

**VYSOKÁ ŠKOLA BÁŇSKÁ – TECHNICKÁ UNIVERZITA OSTRAVA**

**Hornicko – geologická fakulta**

**Katedra geoinformatiky**

## **PROSTOROVÁ ANALÝZA DAT**

### **Deterministické metody interpolace**

Lucie Orlíková

Ostrava, 2019

## Úvod do problematiky

Jednou z nejjednodušších a nejčastěji používaných metod pro prostorovou interpolaci je metoda inverzních vzdáleností (v ang. Inverse Distance Weighted). Tato metoda je založena na váženém průměru, jehož váhy jsou nepřímo úměrné nějaké mocnině vzdálenosti mezi interpolovaným bodem a datovými body. Jedná se o vyjádření představy, že bližší měření mají mít větší vliv na výsledek. Jednotlivé body měření se nacházejí různě daleko od body, který chceme interpolovat, a je důležité zvolit váhy tak, aby odrážely tuto skutečnost.

Nejjednodušší způsob je zvolit nepřímo úměrné nějaké mocnině vzdálenosti. Pokud je interpolovaný bod totožný s některým z bodů měření, je vzdálenost mezi body nulová. V tomto případě se příslušná váha položí rovna 1 a ostatní váhy jsou rovny 0. Tím je zajištěno, že interpolační křivka prochází datovými body.

Metody inverzních vzdáleností se tedy řadí mezi přesné interpolátory. U metody IDW platí, že malé hodnoty mocniny  $p$  zvýrazňují vliv vzdálených bodů, zatímco velké hodnoty  $\alpha$  zvýrazňují vliv blízkých datových bodů.

Metoda radiálních bázových funkcí (Radial Basis Function) realizuje představu, že každý datový bod ovlivňuje interpolaci ve svém okolí ve všech směrech stejným způsobem a úměrně vzdálenosti od něj. Při interpolaci dat se používají různé typy bázových funkcí. Mohou být zvoleny zkusmo nebo se k nim dojde na základě určitého požadavku.

Interpolační metoda Spline odhaduje neznámé hodnoty použitím matematických funkcí. Metoda je založena na dvou podmínkách – matematická funkce prochází naměřenými body a má minimální křivost. Každá část povrchu je reprezentována samostatnou polynomickou funkcí odvozenou z lokálních hodnot, přitom musí být zajištěna spojitost sousedních polynomických funkcí na jejich styku.

Pro posuzování kvality interpolace se používá postup zvaný cross-validation. Jde o posouzení na základě toho, jak je interpolační metoda schopna reprodukovat původní data. Velmi častou metodou je metoda nazývaná leave-one-out cross validation. Její podstatou je vynechání nějakého datového bodu a z ostatních dat se vypočítá jeho odhad, který většinou vyjde odlišně od původní hodnoty. Vzniklý rozdíl je tedy možné považovat za chybu.

Další věcí, která nás bude zajímat je průměrná chyba, u dobré interpolace by průměrná chyba měla být blízko nule. Pokud tomu tak není, znaménko průměrné chyby nám říká, zda metoda systematicky podhodnocuje či nadhodnocuje.

### Zadání cvičení:

- Seznamte se s vlastnostmi deterministických interpolačních metod – s metodou inverzních vzdáleností, s metodou radiálních bázových funkcí a se spliny
- Na cvičných datech si vyzkoušejte všechny výše zmíněné interpolační metody, vyzkoušejte si vliv změny parametrů
- Výsledky jednotlivých metod mezi sebou porovnejte a popište rozdíly

## Zpracování cvičení v prostředí R studia:

Pro práci v R Studiu budete potřebovat následující knihovny: ggplot2, gstat, sp, maptools, readxl, fields

```
library(ggplot2)
```

```
library(gstat)
```

```
library(sp)
```

```
library(maptools)
```

```
library(readxl)
```

```
library(fields)
```

## Načtení dat v excelu a vykreslení oblasti

```
oblast <- read_excel("data.xls")
```

```
View(oblast)
```

```
coordinates(oblast) = ~X+Y
```

```
plot(oblast)
```

## Vytvoření gridu pro interpolaci

```
locator(4)
```

```
x.range <- as.integer(c(-20,220))
```

```
y.range <- as.integer(c(-20,220))
```

```
grd <- expand.grid(x=seq(from=x.range[1], to=x.range[2], by=1), y=seq(from=y.range[1], to=y.range[2], by=1))
```

```
coordinates(grd) <- ~ x+y
```

```
gridded(grd) <- TRUE
```

```
plot(grd, cex=1.5)
```

```
points(oblast, pch=1, col='red', cex=1)
```

## Metoda inverzních vzdáleností

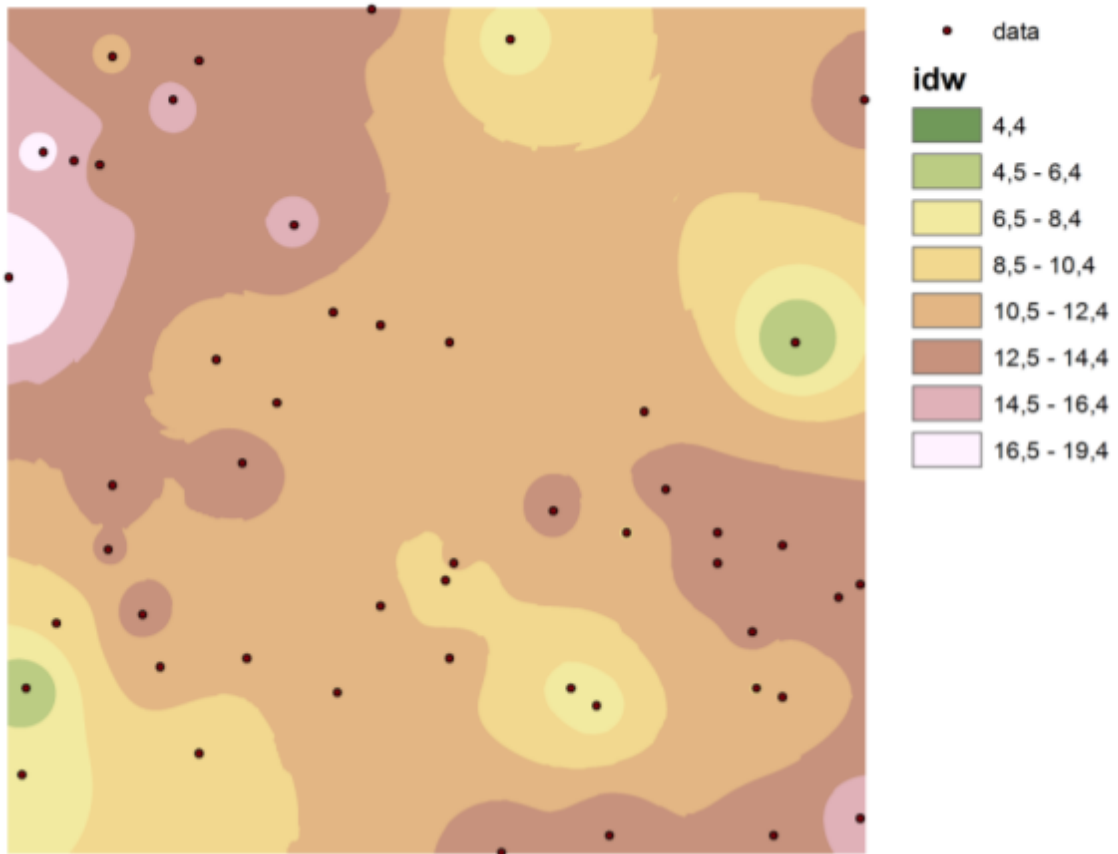
```
idw <- idw(formula = Hodnota ~ 1, locations = oblast, + newdata = grd)
```

```
idw.output = as.data.frame(idw)> names(idw.output)[1:3]<-c("X", "Y", "var1.pred")
```

```
plot<-ggplot(data=idw.output,aes(x=long,y=lat))
```

```
layer1<-c(geom_tile(data=idw.output,aes(fill=var1.pred)))
```

```
plot+layer1+scale_fill_gradient(low="#FEEBE2", high="#7A0177")+coord_equal()
```



### Metoda Spline

```
x <- oblast$X y <- oblast$Y
```

```
z <- oblast$Hodnota
```

```
tp <- cbind(x,y)
```

```
tpo <- list(x=tp,y=z)
```

```
tout <- Tps(tpo$x,tpo$y)
```

