

GIS

Geografické informační systémy



GIS – Prostorové modely

- Rastrový model
- Pravidelné, nepravidelné buňky
- Způsoby uložení
- Komprese dat

Rastrový model

Intuitivně chápání - množina elementů (obecně různého tvaru a velikosti), které na sebe navazují tak, že kompletně vyplní sledovanou plochu nebo prostor.

Jedná se o definování *hodnot* sledovaných útvarů i jevů v konkrétní poloze prostoru (2D, 2.5D, 3D).

Základní informační jednotkou je **prostorový element** - buňka svázána s hodnotami popisujícími stav tohoto elementu (atributy).

Obrazová data: letecké a satelitní snímky, pravidelně rozmístěné výškové informace, naskenované mapy.

Rastrová buňka

Spojitost plošných dat x prázdná místa s nulovými daty („no data”).

Buňka nese libovolnou hodnotu, obsah buňky je však homogenní.

Zachycuje geografickou plochu v závislosti na zvoleném rozlišení rastru (např. 12×12m). Zvolené rozlišení má dopad na informační obsah, popř. kvalitu dat.

Dvoudimenzionální rastrová buňka = „*pixel*” („Picture Element”)

Třídimenzionální buňka = „*voxel*” (Volume Pixel).

Rastrový model

Každá buňka je popsána pozičním indexem (uloženo číslo řádku a sloupce) a kódem vyjadřujícím hodnotu prvku.

Jednotlivé buňky obsahují hodnoty (values), které zastupují zkoumanou lokalitu.

Číslování většinou začíná v *levém horním rohu*.

Rastrový model

Na rozdíl vektorového modelu buňky svým tvarem neodpovídají krajinným prvkům. Uložení dat je provedeno pomocí imaginární sítě - rastru.

Základní rozdělení rastrových modelů je dáno tvarem buněk. Hovoříme pak o **pravidelném** nebo **nepravidelném** členění prostoru.

V případě pravidelného členění prostoru sledujeme velikost buněk.

Rastrový model

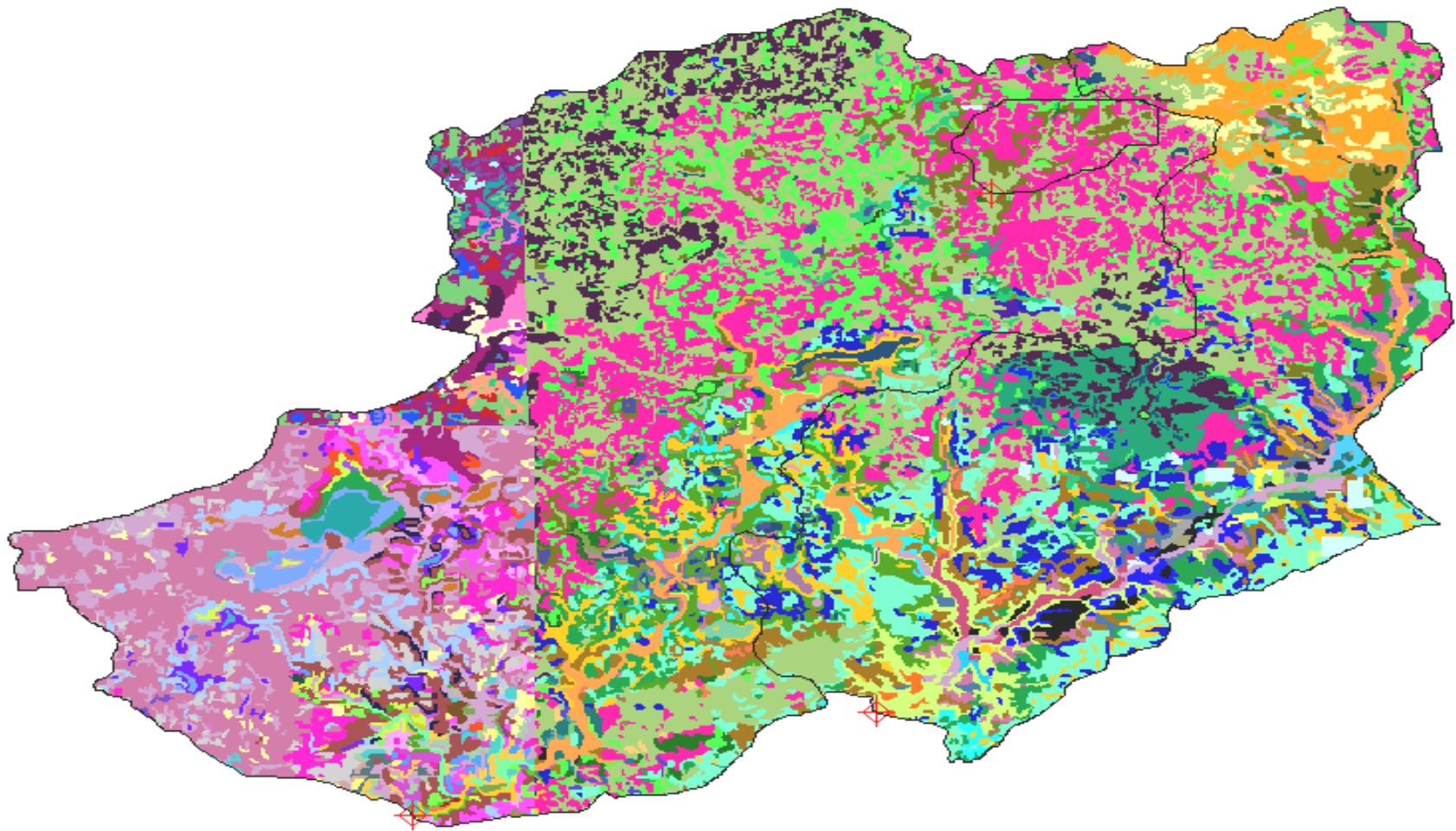
Má-li být pro každou buňku rozlišovací úroveň stejná, pak se použijí stejně velké buňky.
V případě různé rozlišovací úrovně (či hierarchického členění) dochází ke změně velikosti definovaným způsobem. Ten pak ovlivňuje způsob reprezentace.

Typ zpracovaných údajů

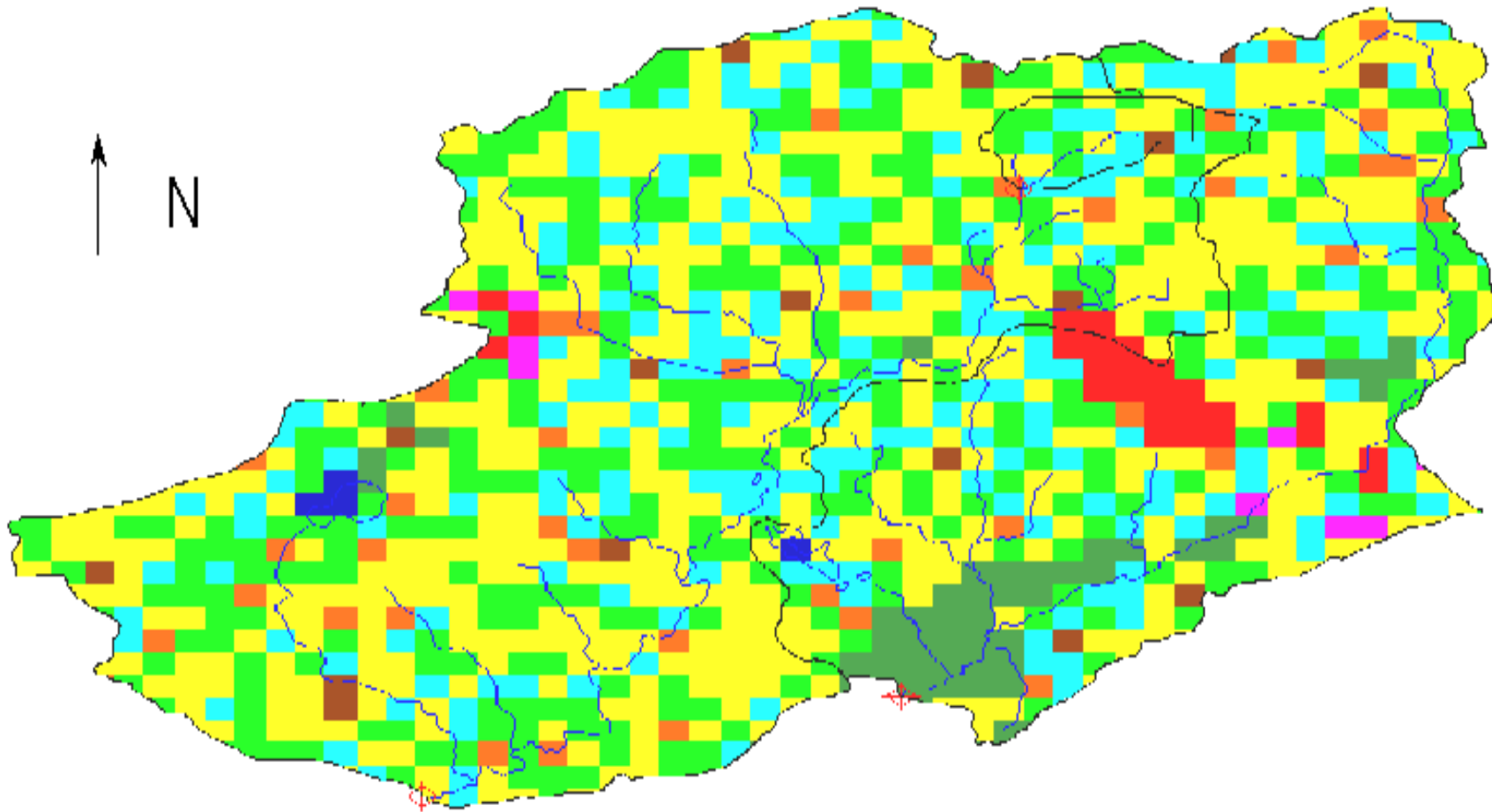
- Fyzikální data – srážky, ...,
- administrativní regiony, ...
- klasifikace půdy,
- družicová data,
- letecké snímky, ...

Údaje jednoho typu jsou uložena v **tématické vrstvě**.

Rastrový model



Rastrový model



Rastrový model - síť

Vlastnosti sítě - vliv na reprezentaci zobrazeného jevu – popisné složky:

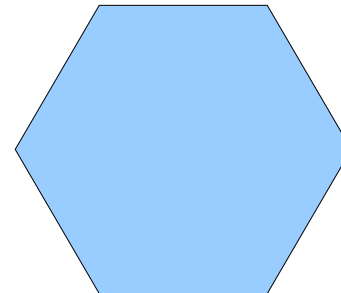
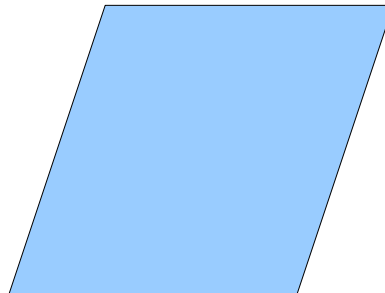
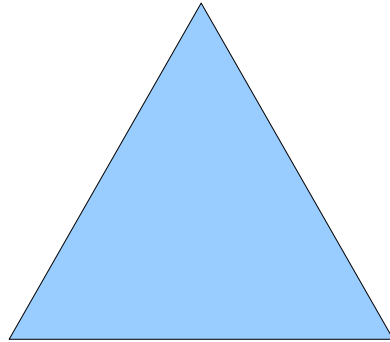
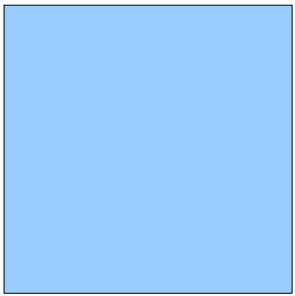
- rozměrnost sítě (2D, 2.5D, 3D)
- typ sítě – pravidelná, nepravidelná
- velikost sítě - počet buněk v jednotlivých dimenzích, případně celkový počet buněk
- způsob přiřazování hodnot vlastností buňkám
- interpolace v rámci sítě (použitý interpolační mechanismus v rámci sítě)

Podmínky pro tvar buněk

1. Nekonečná opakovatelnost buněk v prostoru.
2. V některých případech i nekonečně rekurzivně rozložitelná buňka na menší buňky stejného tvaru.

Při splnění první podmínky – reprezentace oblasti libovolné velikosti rastrem bez zbytku, druhá umožňuje použití hierarchické datové struktury.

Tvar buněk



Které tvary splňují obě podmínky?

Rastrový model - buňka

Popis vlastností buňky sítě

- pravidelné členění – pro všechny buňky stejné
- nepravidelné členění – každá buňka má svůj popis

Rastrový model

Hlediska popisu buňky

- tvar buňky – pravidelný, nepravidelný, nebo charakterizován vzorem
- velikost buňky – délka stran buňky či úhly sevřené hranicemi
- interpolace hodnot v rámci buňky – vztaženo k vrcholům, ...

Pravidelné dělení prostoru

Pravidelná buňka má vždy přesně definován stejný tvar:

- stejná délka všech hran,
- vnitřní úhly mají stejnou velikost.

Použití rozlišovací schopnosti reprezentaci

- se stejnou rozlišovací schopností,
- s odlišnou rozlišovací úrovní (velikost buněk se definovaně mění).

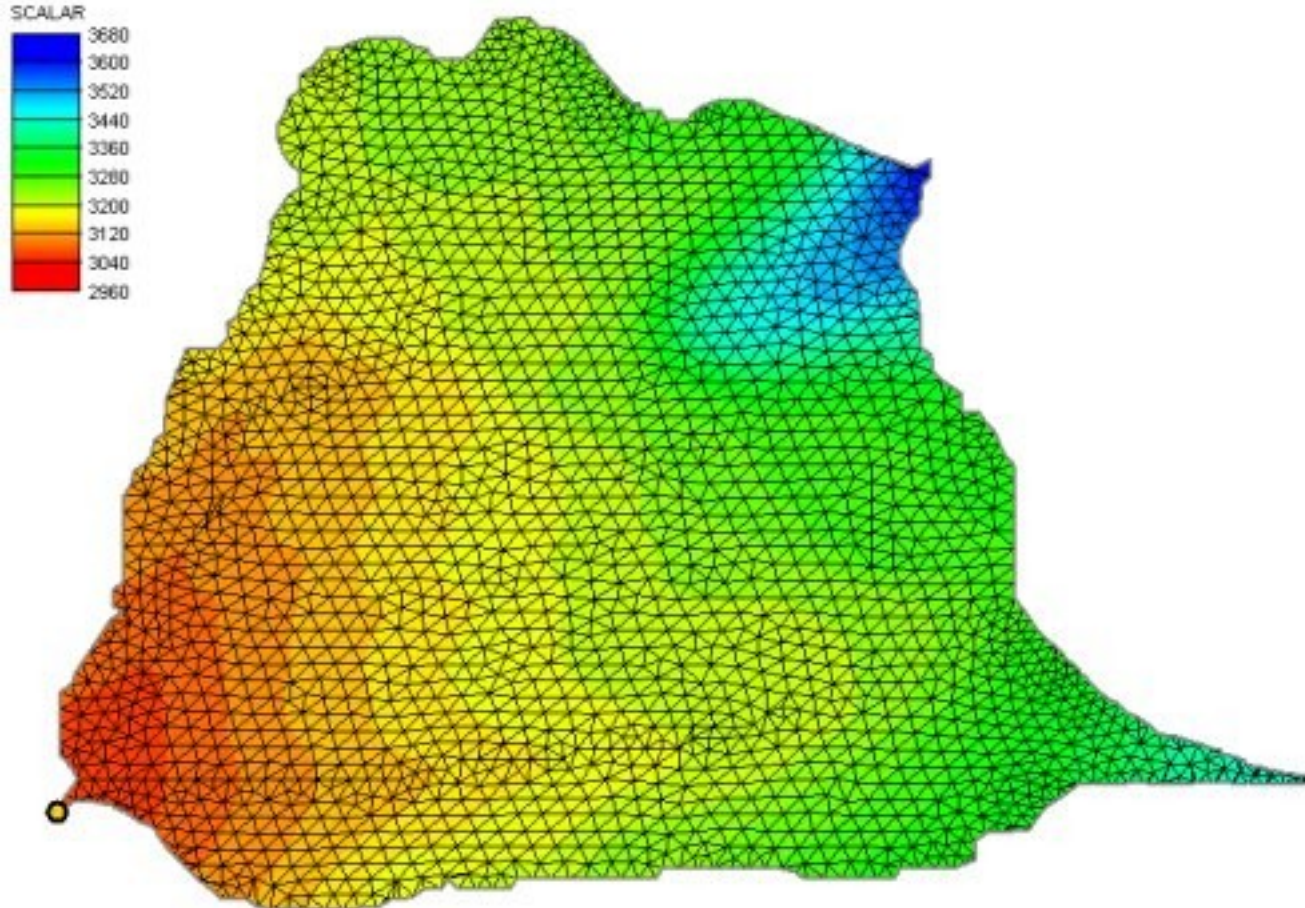
Nepravidelné dělení prostoru

Vytvářené buňky mohou mít různý tvar i velikost.

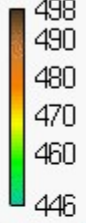
Používá se hlavně u nepravidelné trojúhelníkové sítě – TIN (triangulated irregular networks), kdy se u trojúhelníků využívá jejich různé orientace

Reprezentuje rovinu + kopce – tj.povrchy (2.5D).

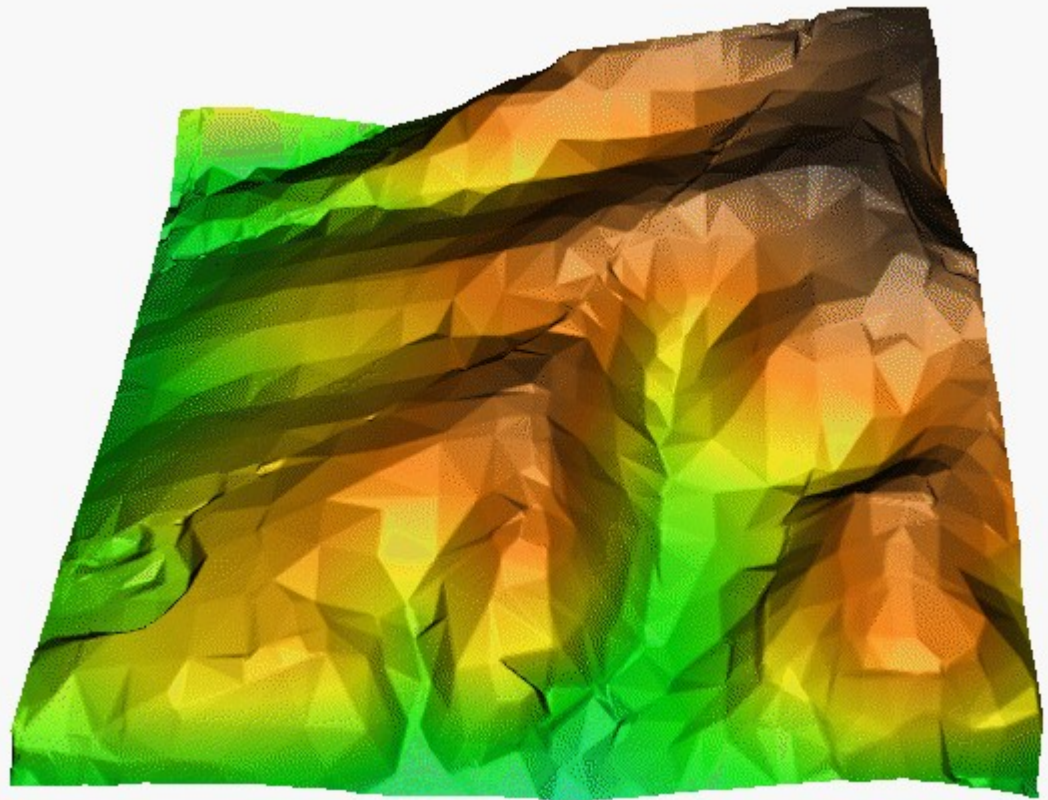
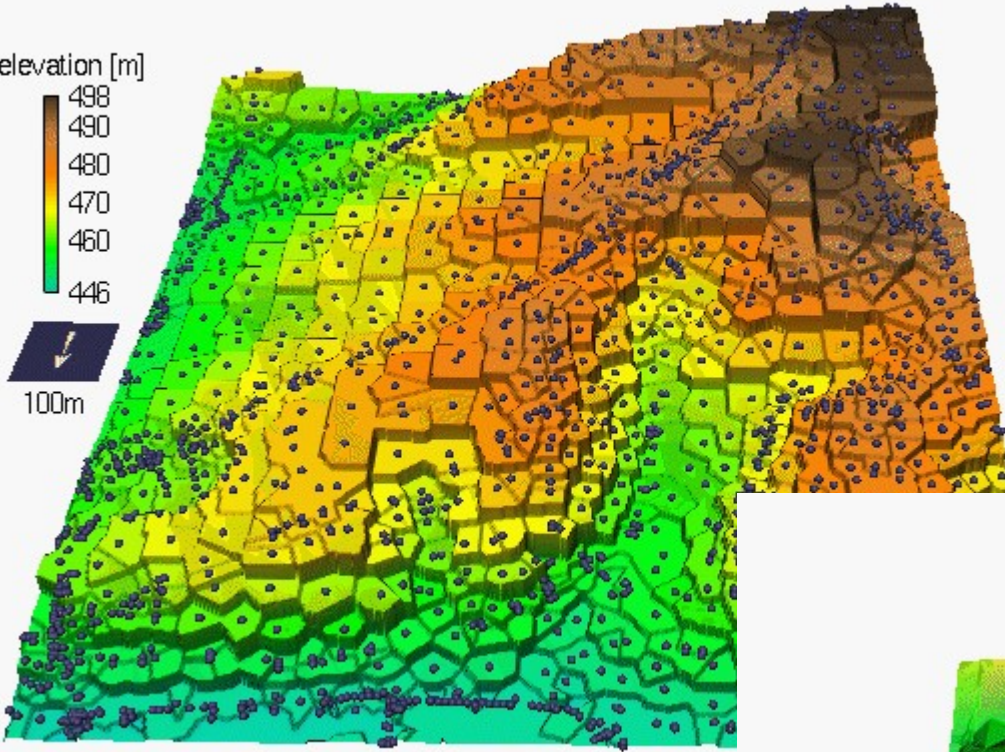
TIN - triangulated irregular networks



elevation [m]

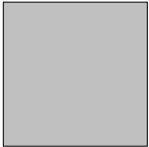
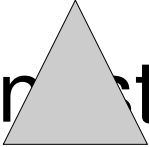
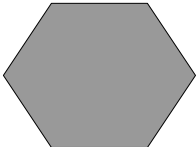


100m



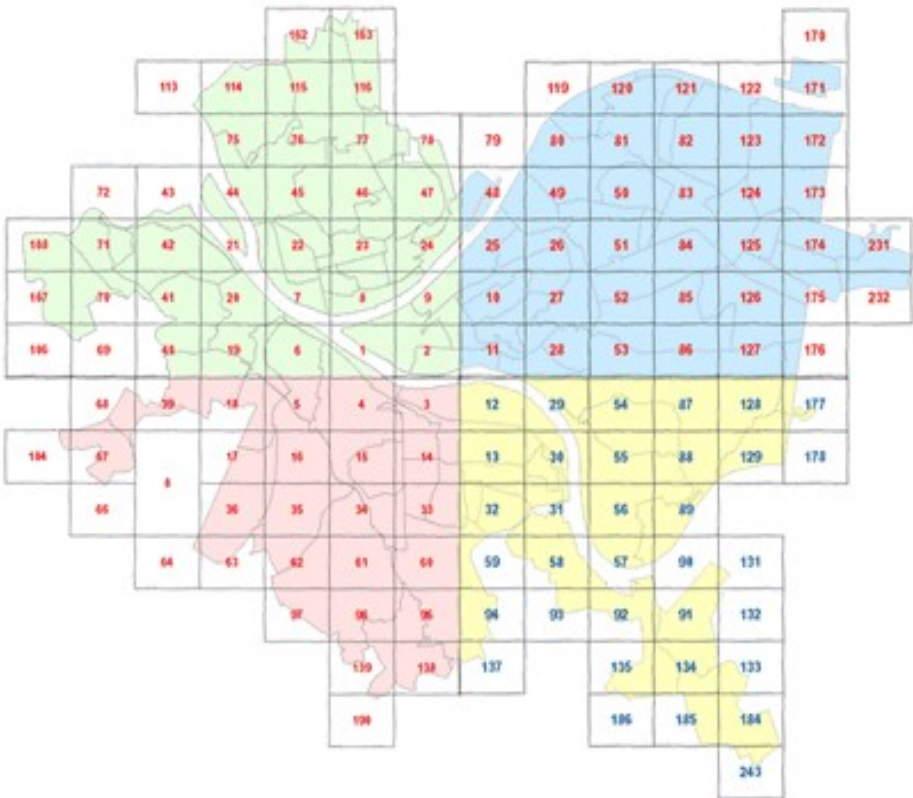
Pravidelné členění prostoru

Ve většině praktických aplikací se používá pravidelné dělení, kde se používají tvary buněk:

- *čtverec* (pixel) - gridový model (mříž), 
- *trojúhelník* (voxel) - nestejná orientace, 
- *šestiúhelník* (hexagonální buňka) - vzdálenost středů buněk je pro všechny buňky stejná. 

Pravidelné členění prostoru

Informace jsou uloženy v databázi, přičemž každá buňka je popsána pozičním indexem (číslo řádku a sloupce) a kódem vyjadřujícím hodnotu prvku.



Čtvercová mřížka

U pravidelného členění se nejčastěji používá čtvercová mřížka. Důvody použití:

- je kompatibilní s kartézským souřadnicovým systémem,
- 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).

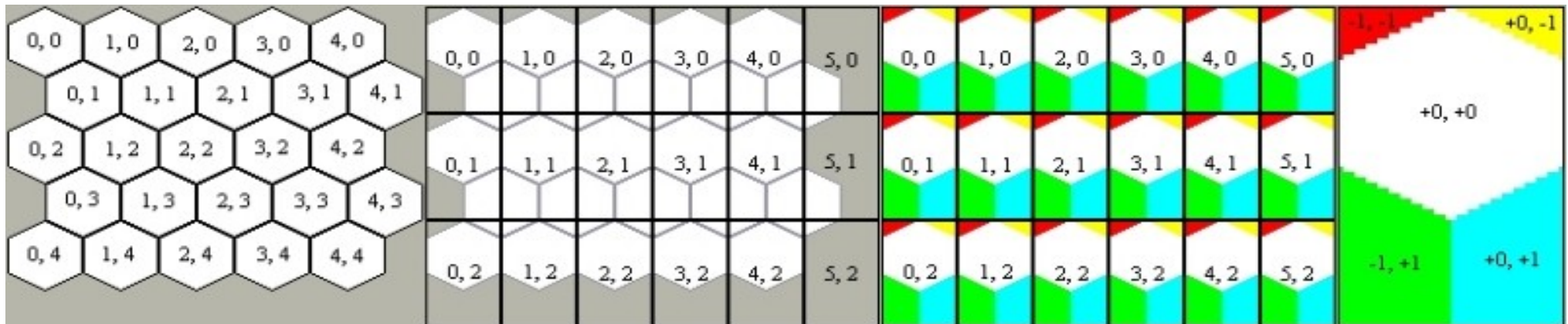
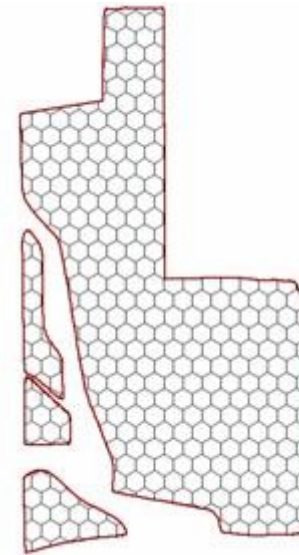
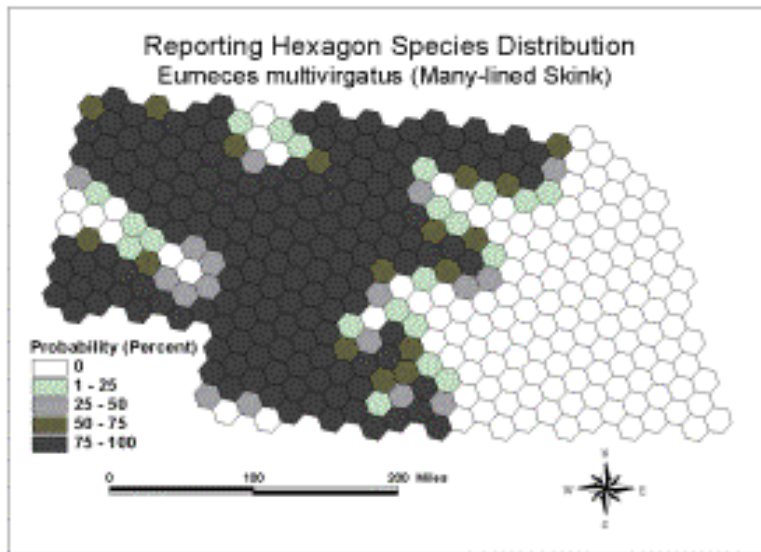
Hexagonální mřížka

Středy všech sousedních buněk jsou ekvidistantní (stejně od sebe vzdálené), což je výhodné pro některé analytické funkce – např. pro paprskové vyhledávání.

Ve čtvercové mřížce je toto nemožné a tato vlastnost se musí kompenzovat nebo se prostě zanedbává.

Hexagonální tvar buňky se používá zřídka, brání tomu:

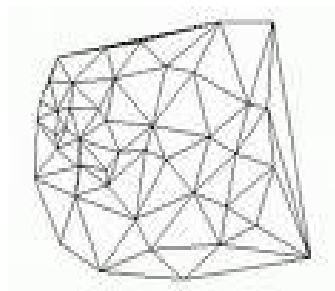
- připravenost hw zařízení na čtvercovou mřížku



Nepravidelná mřížka

Nepravidelné rastrové reprezentace se prakticky používají jen v několika určitých případech.

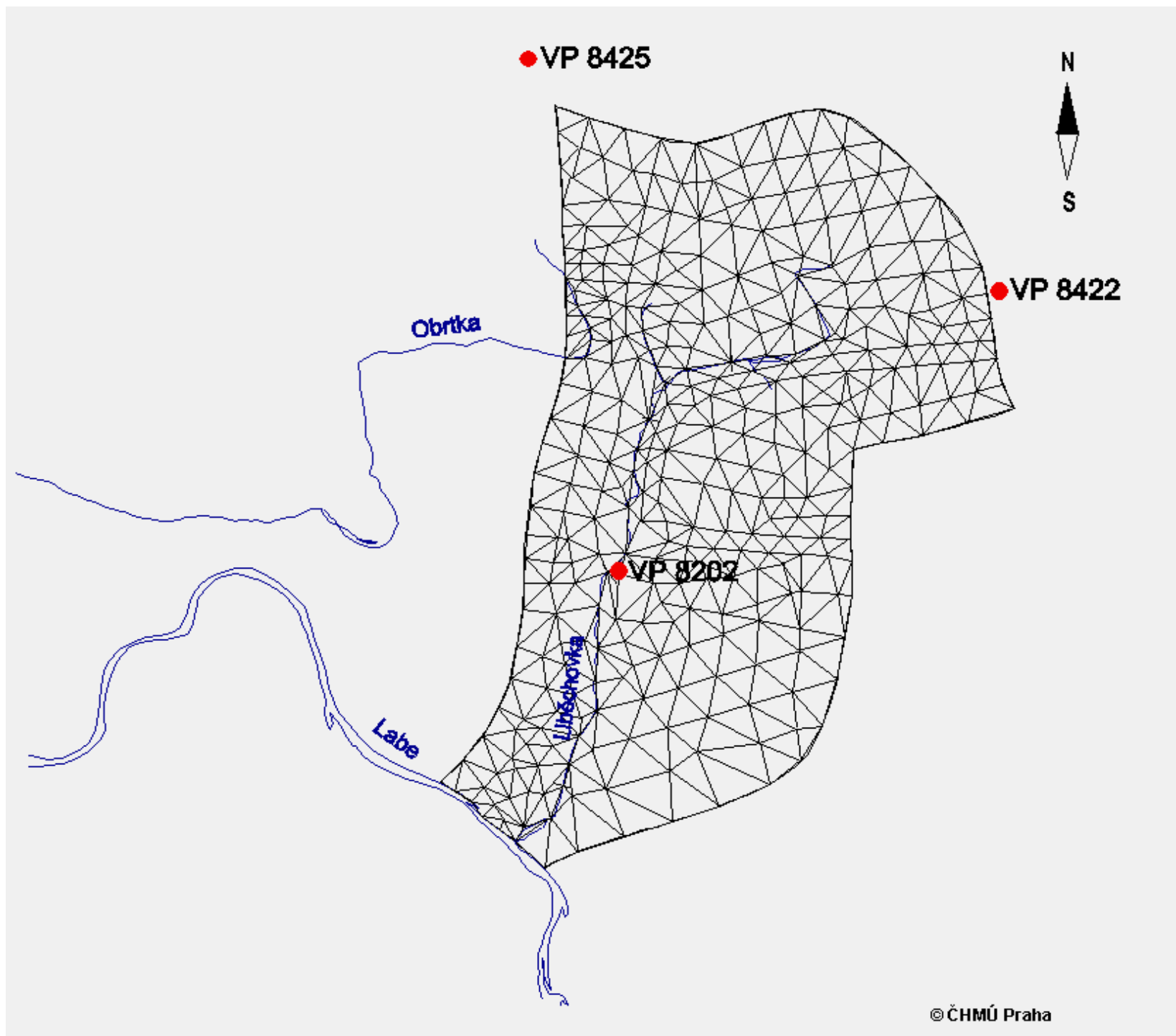
Všechny algoritmy (pro vytvoření, uložení i analýzy) jsou zde mnohonásobně složitější, než u pravidelných rastrů.



Nepravidelná mřížka

Nepravidelná trojúhelníková síť (Triangulated Irregular Network – TIN) představuje speciální případ použití.

Svou strukturou na pomezí mezi vektorem a rastrem a velice často se používá pro reprezentaci povrchů, například digitálního modelu reliéfu - DMR.



Modelová síť a poloha srovnávacích vrtů s dlouhodobým pozorováním, ČHMÚ

Trojúhelníková mřížka

Jednotlivé buňky nemají stejnou orientaci, což je právě výhoda při reprezentování digitálního modelu reliéfu (terénu).

Každému vrcholu o souřadnicích x,y je přiřazena funkční hodnota z (výška $z = f(x,y)$).

Jednotlivé trojúhelníky pak implicitně obsahují údaje o svém sklonu a směru tohoto sklonu. Právě tato vlastnost má za následek mnohem větší složitost všech algoritmů pracujících s tímto modelem.

TIN

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).

Vlastní proces vytvoření sítě se nazývá triangulace - náročný na výpočetní výkon.

TIN

V počítači je TIN uložena ve třech seznamech:

- seznam trojúhelníků,
- seznam hran těchto trojúhelníků,
- seznam souřadnic vrcholů těchto trojúhelníků.

Zhodnocení TIN

Výhody proti pravidelným rastrovým reprezentacím:

- zmenšení objemu uložených údajů při reprezentaci nehomogenních povrchů,
- větší přesnost a věrnost pro nehomogenní povrchy, struktura obsahuje informace o sklonu a směru sklonu,
- kompatibilita s moderními grafickými kartami (které podporují 3D zobrazení v reálném čase).

Nevýhody:

složitosť datové struktury a s tím spojených algoritmů.

Zhodnocení mřížek

Pravidelné mřížky jsou mnohem jednodušší pro ukládání a zpracování údajů, zabírají ovšem na disku mnoho místa.

Nepravidelné mřížky mohou mnohem lépe reprezentovat lokalitu (příklad roviny a zpracování ploch s výškopisem), jejich zpracovávání je však algoritmicky i výpočetně náročné.

Topologie je v rastrovém modelu definována implicitně (kdo je čí soused), není nutné ji explicitně ukládat jako pro vektorový model.

Vliv volby rozlišení

Při reprezentaci prostorových objektů rastrem hraje roli volba vhodného rozlišení - velikost pixelu. Nevhodná volba rozlišení:

- zbytečné ukládání mnoha dat na disk
- ztráta prostorových informací (o tvaru jednotlivých objektů).

Některé informace se mohou dokonce ztratit úplně!

Neurčené hodnoty

Případy, kdy buňka nenabývá žádné hodnoty (NoValue/NoData) se v počítači řeší zavedením speciální hodnoty NODATA s předem určenou hodnotou.

Například v ARC/INFO GRID (modul na zpracování rastrového datového modelu) je tato hodnota určena jako -9999. Tato hodnota má svůj speciální význam.

Rastrový model – struktura popisných údajů

Kódování popisných atributů znamená přiřazení jednoho kódu každé buňce i v případě více hodnot.

Rozlišujeme dva základní způsoby přiřazení hodnoty:

- hodnota je platná na větší „ploše“ buňky,
- hodnota je získána výpočtem (např. průměrná hodnota, maximální hodnota – vojenská mapa s výškopisem).

Rastrový model – popisné údaje

Binární rastr – ne/přítomnost atributu (0/1),
zaznamenání hodnoty v jednom bitu.

Osmibitový rastr – 256 různých celočíselných hodnot
sledovaného atributu, potřebujeme 1 bajt.

Dvacetičtyřbitový rastr – každou buňku je možné
ohodnotit 1.6 miliony různých celočíselných hodnot
sledovaného atributu, potřebujeme 3 bajty.

Kontinuální rastr - rozlišujeme XX množství reálných
různých hodnot sledovaného atributu – desetinná
čísla, používá se 4 nebo 6 bajtů.

Typy rastrových modelů – použití

Binární rastr – naskenované katastrální mapy.

Osmibitový rastr – naskenované barevné předlohy, panchromatické letecké a družicové snímky, případně produkty rastrových systémů při rastrových analýzách.

Dvacetičtyřbitový rastr – zpracování multispektrálních družicových snímků.

Problémové operace nad rastorvými daty

Prostorové rozlišení – parametrem je velikost buňky, což má vliv

- na přesnost umístění objektu,
- na tvar objektu.

Vztahy mezi objekty – implicitně umístěním v sousedících buňkách.

- zjištění jednoho polygonu – bez problému,
- zjištění obvodu tohoto polygonu - problém.

Metody komprese údajů u rastru

Pro uložení dat - metody komprese pro snížení kapacity pro ukládání dat beze ztráty informace (hlavně na disk). Vlastní zpracování v paměti – reprezentace dvourozměrnou maticí.

- Řádková komprese dat,
- metoda délkových kódů (RLE),
- Mortonovo pořadí,
- Peanovo pořadí,
- kvadrantový model (čtyřstrom – quadtree).

Kompresní techniky

Uložení pravidelných rastrových dat - náročné na prostor, vysoké režijní náklady.

Velikost dat záleží na rozloze představované oblasti a na rozlišení, ale nezávisí na obsahu.

Zmenšení náročnosti vede k zavedení účinného způsobu uložení (hlavně reprezentace řídce rozptýlených bodů).

Před zavedením kompresních technik pro operace využívající okolí zpracovávaného bodu musí být data nejdříve dekódována, teprve pak je lze zpracovat.

Metoda délkových kódů (RLE)

Run Length Encoding - RLE

Technika pro odstranění neefektivnosti uložení rastrových dat pomocí matic.

Principem je “stlačení” stejných opakujících se informací do jedné souhrnné.

Rozsáhlé homogenní objekty jsou reprezentované velkým množstvím pixelů (ukládá se hodnota každého bodu).

Metoda délkových kódů - RLE

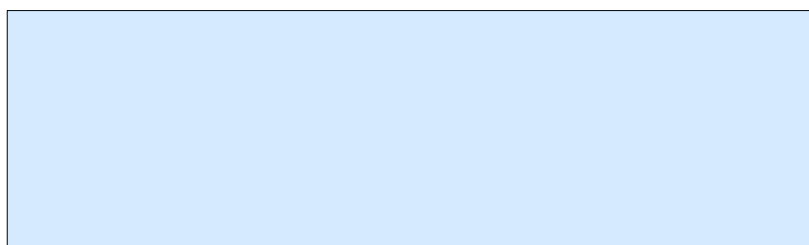
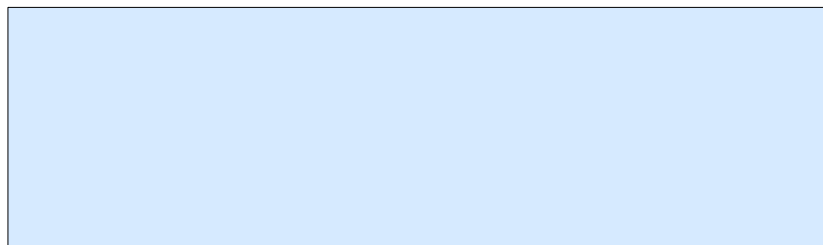
RLE organizuje řádek pixelů o stejné hodnotě do skupiny, kterou uloží jako počet stejných pixelů s danou hodnotou.

Výsledkem je datová struktura, kde původní data nahrazuje seznam dvojic $(n, v) = (\text{počet pixelů}, \text{hodnota pixelu})$.

Příklad:

9 9 9 9 9 9 1 1 1 1 5 5 9 9 9 (6 9)(4 1)(2 5)(3 9)

9 2 9 9 9 4 4 1 1 5 5 5 5 5 5 (1 9)(1 2)(3 9)(2 4)(2 1)(6 5)



Problémy RLE

Při častých změnách v průběhu dat dochází ke zdvojnásobení velikosti dat. Nejhorší případ - každý kódovaný úsek má velikost jeden pixel.

Například:

Z tohoto důvodu se používají současně s RLE i jiné techniky komprese.

Čtyřstrom (QuadTree)

Hierarchická rastrová struktura.

Využívá model "rozděl a panuj". Prostor je dělen rekurzivně do kvadrantů podle ne/homogenosti oblasti reprezentované příčným uzlem stromu. Ukončení rekurze je dáno zvoleným rozlišením, tj. velikostí pixelu.

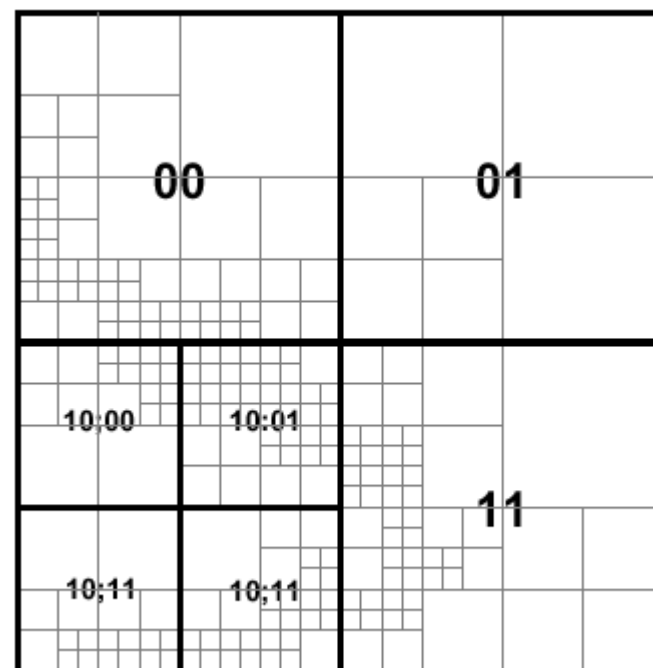
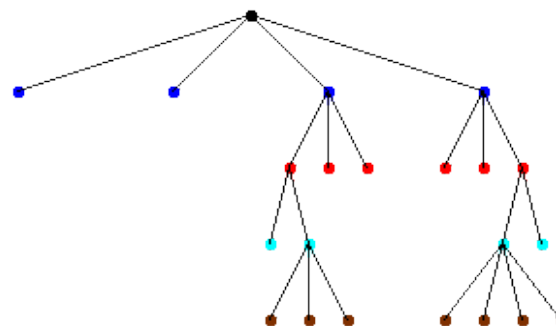
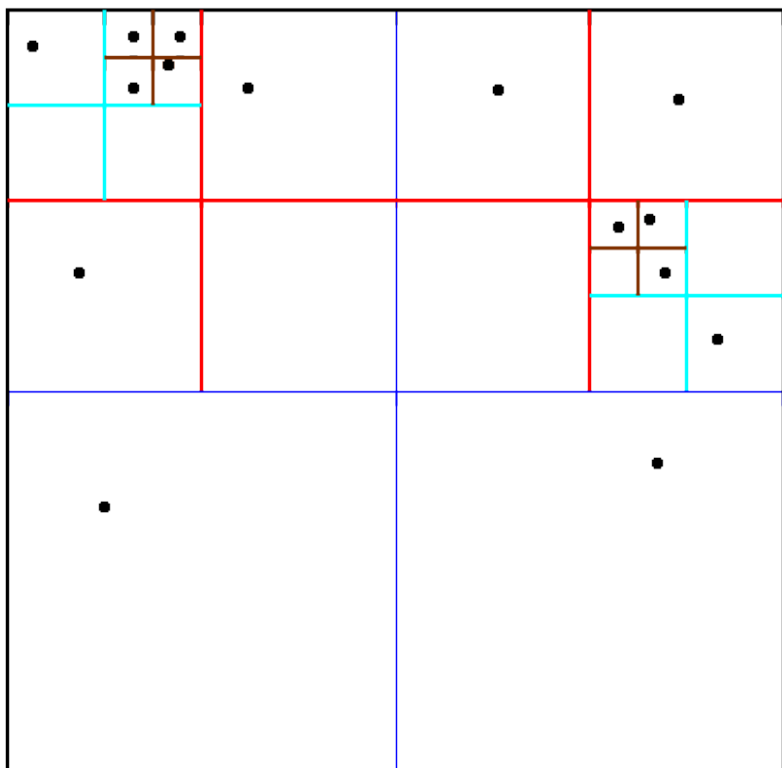
QuadTree

Struktura tak vytváří strom s uzly reprezentujícími heterogenní oblasti a listy oblasti se stejnou hodnotou.

Nevýhody QuadTree - struktura není invariantní s operacemi translace, rotace a změna měřítka.

To se týká všech hierarchických modelů reprezentace rastrových dat využívajících dělení prostoru.

Adaptive quadtree where no square contains more than 1 particle



Rozdělení rastrových dat

Rastrová data mohou reprezentovat hodnotu buď pouze pro střed buňky (lattices) nebo pro celou oblast buňky (grids).

Z hlediska obsahu lze rastrová data dělit na data reprezentující jeden jev - "*klasické rastry*" a na data *obrazová*.

Obrazová data

Snímek dálkového průzkumu Země (DPZ) -
vícepásmové obrazy (snímky z družice Landsat)
používané pro vyhodnocení polohy jednotlivých
geografických prvků v krajině (vodstvo, lesy, silnice, ...).
Takto vyhodnocené vrstvy jsou pak uloženy do
"klasického rastru", případně vektoru.

Ortofoto - ve většině případů používáno jako
jednopásmový obraz pro vyhodnocení polohy
jednotlivých geografických prvků v krajině, využívají se
jako podkladová data pod tématické vrstvy.

Obrazová data

Scannované plány - podkladová data pod
tématické vrstvy, musí být georeferencovány.

Dokumentace - nereferencovaná obrazová data,
sloužící pouze jako archivní dokumentace
projektu.