



Bi9000

GIS v botanice a zoologii



INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Datové modely

- Vektorový datový model
- Rastrový datový model...

...a ještě...

- ...TIN



GISová data - geodata

- **Atributová data** - data prostorově nijak neurčená, v převážné většině případů však prostorově určitelná.
- **Prostorová data** - data obsahující prostorové určení (geometrii) a prostorové vztahy (topologii) objektu (prvku).
- **Geografická data (geodata)** - propojená atributová a prostorová data.

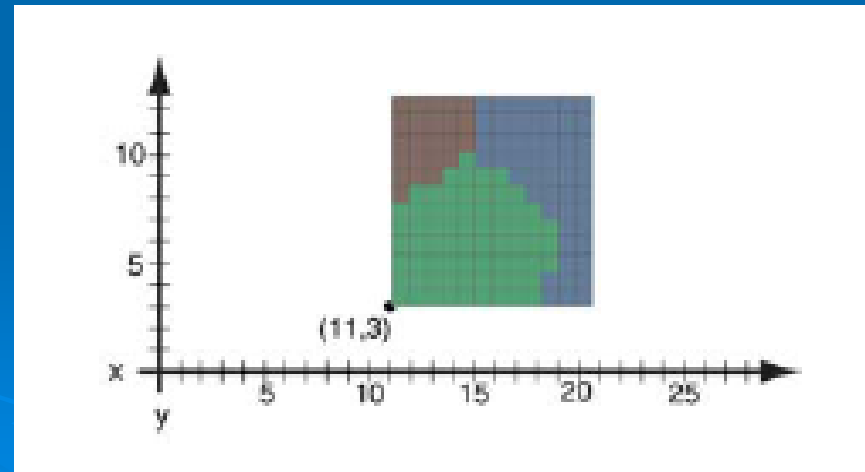
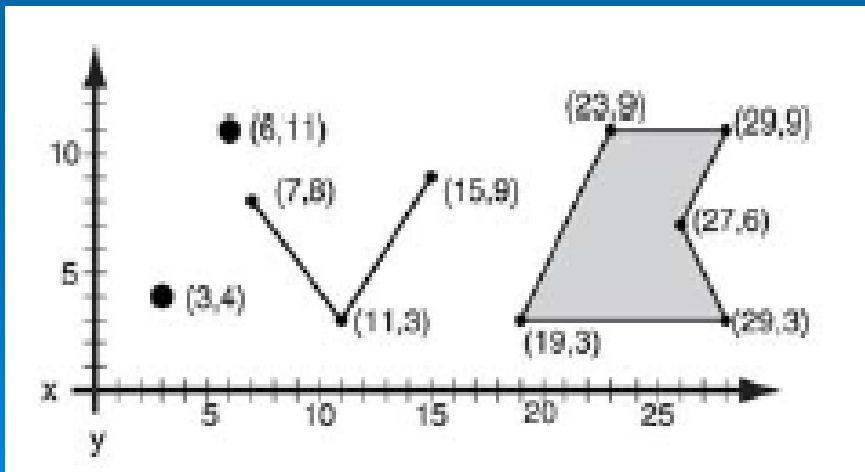


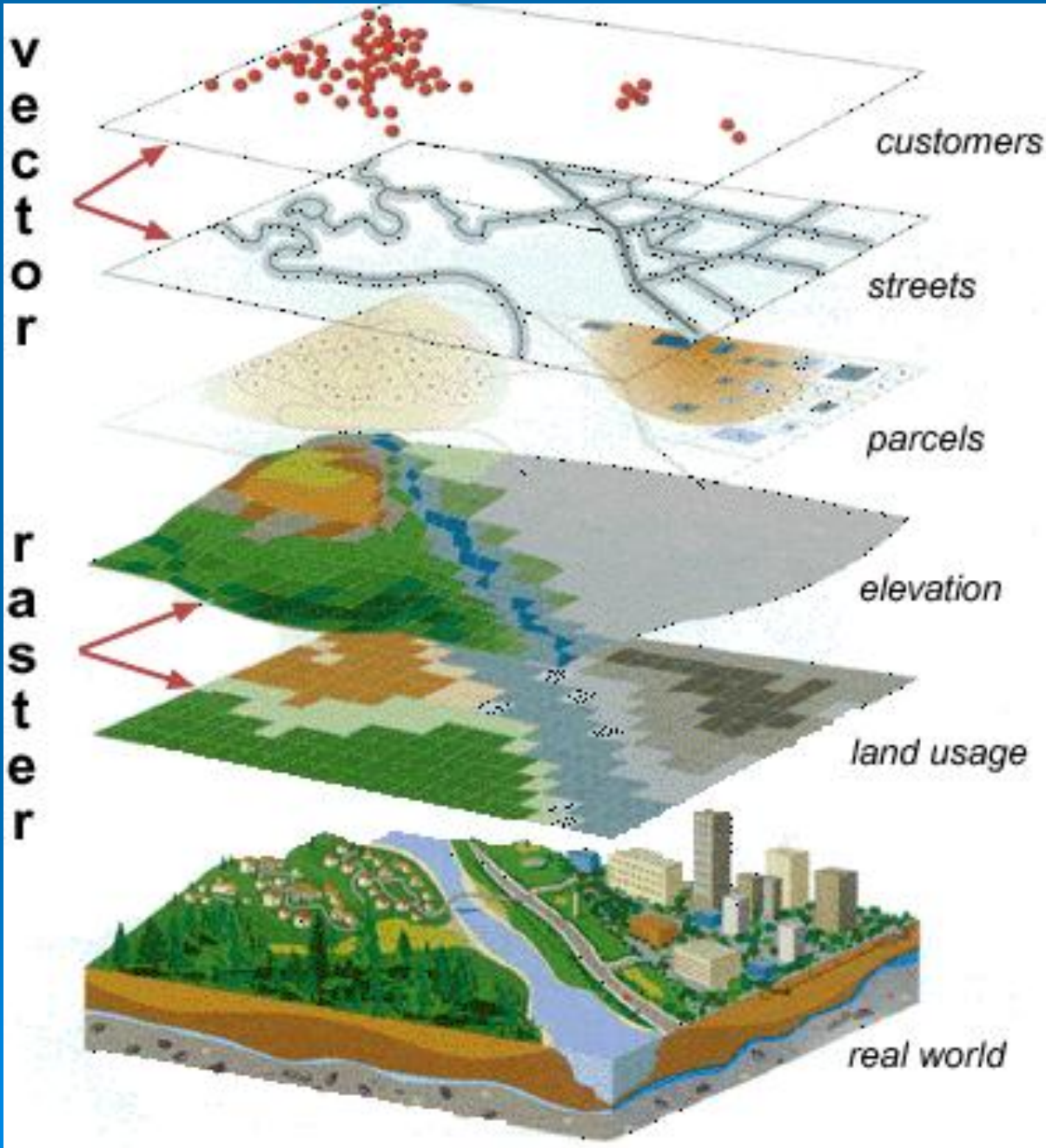
Datové modely:

Pro prostorová data jsou používány převážně 2 modely jejich reprezentace v digitální podobě.

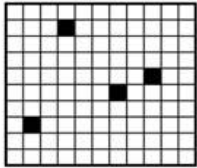

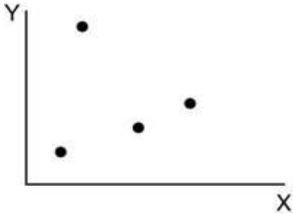
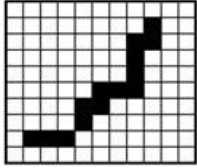


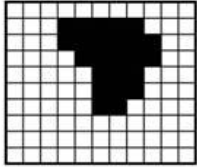
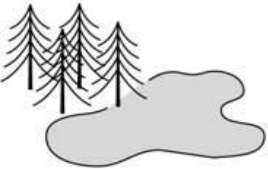
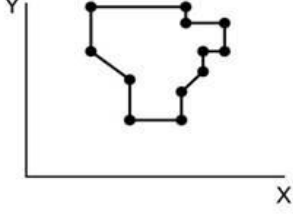
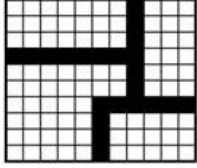
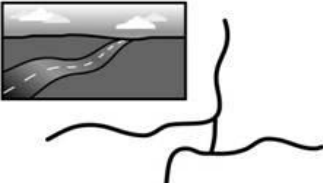
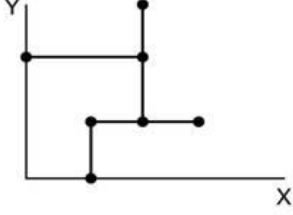
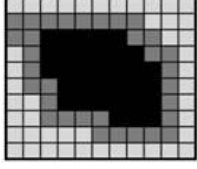

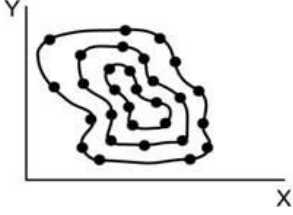
Liší se jak způsobem uložení v databázi, tak charakterem přiřazení tématické (atributové) informace.

Jsou to **VEKTOROVÁ** a **RASTROVÁ** data.





RASTR

The raster view of the world	Happy Valley spatial entities	The vector view of the world
	 x x Points: hotels	
	 Lines: ski lifts	
	 Areas: forest	
	 Network: roads	
	 Surface: elevation	

VEKTOR

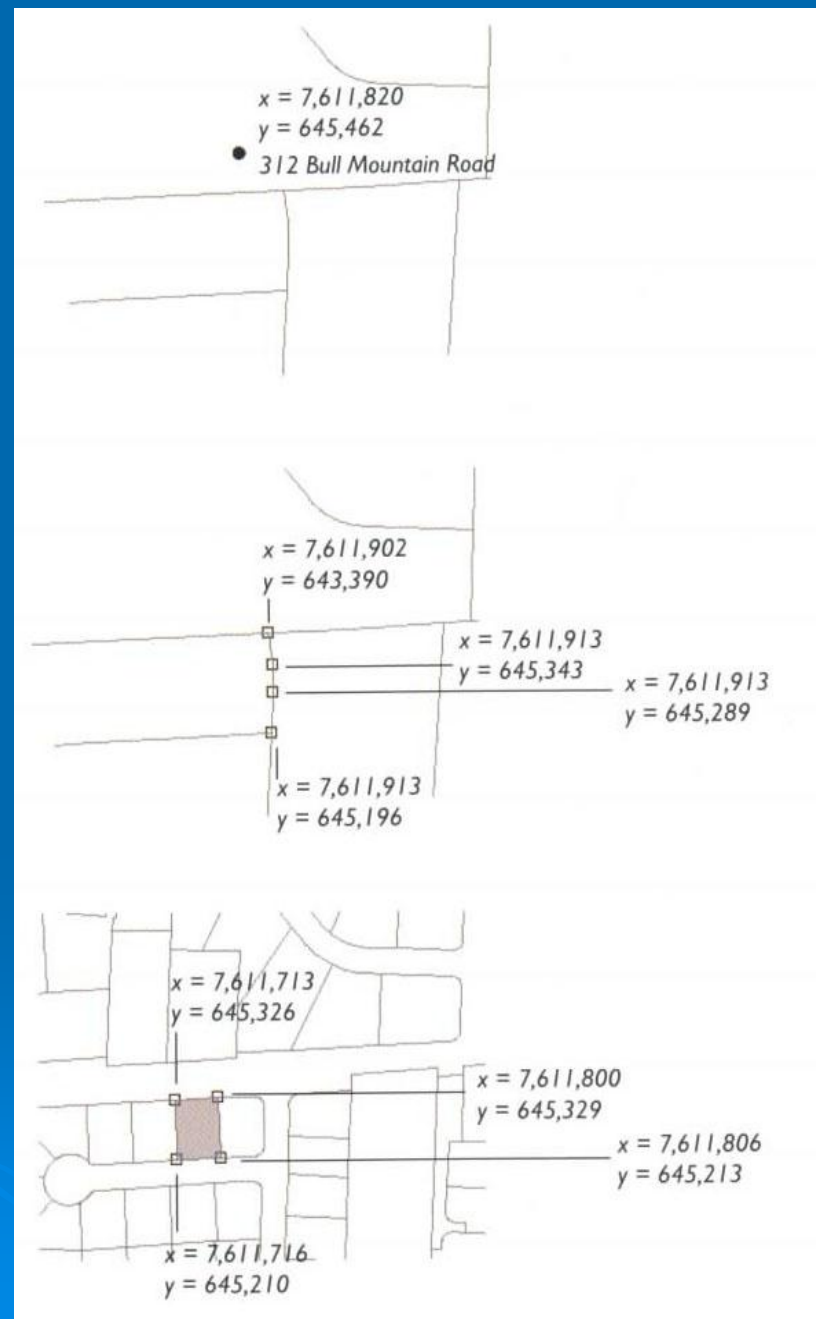
Vektorový datový model

- Základní myšlenkou při použití vektorových dat je snaha vyjádřit geometrické vlastnosti jevů na zemském povrchu pomocí lineárních charakteristik.
- Jednotlivé základní geometrické prvky jsou vyjádřeny jednotlivými body, sestavou bodů, body spojených úsečkami
- Základními prvky vektorových dat jsou:
 - **BOD** (zdroj znečištění, hnízdiště, pramen, památný strom...)
 - **LINIE** (řeka, bariéra podél dálnice, polní cesta)
 - **PLOCHA** (jezero, povrchový důl, fyto geografický okres, povodí, areál výskytu určitého druhu)

BOD...definován dvojicí souřadnic, reprezentuje objekty, které **nemají žádný rozměr** nebo takové, jejichž rozměr je tak malý, že v měřítku mapy nejsou vyjádřitelné plošně; bod má **dimenzi 0** - nelze u něj měřit žádný rozměr

LINIE...je sled orientovaných úseček, definovaných souřadnicemi počátečního a koncového bodu, reprezentuje objekty jako řeky, silnice, potrubí, vedení, tedy objekty tak úzké, že je v měřítku mapy není vhodné reprezentovat plochami nebo také objekty, které nemají definovanou šířku (vrstevnice); linie má **dimenzi 1** - lze u ní měřit délku jen v jednom rozměru

PLOCHA...zvláštní případ linie, jejíž první a poslední bod jsou definovány stejnou dvojicí souřadnic, reprezentuje objekty, jejichž hranice uzavírá nějakou homogenní oblast (například jezera, lesy, zastavěná plocha, ...); Plocha má **dimenzi 2** - lze na ní měřit ve dvou rozměrech.

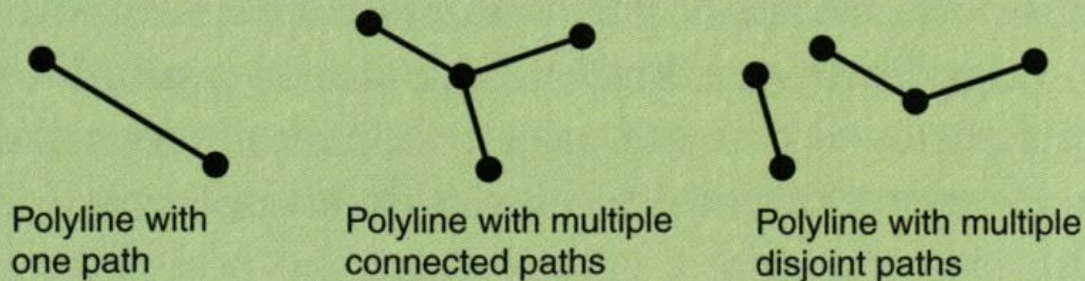


Vektorový datový model

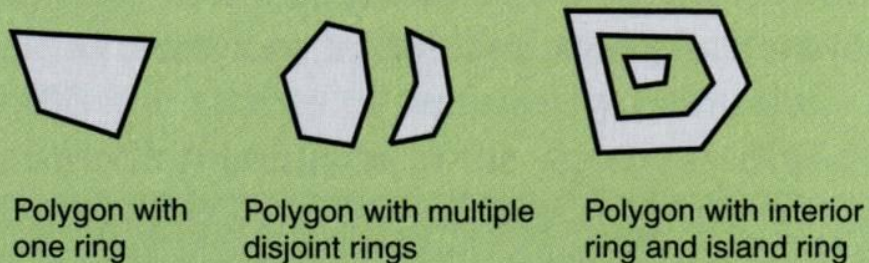
Bod



Linie



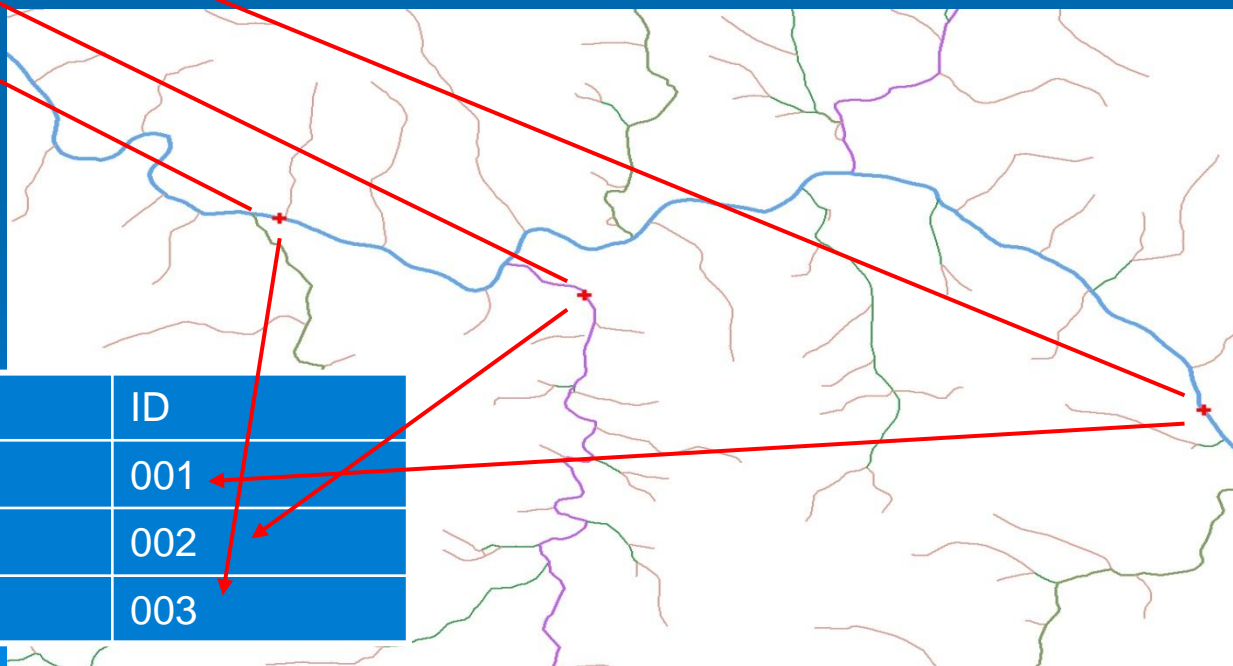
Plocha



Pro **vektorová data** je charakteristický zprostředkovatelský vztah mezi prostorovou a atributovou složkou datového modelu GIS. Prostorová i tematická informace je vázána k **identifikátoru objektu**, který je **základním prvkem vektorových dat**.

ID	lokalita	datum	teplota	pozn.
001	Hrádek	5.5.2005	8	sbíráno metodou dle Smitha
002	Loučka	5.7.2005	15	sbíráno standardní metodou
003	Myslivna	27.11.2005	4	sbíráno standardní metodou

X	Y	ID
56233,2	11474,0	001
56338,1	11421,9	002
56501,5	11388,0	003



Relační databáze

Druhy

ID	rod	druh	cesky
1	Picea	omorika	smrk omorika
2	Picea	abies	smrk ztepilý
3

Záznamy

ID	datum	druh	lokalita	sebral
1	2.7.1996	1	13	8
2

Lokality

ID	lok_jmeno	delka	sirka	popis
...
13	Dolní Vidim	14,675	50,458	palouk ...
14	Odřepsy	12,345	53,658	náves

Sběratelé

ID	prijmeni	jmeno	adresa
...
8	Tichy	Lubos	Chytrakov
9	Chytry	Milan	Ticha Lhota

Kolekce tabulek vzájemně propojených relacemi přes klíčová pole

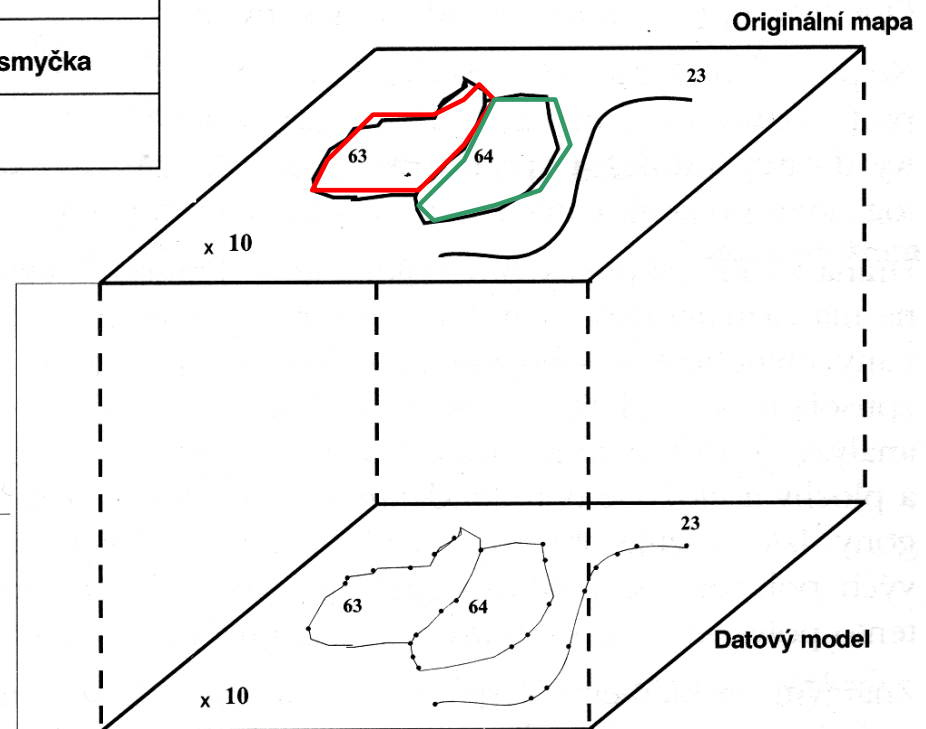
Vektorové datové modely

- **Špagetový** - redundantní uložení dat, žádná topologie (sousedé, konektivita)
- **Topologický** - neredundantní, ale neuspořádané záznamy, dlouhé prohledávání
- **Hierarchický** - neredundantní údaje zvlášt' o bodech, liniích a polygonech v hierarchické struktuře

Špagetový model

Datová struktura

Objekt	Číslo	Poloha
Bod	10	X, Y Jednotlivý bod
Čára	23	$X_1 Y_1, X_2 Y_2, \dots, X_n Y_n$ Řetězec
Polygon	63	$X_1 Y_1, X_2 Y_2, \dots, X_1 Y_1$ Uzavřená smyčka
	64	$X_1 Y_1, X_2 Y_2, \dots, X_1 Y_1$



Topologický model

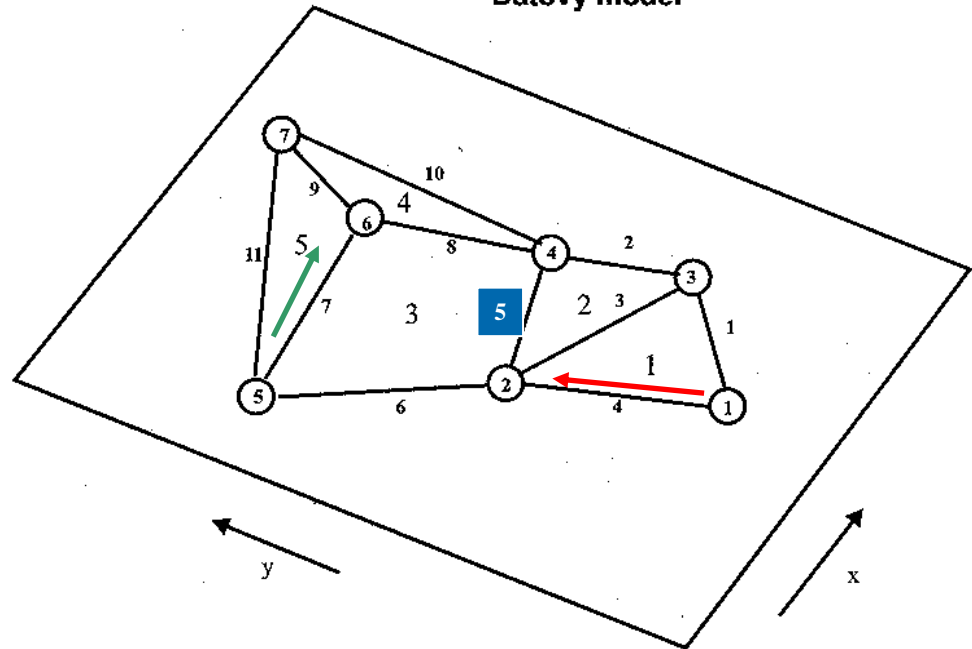
Soubor topologických vztahů

Hrana	Pravý Polygon	Levý Polygon	Počátek v bodě	Konec v bodě
1	1	0	3	1
2	2	0	4	3
3	2	1	2	2
4	1	0	1	2
5	2	2	4	2
6	3	0	2	5
7	3	5	5	6
8	3	4	6	4
9	4	5	7	6
10	0	4	7	4
11	5	0	5	7

Soubor souřadnic bodů

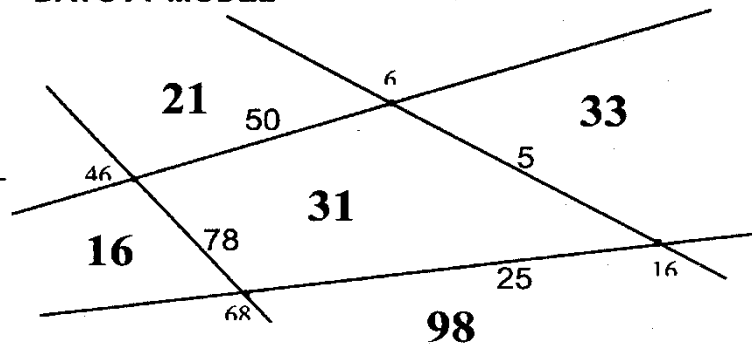
Uzel	X souřadnice	Y souřadnice
1	23	8
2	17	17
3	29	15
4	26	21
5	8	26
6	22	30
7	24	36

Datový model



Hierarchický model

DATOVÝ MODEL



DATOVÁ STRUKTURA

soubor záznamů bodů

Bod	X	Y
6	64	390
46	724	64
68	543	4
16	54	75
atd.		

soubor záznamů linií

Linie	z bodu	do bodu	levý polygon	pravý polygon
50	6	46	31	21
78	46	68	31	16
25	16	68	98	31
5	6	16	33	31
atd.				

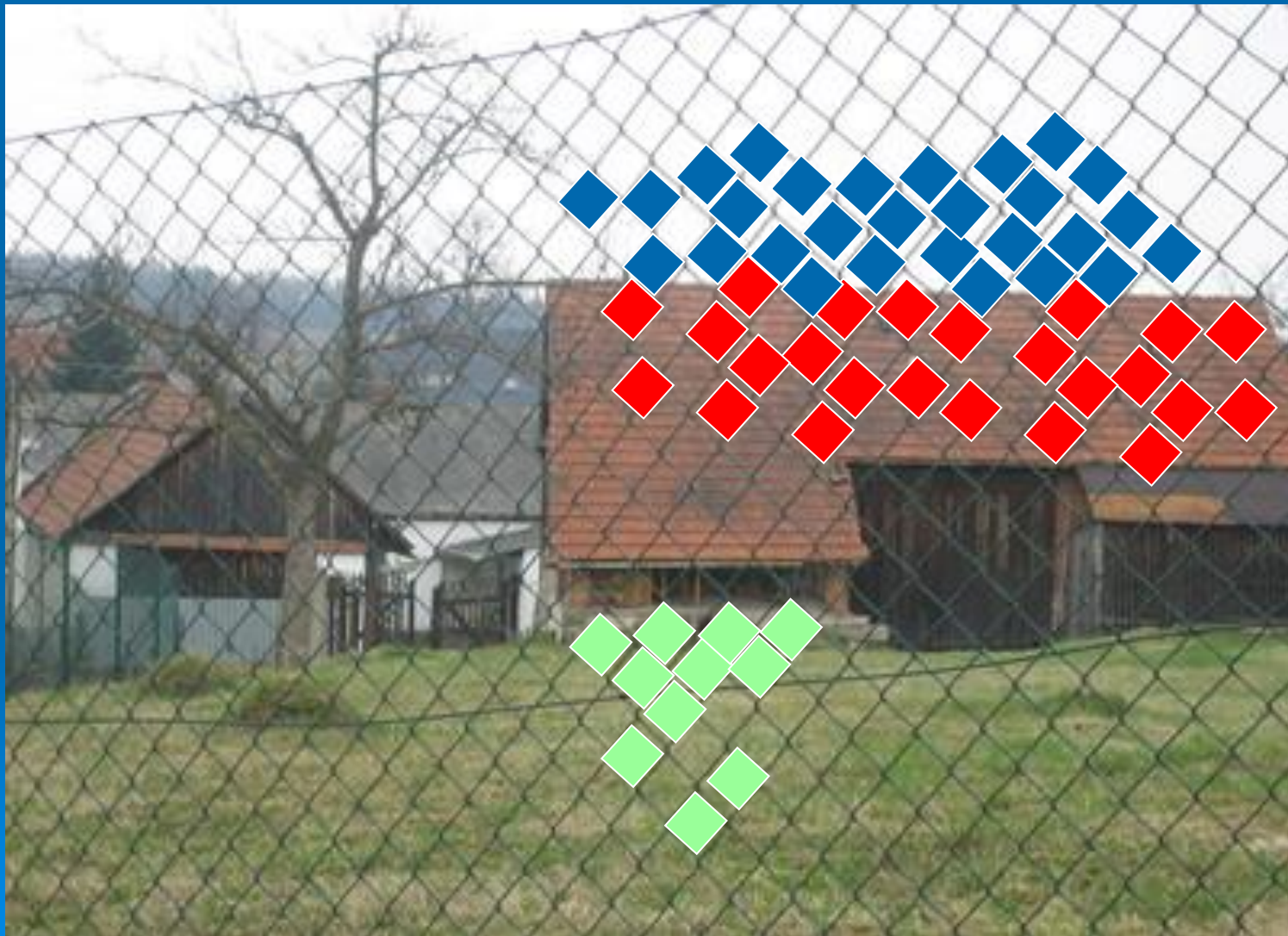
tabulka atributů polygonů

Polygon	Linie	Obvod	Plocha	Majitel	Využití
31	50,5,25,78	546	12698	Černý J.	garáž
atd.					

Rastrový datový model

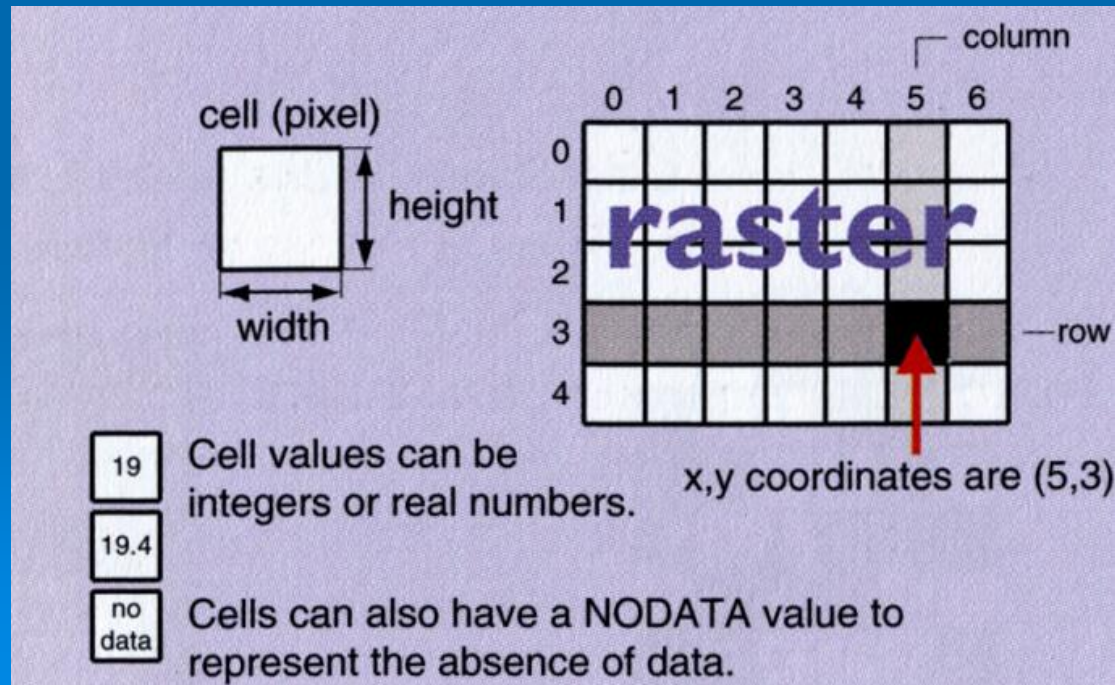


Rastrový datový model



Rastrový datový model

- Základem rastrových dat je **překrytí** zemského povrchu pravidelnou sítí dostatečně malých ploch. Zkoumaný jev na zemském povrchu je pak popsán hodnotami, které jsou vztaženy k plochám této sítě.
- Každá plocha sítě má svou jednoznačnou adresu, danou sloupcovým a řádkovým indexem.



Hodnoty v rastrech

- Diskrétní prvky – Tematický rastr

Jasně rozlišitelné hranice

Hodnoty ukládány jako celé číslo (integer grid)

Landuse, fyto geografické celky, typy porostů, rastry vzniklé reklasifikací...

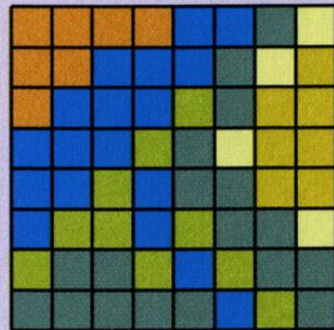
- Spojitě jevy – Kontinuální rastr

Plynule se mění v prostoru, hodnoty sousedních buněk se liší minimálně




Hodnoty ukládány jako reálná čísla (float grid)

DEM (Digital Elevation Model), klima, koncentrace znečišťujících látek...

Uložení atributu v rastru



The attribute table

	Value	Count	Type	Code
	23	7	Fir	400
	29	18	Juniper	410
	31	10	Aspen	420
	37	18	Piñon	500
	41	4	Cottonwood	510
	43	7	Walnut	600

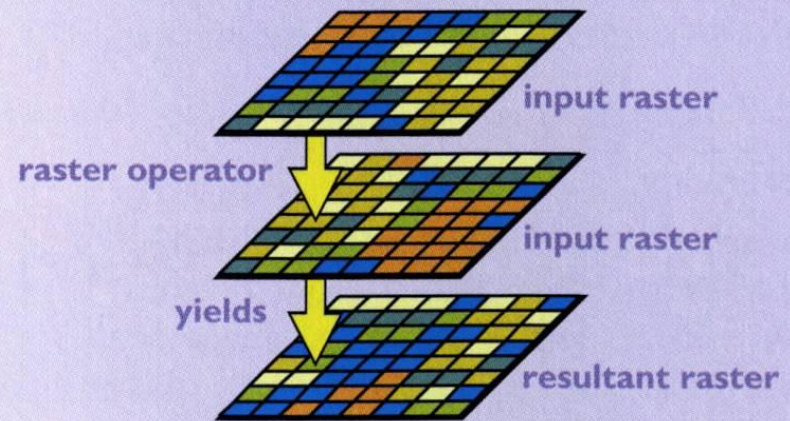
Rasters that have integer valued cells can be defined with an optional attribute table, which records attributes for each unique cell value.

You can add custom fields to the attribute table.

Operace s rastry

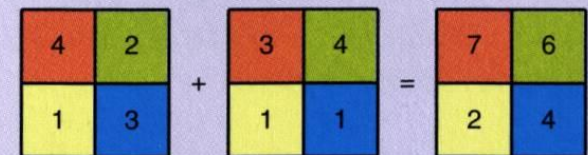
Raster operators

When you are studying an area, you may want to apply a suitability analysis. To do this, you would select rasters with values such as rainfall, soil alkalinity, and insolation and apply a series of operators according to your formula for suitability. Operators can be arithmetic, Boolean, relational, bitwise, combinatorial, logical, accumulative, and assignment.



Map calculation

Mathematical operations can be applied to two rasters and the result is in the output raster. Functions include +, -, /, *, Log, Exp, Sin, Cos, and Sqrt.

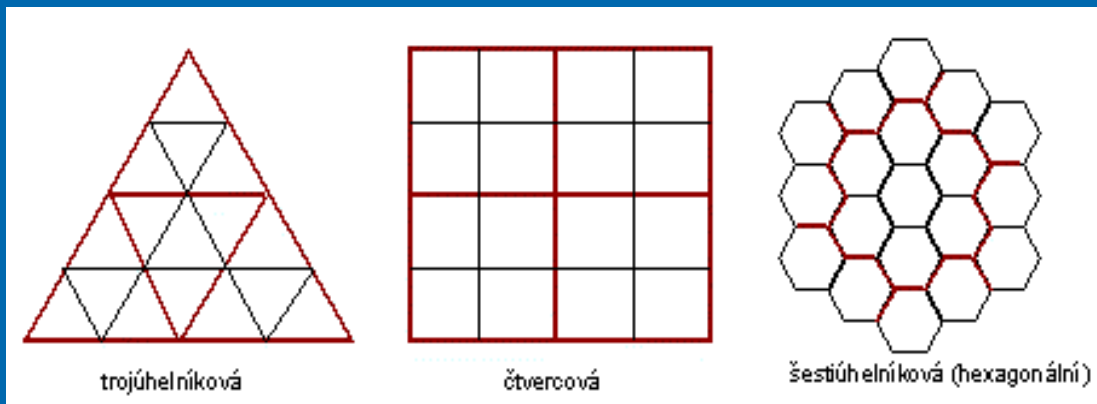


Map query

You can apply Boolean and logical operators on two rasters to create an output raster with true/false values. Operators include And, Or, XOr, Not, >, >=, =, <>, <, and <=.



Typy tvaru buněk



- Nejčastěji se používá **čtvercová mřížka**, protože je kompatibilní s datovými strukturami programovacích jazyků používaných pro tvorbu GIS software, je kompatibilní s mnoha zařízeními pro vstup a výstup dat (monitory, scannery, plottery), je kompatibilní s kartézským souřadnicovým systémem. Pouze buňky tvaru čtverce lze rozdělit na menší čtverce se stejným tvarem a orientací.
- **Trojúhelníková mozaika (TIN)** má tu unikátní vlastnost, že jednotlivé buňky nemají stejnou orientaci, což je výhoda při reprezentování digitálního modelu reliéfu (terénu), kde je každému vrcholu o souřadnicích x, y přiřazena funkční hodnota z (výška $z = f(x, y)$). Jednotlivé trojúhelníky pak implicitně obsahují údaje o svém sklonu a orientaci tohoto sklonu. Daní za tuto vlastnost mnohem větší složitost všech algoritmů pracujících s tímto modelem.
- **Hexagonální mozaika** má tu výhodu, že středy všech sousedních buněk jsou ekvidistantní (stejně od sebe vzdálené), což je výhodné pro některé analytické funkce (např.: paprskové vyhledávání). Ve čtvercové mřížce je toto nemožné a tato vlastnost se musí kompenzovat nebo se prostě zanedbává. Tento tvar buňky se používá jen velmi zřídka.

V rastrovém modelu obecně **neexistuje** popis jedinečných **geoprvků**, ležících v zájmové oblasti, ale jen popis rozložení jedinečných **atributů** v této oblasti.

Faktory, ovlivňující kvalitu zobrazení reálného světa v rastru:

- způsob přiřazení hodnot atributu jednotlivým buňkám (aritmetický průměr, vážený průměr, max. nebo min. hodnota...)
- velikost základní buňky rastru (čím menší buňka, tím přesnější rastr, ale větší soubor - při zmenšení hrany pixelu na polovinu se zvětší objem dat čtyřnásobně - soubor na obr. vlevo potřebuje 16x více diskového prostoru, než soubor na obr. vpravo)
- barevná hloubka (binární, osmibitové, čtyřicetibitové), opět samozřejmě platí, že čím podrobnější, tím větší nároky na paměť počítače



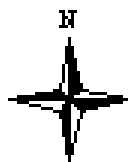
velikost pixelu 50m

velikost pixelu 200m

Land Cover



Note: The land cover data were derived from Landsat TM satellite imagery acquired on 10 July 1989.



Scale bar for image at right which depicts the full extent of the data:

20 0 20 40 60



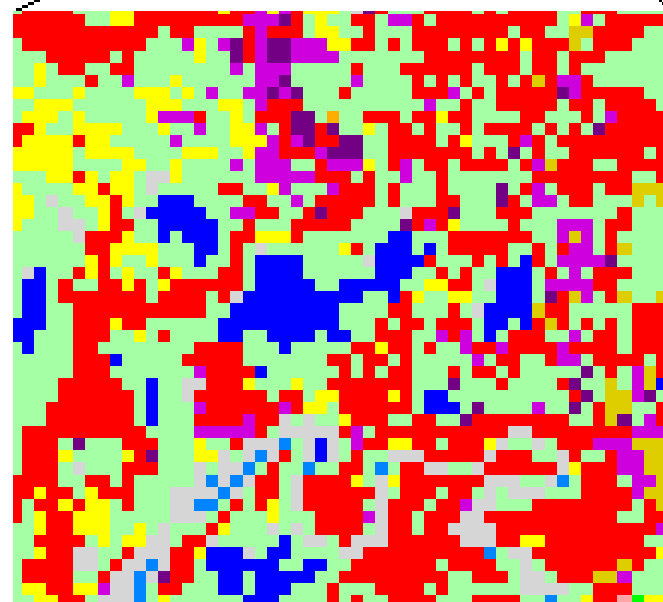
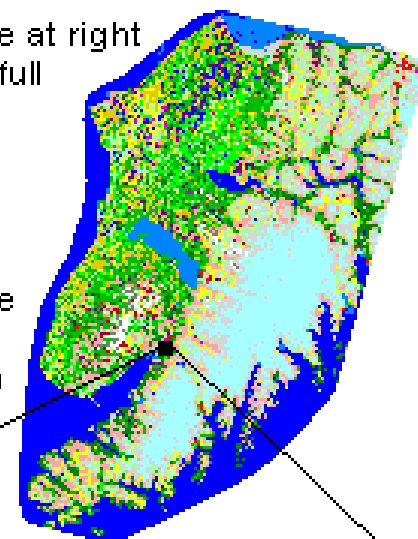
Kilometers

Scale bar for image below:

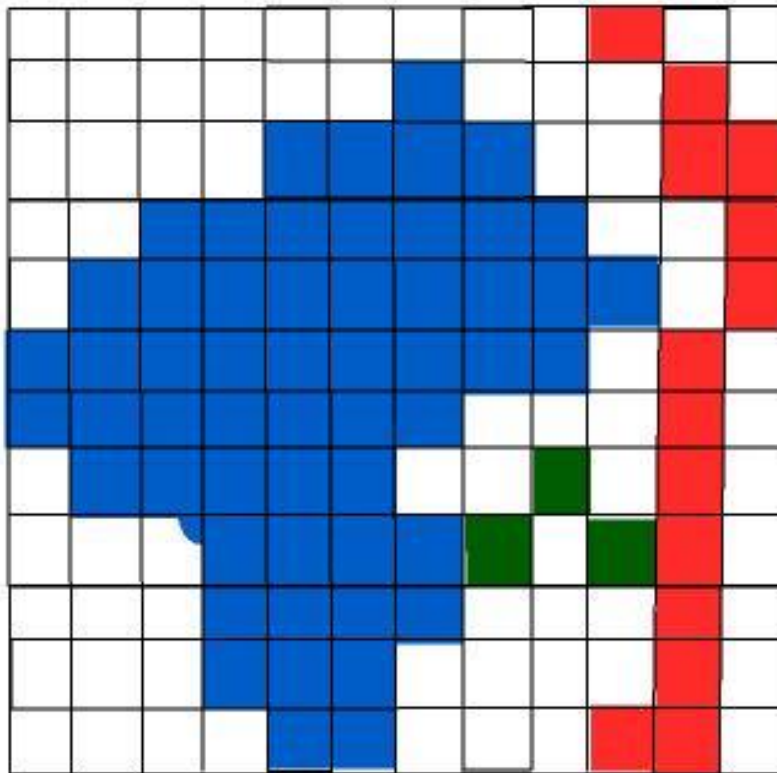
200 0 200 400



Meters



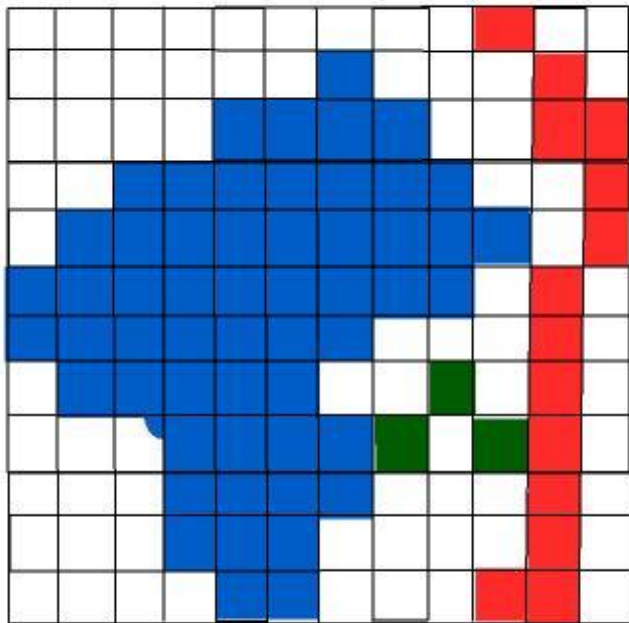
RASTR – bez komprese



0	0	0	0	0	0	0	0	0	0	1	0	0
0	0	0	0	0	0	3	0	0	0	1	0	0
0	0	0	0	3	3	3	3	0	0	1	1	0
0	0	3	3	3	3	3	3	3	0	0	1	0
0	3	3	3	3	3	3	3	3	3	0	1	0
3	3	3	3	3	3	3	3	3	0	1	0	0
3	3	3	3	3	3	3	0	0	0	1	0	0
0	3	3	3	3	3	0	0	2	0	1	0	0
0	0	0	3	3	3	3	2	0	2	1	0	0
0	0	0	3	3	3	3	0	0	0	1	0	0
0	0	0	3	3	3	0	0	0	0	1	0	0
0	0	0	0	3	3	0	0	0	1	1	0	0

RASTR – RLE komprese

ukládání založené na dekompozici původního rastru na úseky se stejnou hodnotou rastru (**RLE** - Run Length Encoding), vhodné zejména pro práci s binárními rastry, vysoká úspora místa na disku.

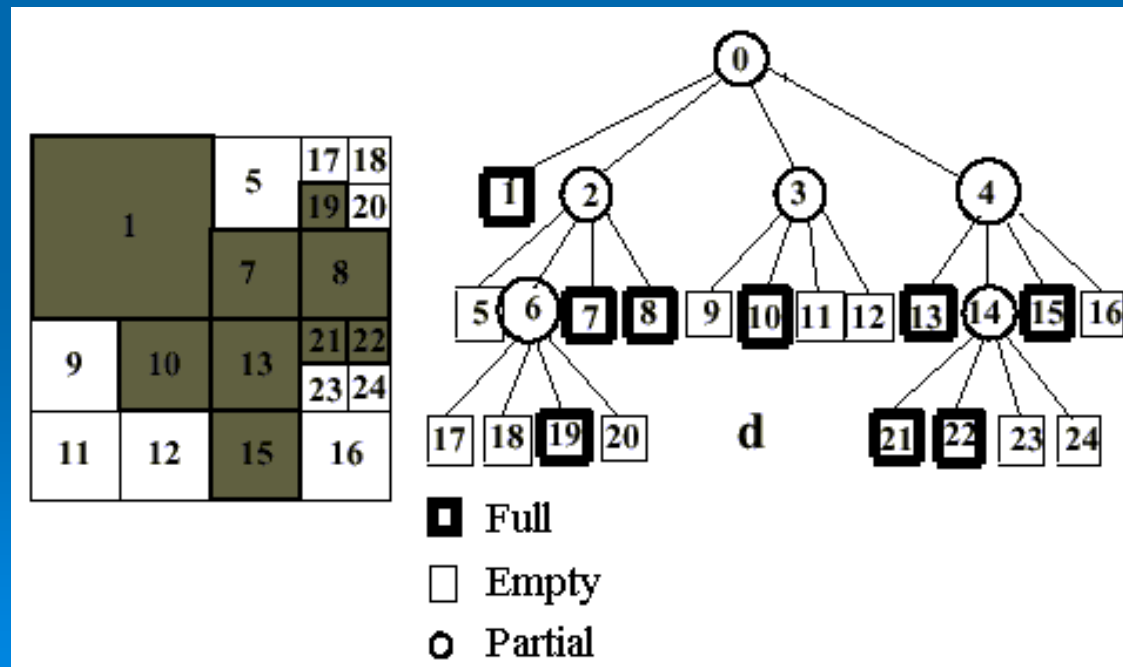


0	0	0	0	0	0	0	0	0	0	1	0	0
0	0	0	0	0	0	3	0	0	0	1	0	
0	0	0	0	3	3	3	3	0	0	1	1	
0	0	3	3	3	3	3	3	3	0	0	1	
0	3	3	3	3	3	3	3	3	3	0	1	
3	3	3	3	3	3	3	3	3	0	1	0	
3	3	3	3	3	3	3	0	0	0	1	0	
0	3	3	3	3	3	0	0	2	0	1	0	
0	0	0	3	3	3	3	2	0	2	1	0	
0	0	0	3	3	3	3	0	0	0	1	0	
0	0	0	3	3	3	0	0	0	0	1	0	
0	0	0	0	3	3	0	0	0	1	1	0	

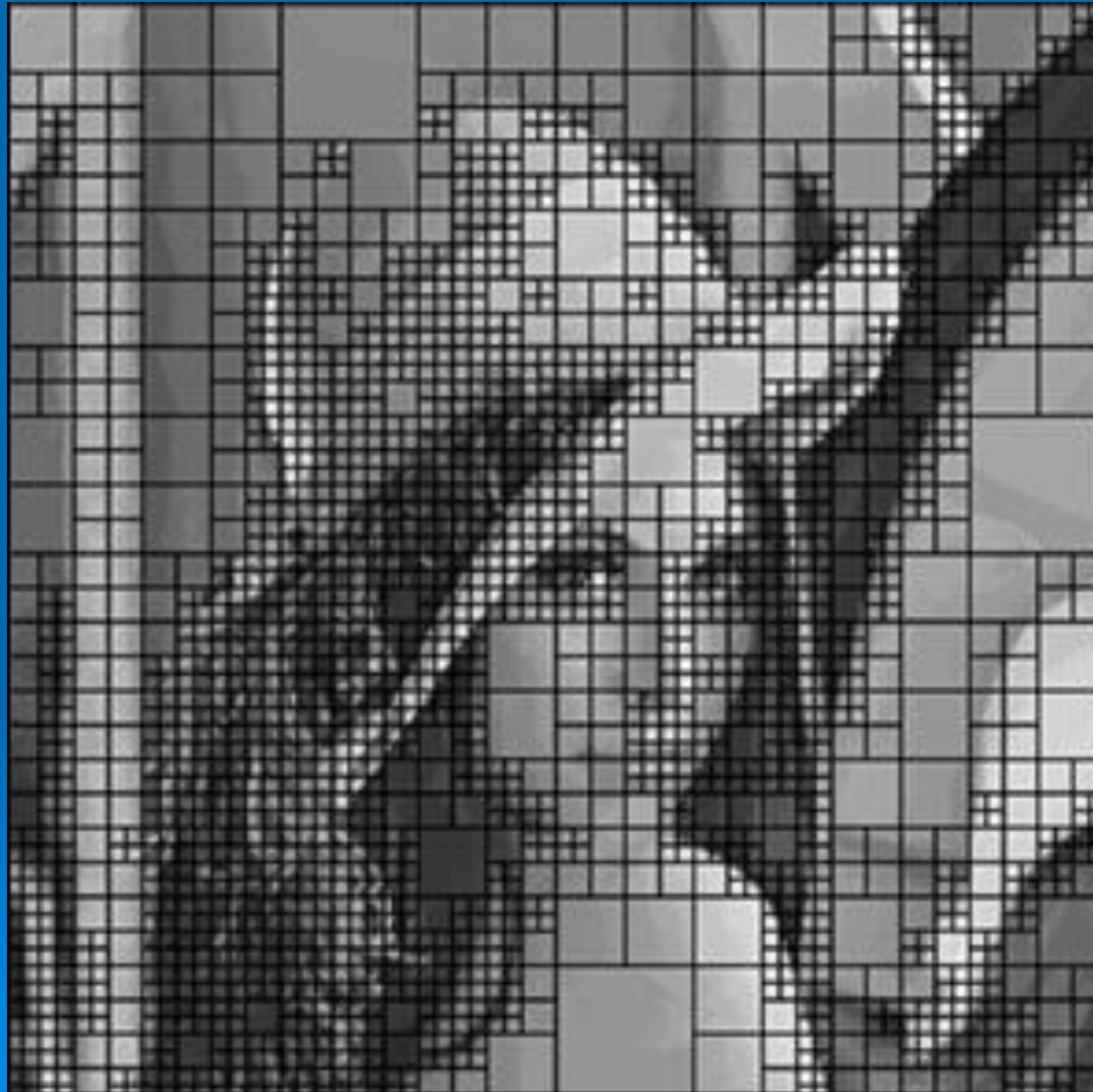
901120
6013201110
40432021
20732011
10931011
93101110
73301110
10532012101110
30431210121110
3043301110
3033401110
4023302110

RASTR – quadtree komprese

postupné pravidelné dělení čtvercové oblasti na kvadranty tak dlouho, až je v celém kvadrantu stejná hodnota, nebo už nelze dále dělit



RASTR – quadtree komprese

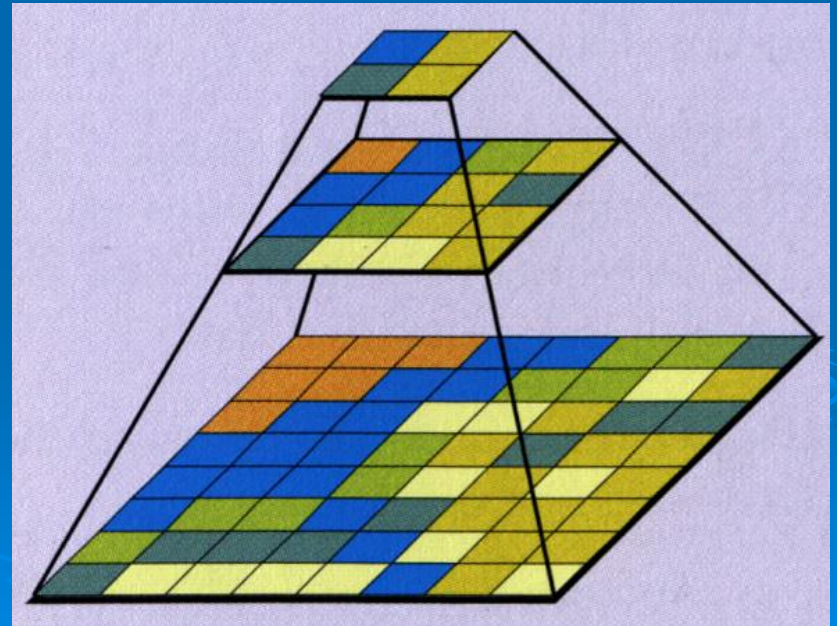


Pyramidy pro rastry

Pyramidy jsou sady resamlovaných rastrů

Používají se při zoomování

Zvyšují výkon při překlesování

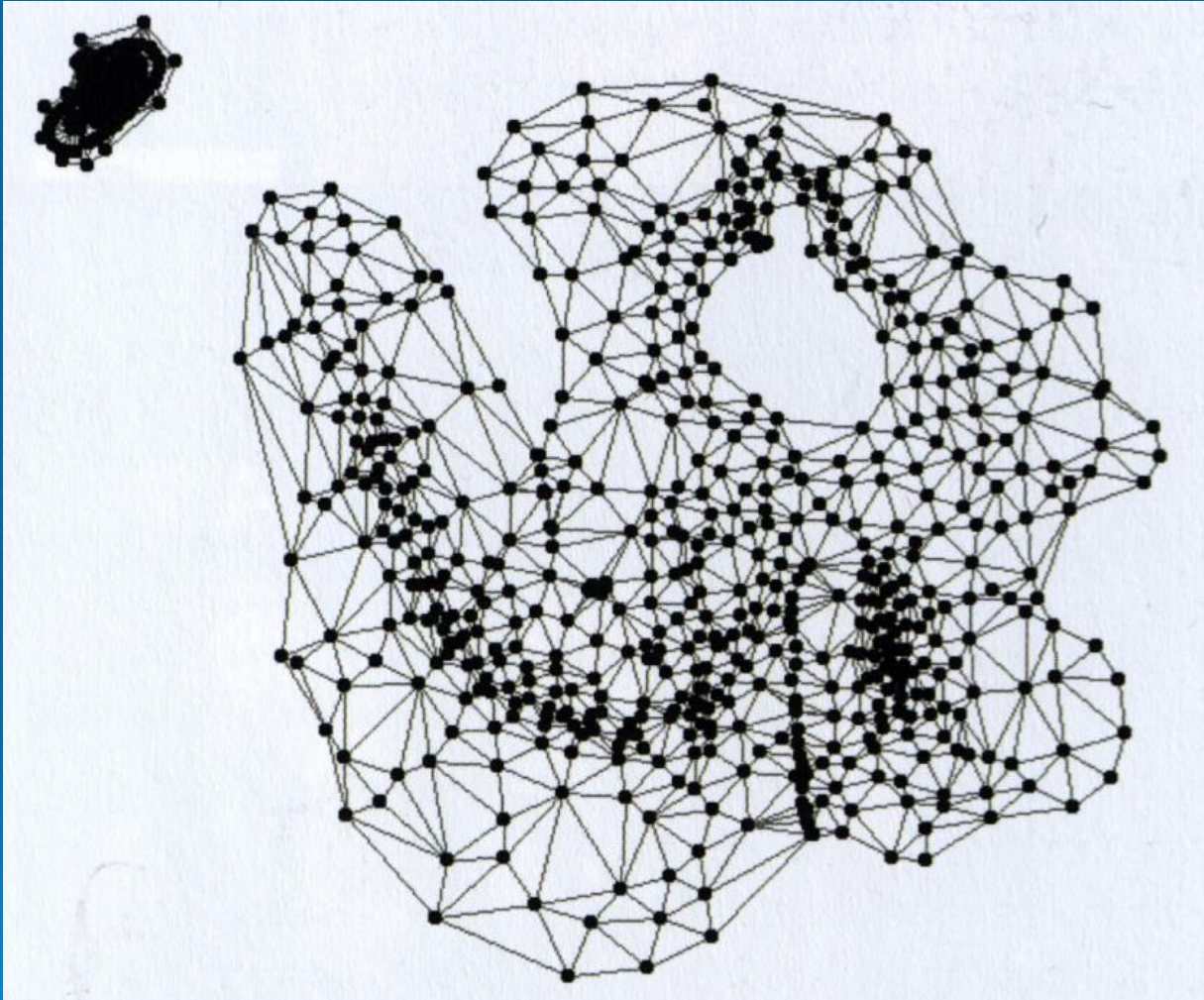


Nepravidelná trojúhelníková síť (Triangulated Irregular Network)

TIN

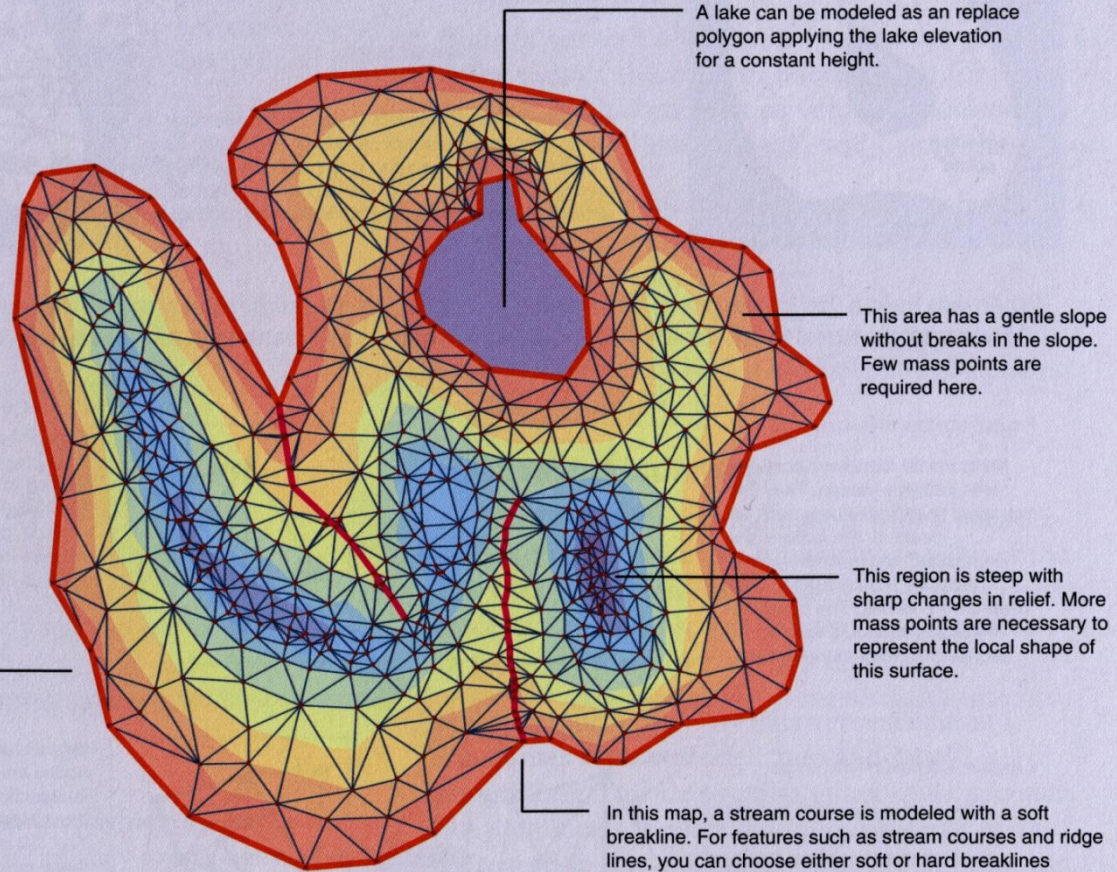
- Svou povahou leží na pomezí mezi **vektorem** a **rastrem**, a velice často se používá pro reprezentaci povrchů, například digitálního modelu reliéfu - DEM
- TIN reprezentuje povrch jako **soubor trojúhelníků** (odtud trojúhelníková), které jsou definovány třemi body umístěnými **kdekoliv** v prostoru (odtud nepravidelná) a pro tyto trojúhelníky uchovává **topologické vztahy** (odtud síť).
- **Zmenšení** objemu uložených údajů při reprezentaci homogenních povrchů, **větší přesnost** a věrnost pro nehomogenní povrchy, struktura automaticky obsahuje informace o sklonu a směru tohoto sklonu.

TIN





A TIN can model a discontinuous surface. Each discrete land mass is represented with one hull and an integrated set of nodes, edges, and faces.



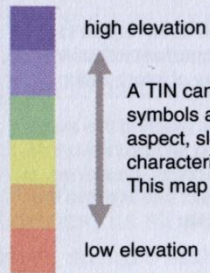
A lake can be modeled as an replace polygon applying the lake elevation for a constant height.

This area has a gentle slope without breaks in the slope. Few mass points are required here.

This region is steep with sharp changes in relief. More mass points are necessary to represent the local shape of this surface.

A shoreline for an island or land mass is modeled by adding an clip polygon. This polygon defines the hull, or outer edge of a set of faces. The hull defines a shore or a project boundary.

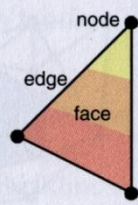
In this map, a stream course is modeled with a soft breakline. For features such as stream courses and ridge lines, you can choose either soft or hard breaklines depending on the effect you desire when drawing a TIN with smoothed interpolation of elevations.



high elevation

A TIN can be drawn with symbols applied to elevation, aspect, slope, or other characteristics of a surface. This map shows elevations.

low elevation

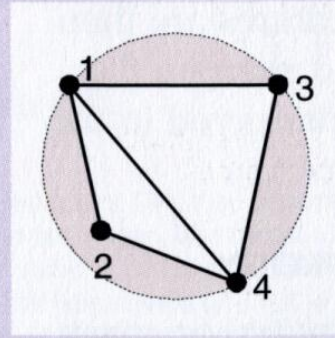
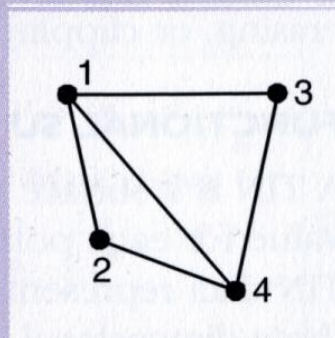


To illustrate the structure of a TIN, this map symbolizes nodes, edges, and faces. Normally, nodes and edges are not displayed on a map. In this map, faces are drawn with interpolated elevation ranges.

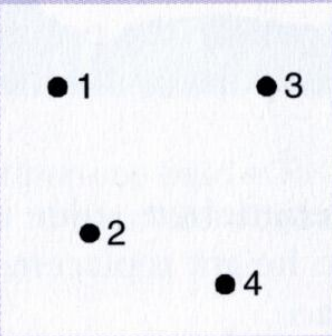
- Hard breaklines or polygons represent features that mark a significant change in slope.
- Soft breaklines or polygon represent surface features that do not mark a slope discontinuity, but add edges to a TIN to better model a surface.

Princip Delaunay triangulace

A Delaunay triangulation uses an algorithm to optimize the surface representation.

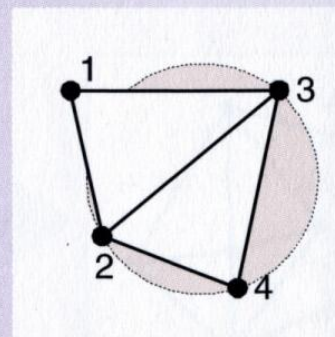
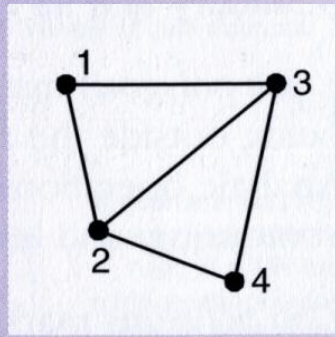


This triangulation fails the Delaunay test because the circle bounding nodes 1, 3, and 4 includes node 2.



From a simple case of four mass points, two triangulations are possible. Which is the valid TIN?

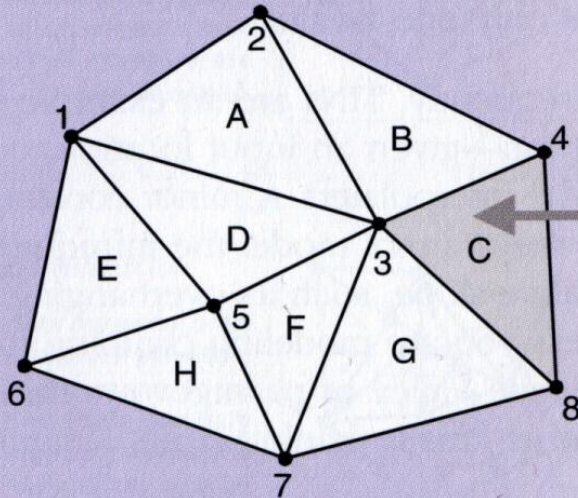
The definition of the Delaunay triangulation specifies that any circle around three nodes in a triangle will not include any other node.



This triangulation satisfies the Delaunay test because a circle around each triangle contains no other nodes. This is the valid triangulation.

Uložení TINu

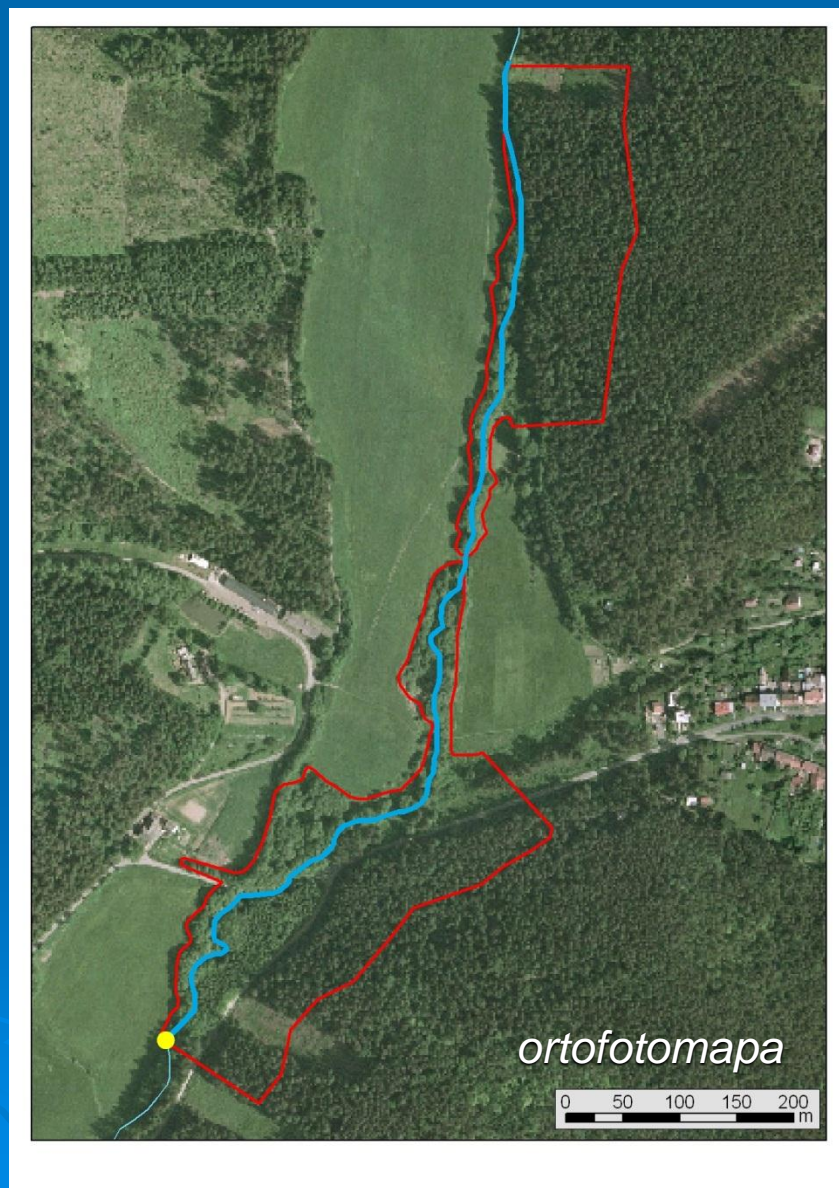
A TIN is a topological data structure that manages information about the nodes that comprise each triangle and the neighbors to each triangle.



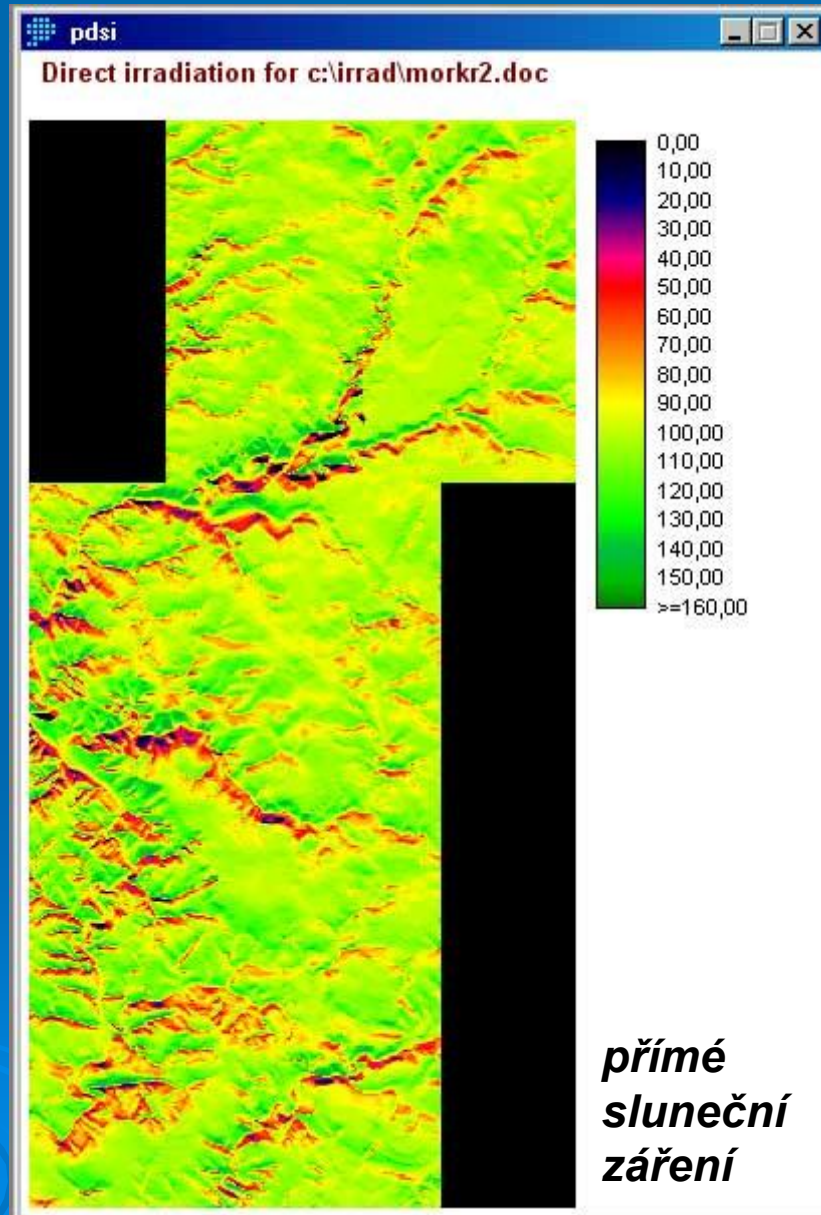
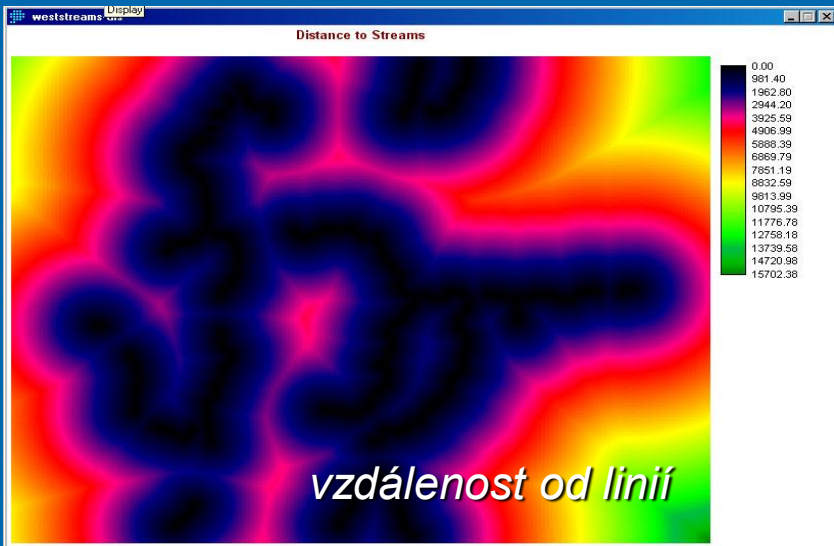
Triangle	Node list	Neighbors
A	1, 2, 3	-, B, D
B	2, 4, 3	-, C, A
C	4, 8, 3	-, G, B
D	1, 3, 5	A, F, E
E	1, 5, 6	D, H, -
F	3, 7, 5	G, H, D
G	3, 8, 7	C, -, F
H	5, 7, 6	F, -, E

Triangles always have three nodes and usually have three neighboring triangles. Triangles on the periphery of the TIN can have one or two neighbors.

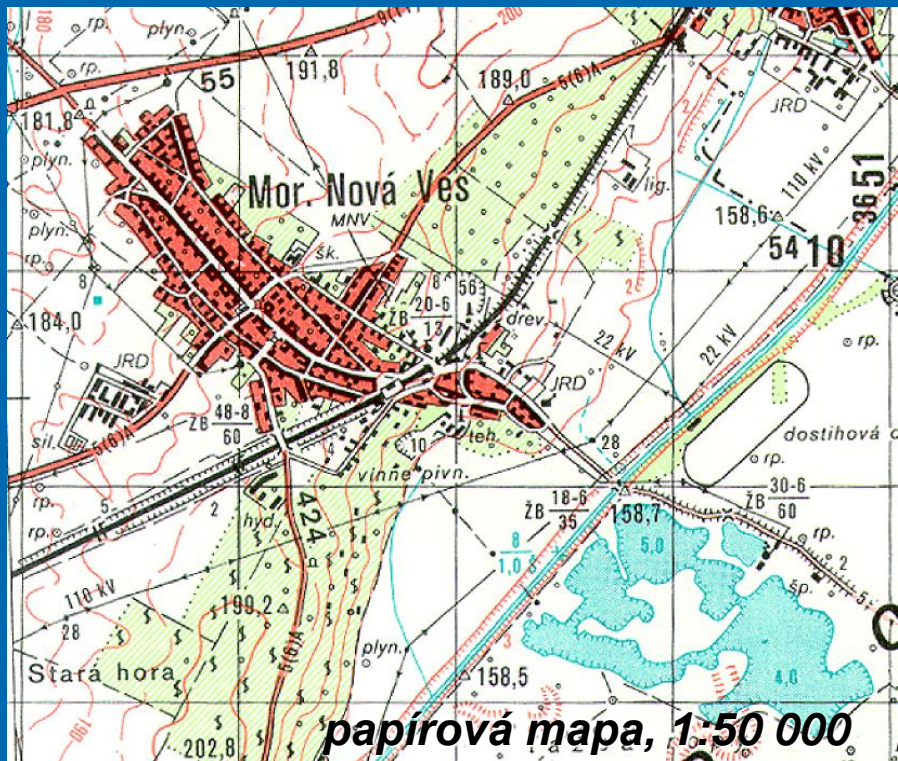
Různé formy rastrových dat



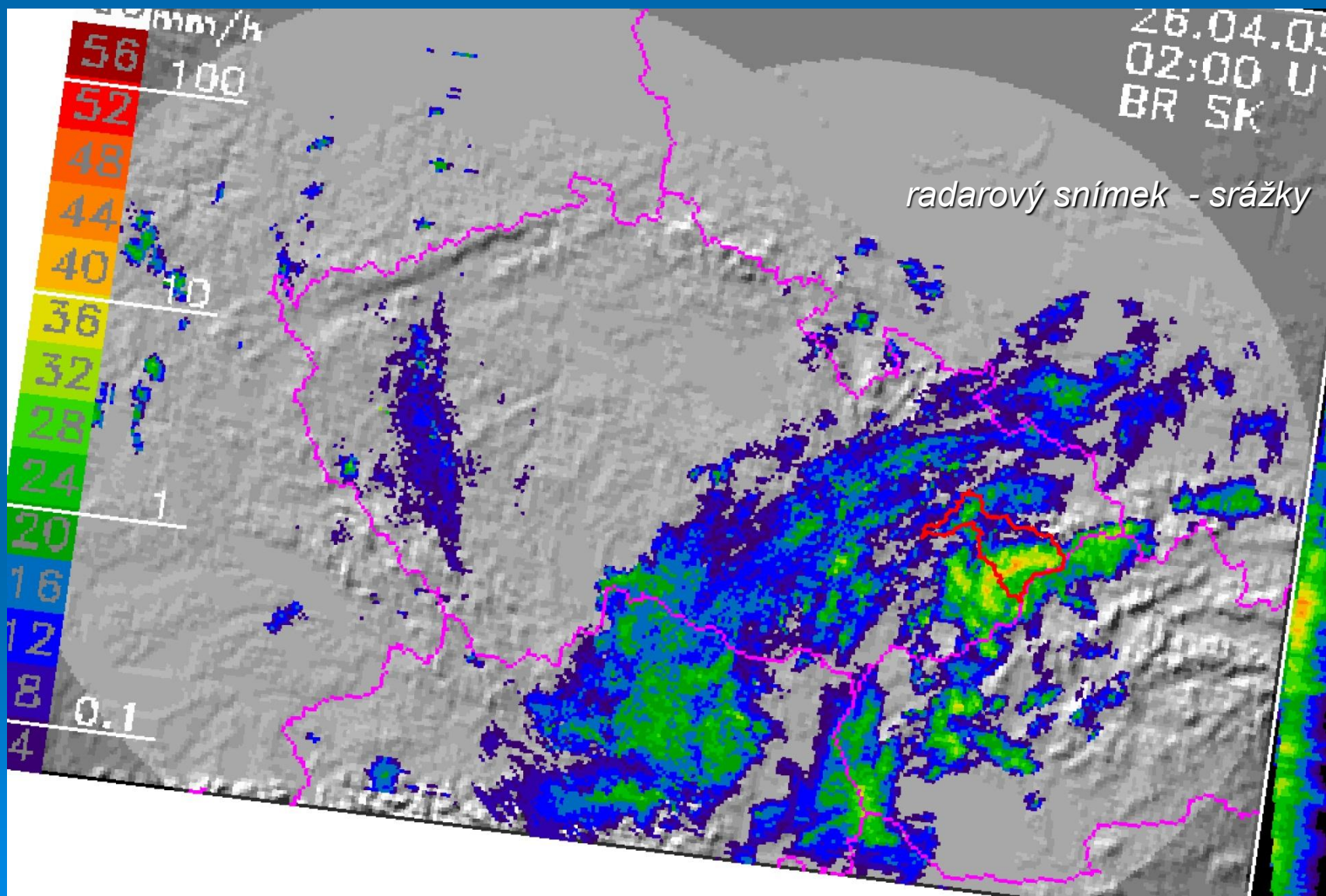
Různé formy rastrových dat



Různé formy rastrových dat



Různé formy rastrových dat



VEKTOR

Uložení dat

RASTR

Atributová data

- databáze
(R)DBMS (Relational)
database management
system

Polohová data

- nativní formát
 - ArcGIS: shapefile
***.shp**, coverage
 - Microstation: ***.dgn**
 - Topol: bloky ***.blk**

Atributová data

- DBMS nebo RDBMS nebo
bez atributových souborů

Polohová data

- speciální formát (většinou
komprimovaný)
 - obecné grafické (**tiff**,
jpeg, **bmp**)
 - softwarově specifické
(ArcINFO **grid**, Erdas
***.img**, Topol ***.ras**)

Komplexní formáty ukládají více vrstev různých typů

ArcGIS: geodatabase, Geomatica: ***.pix**

Výhody a nevýhody rastrů

- rychlé dotazování
 - jednoduchost datové struktury
 - jednoduchá kombinace s jinými daty rastrové povahy (DPZ)
 - snadné provádění analytických operací
-
- značná paměťová náročnost (velký objem dat)
 - omezená přesnost, daná rozlišením rastru a orientací rastru (výpočty délek, vzdáleností, ploch ...)
 - kvalita výstupů závislá na rozlišení rastru
 - nižší vizuální kvalita rastrových výstupů
 - nevhodnost pro síťové analýzy
 - náročná generalizace

Výhody a nevýhody vektorů

- **objektová orientace**, s jednotlivými objekty lze pracovat jako se samostatnými celky
- **menší náročnost na paměť**.
- dobrá reprezentace jevové struktury dat
- **kvalitní grafika**, přesné kreslení, znázornění blízké mapám
- **vysoká geometrická přesnost**
- **jednoduché vyhledávání**, úpravy a **generalizace** objektů a jejich atributů

- složitější zjišťování odpovědí na polohové dotazy a **obtížná tvorba překryvů** vektorových vrstev.
- problémy při náročných analytických operacích, **složitost výpočtů**
- problémy při **tvorbě modelů**, komplikované využití pro simulaci jevů

VEKTOR RASTR

	Vektor	Rastr
prezentace jevové struktury	dobrá (nelze spojité povrchy)	záleží na rozlišení (nevhodné pro liniové)
datová struktura	složitá	jednoduchá
kvalita grafiky	dobrá	záleží na rozlišení
topologie	ano	ne (jen sousednost buněk)
objem uložených dat	malý	velký
nároky na software	velké	malé
analýzy	složitější ale komplexnější	jednodušší ale některé neproveditelné (sítě) a nepřesné (plochy, délky)
transformace mezi souř. systémy	přesná	nepřesná (resampling)