() - ptm #Displaying of time taken.

##      user      system elapsed
##  10.52      0.04    10.62

colnames(NTL) <- c( "city name", "boundary type", "state", "X","y", "z","year",
,"NTL_Data_Source","NTL_Count","NTL_Lit_up_Pixel_count","NTL_Percentage_light_cover_Area",
"NTL_Mean","NTL_Min","NTL_Max","NTL_Median","NTL_Std_Dev","NTL_Sum","NTL_Mode","BQ" ) # Change column header

## [29,] "1876076" "580" "11992_529.622795725"

write.csv(NTL,"D:/IDFC work/Bulk Zonal Stat Calculation/INPUT/R_Script_Directory/NTL_Admin_ZS.csv", na="NA") # Enter Output csv file name and path

```