

Ruimtelijke Analyse en GIS 1 (GEO2-4206) TUSSENTOETS 2005
Bacheloropleiding Aardwetenschappen

VRIJDAG 11 MAART 2005 Tijd: 09.00 - 12.00.

Nota bene: U moet alle vragen beantwoorden.

Tekeningen en grafieken kunnen uw antwoorden verhelderen.

1. Geef een korte definitie van de volgende GIS-termen. Geef bij elke definitie ook een voorbeeld van toepassing.

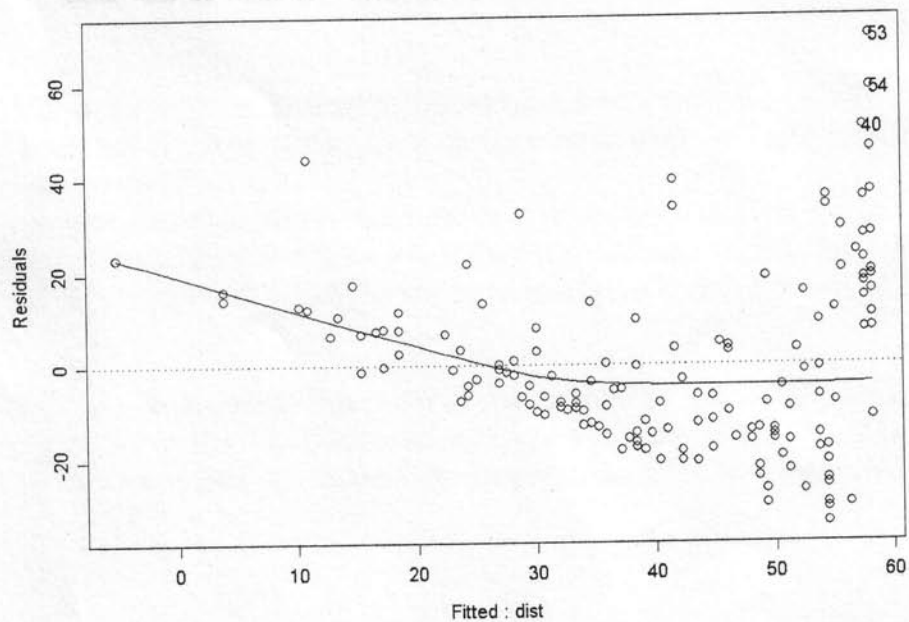
- a) Digitaliseren
- b) Remote sensing
- c) Digital elevation model (DEM)
- d) Geografische entiteiten
- e) Interpolatie
- f) GPS
- g) Thiessen polygoon
- h) Ruimtelijke attributen
- i) Exploratory data analyse
- j) Polygoonkaart

2. Beschrijf en vergelijk twee verschillende methoden voor het coderen van geografische gegevens in een computer. Geef de voor- en nadelen van beide methoden mbt diverse toepassingen zoals

- a) de grenzen van landeenheden en/of de continenten in de computer zetten
- b) het in kaart brengen van het voorkomen van vulkanen in Zuidoost Azië
- c) het opnemen van ruimtelijke gegevens meet een "airborne sensor" of satelliet

3. Beschrijf de achterliggende redenen waarom GIS DBMS (DataBase Management Systems) veel gebruikt worden. Leg de verschillen tussen hiërarchische, netwerk en relationele systemen uit en geef aan waanneer de ene boven de andere is te prefereren.

- a geef de regressievergelijking voor elk van de twee lineaire modellen; vul hierbij de gevonden waarden voor de regressiecoëfficiënten in, en geef (kort) een interpretatie van deze modellen.
- b leg uit hoe het kan dat de helling voor *dist* anders is voor model 1 als voor model 2.
- c bereken de gecorrigeerde R^2 , \bar{R}^2 voor beide modellen.
- d welk van de maten is geschikt voor het vergelijken van de kwaliteit van deze twee modellen: (1) residual standard error, (2) R^2 , (3) \bar{R}^2 , (4) F, (5) p-value van F, (6) residual degrees of freedom.
- e hieronder volgt een van de zgn. *diagnostic plots* voor het eerste model. Leg uit hoe de waarden voor de *x*- en *y*-coördinaat zijn verkregen.
- f beoordeel het eerste model aan de hand van de gegeven figuur: zijn de residuen structuurloos? is de variantie van de residuen constant?



Vraag 4. Hieronder volgt de S-Plus uitvoer van een tweetal lineaire modellen. De afhankelijke variabele copper is koper concentratie in de bodem; de variabele dist is een numerieke variabele met afstand tot de rivier; xkm en ykm zijn de ruimtelijke coördinaten, met eenheid km. **Eerste model:**

*** Linear Model ***

Call: lm(formula = copper ~ dist, data = maas, na.action = na.exclude)

Residuals:

Min	1Q	Median	3Q	Max
-33.43	-13.8	-3.572	10.73	69.73

Coefficients:

	Value	Std. Error	t value	Pr(> t)
(Intercept)	58.9115	2.4529	24.0169	0.0000
dist	-0.0641	0.0067	-9.6082	0.0000

Residual standard error: 18.76 on 153 degrees of freedom

Multiple R-Squared: 0.3763

F-statistic: 92.32 on 1 and 153 degrees of freedom, the p-value is 0

Analysis of Variance Table

Response: copper

Terms added sequentially (first to last)

	Df	Sum of Sq	Mean Sq	F Value	Pr(F)
dist	1	32497.73	32497.73	92.3166	0
Residuals	153	53859.78	352.02		

Tweede model:

*** Linear Model ***

Call: lm(formula = copper ~ dist + ykm + xkm, data = maas, na.action = na.exclude)

Residuals:

Min	1Q	Median	3Q	Max
-34.51	-12.57	-2.512	10.65	67.51

Coefficients:

	Value	Std. Error	t value	Pr(> t)
(Intercept)	48.4870	4.7207	10.2712	0.0000
dist	-0.0529	0.0079	-6.7115	0.0000
ykm	9.3688	3.2843	2.8527	0.0049
xkm	-8.7267	4.5919	-1.9004	0.0593

Residual standard error: 18.26 on 151 degrees of freedom

Multiple R-Squared: 0.417

F-statistic: 36 on 3 and 151 degrees of freedom, the p-value is 0

✓

21

Vraag 5. Hieronder volgt een tabel met waarnemingen van regoliet (bodem) dikte (cm) als functie van hellingklasse (steil, licht hellend, vlak):

steil	licht hellend	vlak
40	50	120
60	70	90
45	60	80
55		90

- Voer een ANOVA-toets uit op deze gegevens. Vergeet daarbij niet aan te geven wat de vooronderstellingen zijn, wat de nul-hypothese is en wat de kritische waarde is voor de toetsingsgrootte. Geef ook de volledige berekening, of geef aan hoe je e.e.a. berekend hebt.
- De computer-uitvoer van een toets op dezelfde data geeft een $Pr(F)$ waarde van 0.003; hoe moeten we deze interpreteren?
- welke van de gemaakte vooronderstellingen is (waarschijnlijk) problematisch voor de variabele regolietdikte?