

► Figur 7.58 Nedbør registreret i 16 samlede punkter inden for en region på ca. 2,5•1,5 km på øen Vagar, Færøerne beliggende ca. 2 km N for øens lufthavn (nord er opad i figuren). Punktobservationerne benyttes til at interpolere værdier mellem disse vha. en såkaldt spline interpolation (teorien bag denne forklares side 224). Beregningen tildeler imidlertid også værdier til områderne uden for de samlede punkter, så her er der tale om ekstrapolation. Interpolationen burde jf. omtalen af det konvekse hylster og figur 7.57 være afgrænset til området inden for de stiplede linier. Uden for dette område er estimeringerne meget usikre og i dette tilfælde utroværdige, fordi nedbøren i det sydøstligste område (som endda ligger inden for det konvekse hylster!) er estimeret til at være negativ, og i det nordøstlige hjørne er de meget overestimerede. Årsagen til de meget lave og meget høje estimater forklares nærmere i afsnittet om spline interpolation. Resultatet af en interpolation på det samme datasæt baseret på den inverse afstandsvægtningemetode (IDW) ses i øvrigt i figur 7.36 side 200.

## 7.5.2 Sampling med henblik på efterfølgende interpolation

### Repræsentationen af det samlede

Som det blev skrevet i kapitel 2 side 27 isoleres et repræsentativt udsnit af en større helhed (en enhed) vha. en sampling (en prøveudtagning). I en statistisk sammenhæng betragtes en sample som et individ, der er medlem af en familie af lignende individer. Hvis der skal indsamles prøver i en helhed med henblik på efterfølgende at give en fladedækkende beskrivelse af variationer inden for denne, må det nøje planlægges efter hvilken metode, der skal samples. Skal der foretages en tilfældig sampling? Skal prøverne være jævnt fordelt, eller er det en fordel at vælge en metode af helt anden karakter?

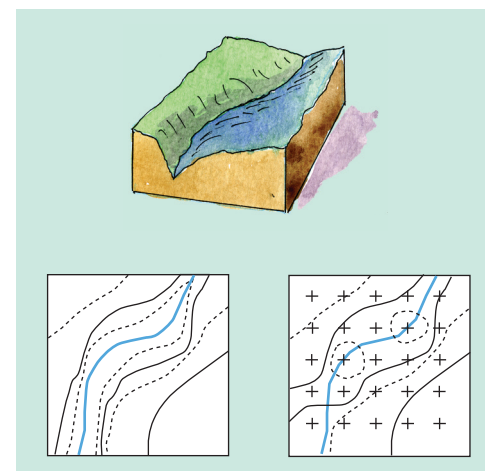
Hvor tæt skal samplingen foretages? Alle disse forhold skal vurderes i forhold til en

senere kvalitetsvurdering af data og de analyser eller tematiseringer, som afledes af den samlede information. Hvis samplingen ikke tages højtideligt, kan det give problemer, som vist i figur 7.59, hvor genskabelsen af en terrænform i et GIS er problematisk.

Ved sampling af meget varierende terrænoverflader må der ofte af tidsmæssige (og økonomiske) årsager foretages et kompromis. Man må i de situationer stille sig tilfreds med en registrering af markante terrænpunkter og ikke udføre en præcis registrering af alle detaljer. Der bør dog ikke efterfølgende være problemer med en genskabelse af terrænet, som vist i figur 7.59. Allerede i samplingsfasen foretages der således en generalisering af fænomener i overensstemmelse med den tematisering, der vælges. Hvis der indsamles højdedata vha. fotogrammetrisk udstyr, klarer udstyret

måske selv målingen af højdeinformation, men som operatør skal det specificeres, hvilken samplingsmetode målingen skal baseres på.

Ved registreringer af skjulte, kvantitative variationer over eller under terræn er det ofte nødvendigt at supplere sine feltindsamlinger med laboratorieanalyser. Resultaterne herfra kan så udbredes til omgivelserne, fx ved interpolation af punktenheder til helheder.



► Figur 7.59 Eksempel på en uheldig sampling af højder i en V-formet kløft med et vandløb i bunden. Her fremvist i et 3D-blokdigram øverst og i en plan afbildning med højdekurver nederst tv. I figuren nederst th. repræsenteres områdets højdevariationer af værdier indsamlet i punkterne markeret med + i henhold til en regelmæssig samplingsstrategi som vist i figur 7.60a. Denne sampling kan føre til en uheldig rekonstruktion af det virkelige terræn baseret på interpolation, som udtrykt ved højdekurverne i figuren nederst th. Der sker, fordi nogle af punkterne indsamlet ved den regelmæssige sampling rammer i bunden af den V-formede kløft. Andre punkter ligger oppe på skrænten og giver anledning til konstruktionen af ikke-eksisterende broer over kløften ved interpolation. For en nærmere beskrivelse af læsningen af højdekurver henvises til figur 2.2 side 24.

## Samplingsstrategier

Det vil ikke være muligt at gennemgå samplingsteori her på meget detaljeret form og med en stærk statistisk fokus. Mht. litteratur på dette område kan der fx henvises til Burrough & McDonnell (1998). Der vil derimod blive givet en række retningslinier for de overvejelser, der bør gøres i tilknytning til samlinger med henblik på efterfølgende interpolation for at (gen)skabe et kontinuert fænomen i et GIS.

Der kan nemlig praktiseres mange forskellige samplingsstrategier afhængig af, hvilket fænomen der ønskes genskabt så troværdigt som muligt i et GIS. En række af disse strategier, som bl.a. benyttes meget af naturvidenskabsfolk (bl.a. biologer, geografer og geologer), ses i figur 7.60.

I figur 7.60a ses en såkaldt **regelmæssig sampling**, som foretages med en ensartet samplingsfrekvens i såvel x- som y-retning i knudepunkterne til et gridnet med ens kantlængde. Denne metode bør bruges, når der ikke er kendskab til områdets kompleksitet, og der ønskes en upåvirket (eng. unbiased) indsamling af data. Der er dog en sandsynlighed for, at der optræder fænomener i området, som ikke berøres af samplingen, hvis udbredelsen af fænomenet er mindre end afstanden mellem de samlede punkter, eller hvis fænomenet er langstrakt, men smalt og forløber parallelt med et af gridnets sider.

For at imødekomme et sådant uheldigt forhold kan der i stedet udføres en **tilfældig sampling** (eng. random sampling), som vist i figur 7.60b, som baseres på en tilfældig spredning af et specifikt antal samplingspunkter inden for en region. Det skal ikke