

Interpolation i laser skannede højdedata

I dette bidrag sammenlignes forskellige metoder til interpolation i laser skannede højdedata. Denne interpolation udgør et vigtigt element i genereringen af en rasteriseret højdemodel.

De tyve datasæt, der er analyseret, dækker ti 100 meter gange 100 meter delområder i det militære øvelsesterræn vest for Oksbøl. Der er analyseret et datasæt fra 2007 og et fra 2013. For data fra 2007 er punkttætheden $\frac{1}{2}$ til ét punkt/m², for 2013 er den 3 til 4 punkter/m².

De sammenlignede metoder er

1. nærmeste nabo interpolation,
2. TIN interpolation,
3. lokal middelværdi interpolation,
4. interpolation ved hjælp af invers afstandsvægtning,
5. interpolation ved hjælp af kvadreret invers afstandsvægtning, og
6. *ordinary kriging* (OK).

Metode 1 bruger altid blot den nærmeste nabo. Metode 2 (TIN står for *triangular irregular network*) bruger normalt Delaunay triangulering til udpegelse af tre naboer, som indgår i interpolationen. Denne metode er meget udbredt i GIS-kredse. For metoderne 3-6 bruges de hhv. 3, 5, 10, 20, 30, 40, 50, 60, 70 og 80 nærmeste naboer i interpolationen.

Metoderne er sammenlignet ved hjælp af såkaldt *leave-one-out* (LOO) krydsvalidering. Alle punkter fjernes her ét ad gangen, hvorefter der interpoleres til det fjernede punkt. Der laves så en sammenligning mellem den fjernede værdi og den interpolerede værdi. Jo tættere den interpolerede værdi ligger på den fjernede værdi, jo bedre. Der findes mange mål for, hvor tæt disse værdier ligger på hinanden.. Her bruger vi den gennemsnitlige differens (kaldet *bias*) og kvadratroden af den gennemsnitlige kvadrerede differens (kaldet *root mean squared error*, RMSE).

De foretagne sammenligninger viser generelt, at målt på bias og RMSE er metode 6 bedst fulgt af metoderne 5, 4 og 3. Det viser sig også, at nogle interpolationsmetoder (også metode 6) i nogle tilfælde over-estimerer lave værdier og under-estimerer høje værdier.

For de relevante metoder viser de her foretagne sammenligninger, at bias og RMSE ikke forbedres ved inklusion af mere end 20-30 naboer.

Metode 6 kræver estimation af et såkaldt semi-variogram med efterfølgende estimation af en model. Den er derfor temmelig arbejdskrævende. Selve interpolationen er ligeledes noget tungere beregningsmæssigt end for de øvrige metoder.

Kort version:

I dette bidrag sammenlignes forskellige metoder til interpolation i laser skannede højdedata. Denne interpolation udgør et vigtigt element i genereringen af en rasteriseret højdemodel. Metoderne er sammenlignet ved hjælp af såkaldt krydsvalidering. De foretagne sammenligninger viser, at generelt er en statistisk baseret metode kaldet kriging bedst. Denne metode er også den mest arbejdskrævende, både for dataanalytikeren og for computeren.

Reference

ftp://ftp.kms.dk/download/Technical_Reports/GST_Technical_Report_19.pdf