

# Metodika vymezení urbánního území na základě klasifikace dat dálkového průzkumu Země

MET.01.2019.07.31

CENIA, česká informační  
agentura životního prostředí



**Autoři:**

Jana Seidlová, Martin Hofrajtr, Zbyněk Stein, Lenka Rejentová, Jiří Kvapil, Tereza Kochová

**Oponenti:**

Tomáš Řezník, Přemysl Štych

Vytvoření metodiky bylo podpořeno z Dlouhodobé koncepce rozvoje výzkumné organizace CENIA, české informační agentury životního prostředí, na období 2018–2022.

CENIA, česká informační agentura životního prostředí

Praha, 2019

## Obsah

Úvod .....	4
Cíl metodiky .....	5
Metodický postup vymezení urbánního prostředí.....	5
Využití klasifikace obrazových dat.....	5
Metodika vymezení urbánního území na základě řízené klasifikace .....	7
Uplatnění metodiky v CENIA .....	8
Ekonomické aspekty metodiky.....	9
Závěr.....	9
Seznam literatury a zdrojů .....	10
Přílohy.....	11

## Úvod

V obcích a městech se v současnosti koncentrují téměř dvě třetiny veškerého obyvatelstva ČR (ČSÚ, 2018). Jedná se o místo centralizující služby, a také oblast, kde dochází k transformaci a výměně energetických toků a materiálů. Obce a města jsou proto velmi zranitelné vůči projevům změny klimatu. Pro potřeby adaptace a také mitigace vůči těmto projevům je nezbytné vymezit zastavěné urbánní území (městské prostředí), kde se většina těchto aktivit odehrává.

Předkládaná metodika je dílem autorského kolektivu CENIA, které vzniklo prvotně pro potřeby hodnocení systému zranitelnosti ČR vůči projevům změny klimatu v městském prostředí, a to z důvodu neexistence aktuálně dostupné, použitelné a v čase opakovatelné metodiky (více viz Technický list). Podkladem pro zpracování této metodiky byla jak rešerše literatury, tak reálné vymezování urbánního prostředí dle vlastních přístupů.

## Cíl metodiky

Cílem předkládané metodiky je nutnost vymezit urbánní území v rámci administrativní jednotky sídel ČR tak, aby bylo možné identifikovat dílčí fenomény odehrávající se v městském prostředí a potřebné analýzy pravidelně opakovat, a také aby vymezení samotné bylo dostatečné pro potřeby dalšího plánování (dostatečná velikost měřítka).

Hlavními cíli metodiky jsou:

- Vymezit vrstvu urbánního území v rámci administrativního členění sídel ČR na základě vydefinovaných tříd povrchu (obsahujících zástavbu, stromy, nízkou vegetaci a vodstvo).
- Vymezit urbánní území ve všech sídlech ČR v různých časových horizontech.
- Využít metodiku při mezinárodním srovnávání.

## Metodický postup vymezení urbánního prostředí

### Využití klasifikace obrazových dat

Klasifikace obrazových dat představuje širokou škálu digitálních metod založených na automatizovaném počítačovém vyhodnocení. Klasifikace družicových snímků využívá jejich spektrální pásma. Při klasifikaci se rozlišují jednotlivé třídy krajinného pokryvu. Existují dvě metody klasifikace snímků – řízená a neřízená. Dalším dělením může být to, zda je objektem, který je přiřazen do tříd, samostatný pixel (per-pixel neboli pixelové klasifikace), nebo skupina pixelů s podobnými vlastnostmi (objektově-orientovaná klasifikace). Klasifikační postupy lze rozdělit do několika kategorií podle různých kritérií (Lillesand, Kiefer and Chipman, 2008).

Řízená klasifikace se skládá ze dvou základních fází – trénovací, fáze učení zahrnující výpočet statistických charakteristik jednotlivých tříd, a klasifikační. Klasifikace probíhá tak, že algoritmus rozpoznává předem definované klasifikované třídy na základě tzv. trénovacích množin, které jsou ručně definovány (Ravindranath and Ostwald, 2006). Algoritmus trénovacích množin potřebuje informaci, která skupina představuje např. jehličnatý les, listnatý les, vodu, zástavbu. Na základě těchto informací se pokouší zařadit všechny části snímku do stanovených tříd (Dobrovolný, 1998). Trénovací plochy pro jednu třídu by nikdy neměly zahrnovat několik spektrálních tříd – například pro zemědělskou půdu je nutné nadefinovat několik podtříd, které budou rozlišovat různé zemědělské plodiny na polích, různou vlhkost, typ půdy atd. Také se v reliéfu ovlivněných scénách musí brát ohled na osluněné a zastíněné svahy, pro stejnou třídu pak definovat jednu kategorii na slunci a druhou ve stínu, a v poklasifikačních úpravách je teprve spojit do jedné informační třídy. Řízená klasifikace také často vyžaduje i terénní mapování. Klasifikační stadium je již automatická práce počítače, který zařadí či nezařadí jednotlivé pixely do předem vybraných tříd a vytvoří tak zcela nový obraz v digitální podobě s novými funkčními hodnotami (Kolář a kol., 1997). K rozřazování pixelů do jednotlivých tříd je využíváno různých klasifikátorů, nejčastěji na základě statistiky, ale nověji i na bázi umělé inteligence a strojového učení. Jsou to klasifikátory minimální vzdálenosti (Minimum Distance), pravouhelníků (Parallelepiped) a maximální pravděpodobnosti (Maximum Likelihood), viz Obr. 2.

Klasifikátor minimální vzdálenosti středů shluků je velmi jednoduchý, avšak není příliš citlivý na různý stupeň rozptylu ve spektrálním chování dat, proto není pro velkou část úloh vhodný. Funguje na principu porovnávání vzdáleností mezi určeným pixelem a mezi středy shluků ve vícerozměrném

prostoru (tzv. centroidy), což jsou vypočítané průměry ze spektrálních hodnot pro každou třídu a každé pásmo. Určovaný pixel se pak podle pravidla o nejmenší vzdálenosti zařadí do toho shluku, k jehož centroidu má nejbližší. Klasifikátor minimální vzdálenosti středů shluků lze vylepšit zadáním hodnoty mezní vzdálenosti, tj. maximální vzdálenosti od centroidu, kdy je ještě pixel do daného shluku zařazen. Pokud je tato mezní vzdálenost od centroidu větší než definovaná mezní vzdálenost, pixel není zařazen do žádného ze shluků a zůstává nezařazený (Dobrovolný, 1998).

Klasifikátor pravoúhelníků bere pro každou třídu a každé pásmo minimální a maximální hodnotu ze všech pixelů, které jsou pro danou třídu natrénovány. Vedením rovnoběžných linií pixely s minimální a maximální hodnotou v každém pásmu vzniknou čtyřúhelníky (v případě dvou pásem) či pravoúhelníky (v případě více pásem), které budou vymezovat oblast, uvnitř které budou všechny pixely patřit do jedné třídy. Pixely mimo takto vymezené oblasti zůstanou nezařazené. Jako neklasifikované zůstávají též pixely, které se nacházejí na překryvu pravoúhelníků. Pixely v těchto oblastech pak musejí být rozřazeny v druhém kroku pomocí jiného rozhodovacího pravidla. Jinou možností pro zařazení těchto sporných pixelů do správné třídy je stanovení priority jednotlivých tříd. Existují i modifikace klasifikátoru, např. definování pravoúhelníků určitým násobkem směrodatné odchylky či nahrazení jednoho jednoduchého pravoúhelníku několika menšími, které lépe vystihují tvar shluků (Dobrovolný, 1998).

Klasifikátor maximální pravděpodobnosti pracuje na principu výpočtu pravděpodobnosti, s jakou klasifikovaný pixel spadá do určité třídy. Vychází z předpokladu, že shluky pixelů v trénovacích datech mají normální rozdělení. Poté při zatřídování každého pixelu klasifikátor maximální pravděpodobnosti kvantitativně hodnotí rozptyl hodnot, kovarianci i korelaci každé třídy. Hodnoty stejných pravděpodobností tvoří ve vícerozměrném prostoru izoliny, které mají často tvar elipsy. Klasifikovaný pixel se zařadí do třídy, do které s největší pravděpodobností spadá. Lze také nastavit prahovou hodnotu pravděpodobnosti, a pokud maximální pravděpodobnost zařazení pixelu je nižší než tato prahová hodnota, pixel zůstane neklasifikován. Klasifikátor maximální pravděpodobnosti je velmi citlivý na případné nedostatky v trénovacích datech, ale pokud jsou trénovací data kvalitní, dává skvělé výsledky. Vylepšením klasifikátoru maximální pravděpodobnosti je Bayesovský klasifikátor, který navíc pracuje s různými váhami tříd. Váhou může být pravděpodobnost výskytu každé třídy v daném území (tj. třídy běžné budou mít větší váhu než třídy, které se na území vyskytují vzácně), nebo třeba „cena“ za špatné zatřídění pixelu (tj. nejvyšší váhu dostanou ty třídy, na jejichž správném určení nejvíce záleží). Klasifikátor maximální pravděpodobnosti předpokládá stejnou pravděpodobnost výskytu pro všechny třídy, Bayesovský klasifikátor pak tyto stejné pravděpodobnosti mění na různé dle zadaných vah (Dobrovolný, 1998).

Neřízená klasifikace je založena na předpokladu, že každý krajinný pokryv má charakteristické spektrální vlastnosti a že stejné typy povrchu mají stejné či velmi podobné spektrální odezvy (Lillesand, Kiefer and Chipman, 2008). Na základě spektrální podobnosti jsou pak pomocí určitých algoritmů pixely slučovány do shluků, které jsou označovány jako spektrální třídy (tj. třídy objektů od sebe spektrálně rozlišitelných). Tyto spektrální třídy ovšem ještě nemají požadovanou informační hodnotu, a proto je následně ručně přiřazen tematický obsah. Největší výhodou neřízené klasifikace je to, že nevyžaduje prakticky žádnou znalost klasifikovaného území a před počátkem procesu nemusí být přesně definován počet tříd krajinného pokryvu jako u klasifikace řízené, takže se nemůže stát, že by bylo na nějakou třídu zapomenuto (Kolář a kol., 1997). Ale i tak je neřízená klasifikace nepřesná a výsledky nemusí být interpretovatelné.

Dalšími typy klasifikací jsou např. klasifikace typu per-pixel. Per-pixel má postupy založené na postupném zařazování jednotlivých obrazových pixelů do klasifikovaných tříd. Patří sem algoritmy pracující s každým pixelem izolovaně, nebo složitější postupy zahrnující do rozhodování o každém

pixelu i jeho okolí (Pham, Prakash and Tien Bui, 2018). Objektově orientovaná klasifikace využívá moderní metody založené na primární segmentaci dat do obrazových objektů a jejich následné klasifikaci (Lillesand, Kiefer and Chipman, 2008). Multispektrální klasifikace pracuje tak, že klasifikační algoritmy využívají spektrálních příznaků klasifikovaných tříd (Harris, 2018). Texturální klasifikace má klasifikační algoritmy, které využívají texturálních příznaků klasifikovaných tříd (Pham, Prakash and Tien Bui, 2018). Hybridní klasifikace má postupy umožňující zahrnout do rozhodovacího procesu různé typy příznaků (např. spektrální, texturální, kontextuální, geometrické) a další informační vrstvy (např. digitální model terénu, tematické mapy a vrstvy GIS). Do této kategorie se řadí výše zmíněný postup objektově orientované klasifikace (Lillesand, Kiefer and Chipman, 2008).

Přesnost klasifikace určuje, jak dobře se algoritmu podařilo zařadit pixely (při pixelové klasifikaci) či objekty (při objektové klasifikaci) do stanovených tříd. Při dobrém výsledku jde o přesnost 85–90 %. Souvislost s klasifikací má nejmenší mapová jednotka (minimum mapping unit MMU). MMU je nejmenší velikost fotografie nebo snímku, která odpovídá požadavkům na tematické rozlišení, bezproblémovou digitalizaci a ekonomické náklady. Určení MMU definuje množství detailů zachycených v procesu interpretace obrazu (St-Louis et al., 2014).

## Metodika vymezení urbánního území na základě řízené klasifikace

---

Základem metodiky pro vymezení urbánního území jsou multispektrální družicové snímky od družice Sentinel-2 (případně pro potřeby vyhodnocení dlouhé časové řady je možné použít i družicové snímky družice Landsat)<sup>1</sup>. Pro úpravy a další zpracování družicových snímků byl použit software ENVI a aplikace ArcMap. Byla využita všechna spektrální pásma multispektrálního snímku družice Sentinel-2 s rozlišením 10 m nebo družice Landsat s rozlišením 30 m. Požadavkem pro toto vymezení jsou snímky bez oblačnosti a sněhu nad zájmovým územím.

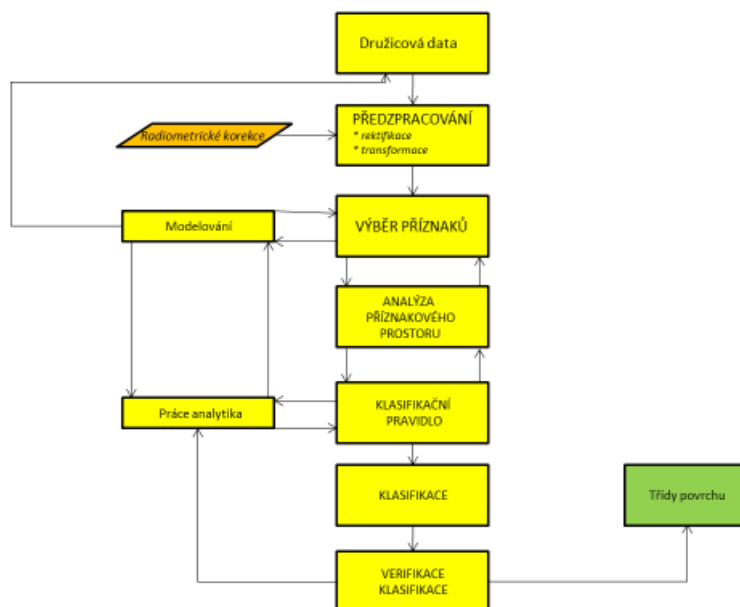
Před použitím klasifikace se družicové snímky musejí upravit z důvodu odrazivosti a atmosférických vlivů. Tři nejčastější radiometrické korekce jsou záření, odrazivost horní části atmosféry a odrazivost povrchu. Výpočty předpokládají, že obrazy byly kalibrovány na odrazivost povrchu. Odrazivost (nástroj Radiometric Calibration – Reflectance) poskytuje rychlý způsob, jak opravit atmosférické efekty. Kalibrace snímků na odrazivost povrchu také zajišťuje konzistenci při porovnávání indexů s časem a od různých senzorů. Korekce atmosférických vlivů (nástroj QUAC) dále eliminuje účinky rozptylu atmosféry a absorpci plynů, které způsobují chyby u vytváření dat při odrazu povrchu. Družicový snímek se nejprve upraví nástrojem Reflectance a poté se upraví nástrojem QUAC.

Upravené družicové snímky se ořízly o vrstvu mraků, z důvodu chybovosti při zpracování. Vytvořily se trénovací množiny dle vizuální interpretace pro čtyři třídy povrchu (Zástavba: vše uměle vytvořené, Stromy: vegetace s výškou nad 1 m, Nízká vegetace: vegetace s výškou pod 1 m, Vodstvo: voda) a spustila se vybraná řízená klasifikace maximální pravděpodobnosti (Maximum Likelihood) s přesností 95 %. Tato klasifikace předpokládá, že statistiky pro každou třídu v každém pásmu jsou normálně distribuovány, a vypočítá pravděpodobnost, že daný pixel patří do určité třídy (každý pixel je přiřazen třídě, která má nejvyšší pravděpodobnost (tj. maximální pravděpodobnost)). Území se klasifikovalo na čtyři kategorie povrchu, tj. na zástavbu, nízkou zeleň, stromy a vodstvo viz Obr. 1. Klasifikované území se ořízlo (nástroj Extract by Mask) na velikost administrativní hranice vybraných měst<sup>2</sup> a zanechala se pouze třída zástavby.

---

<sup>1</sup> Snímky území ČR z družice Sentinel-2 jsou dostupné na adrese webové stránky <https://scihub.copernicus.eu/> (Landsat -<https://earthexplorer.usgs.gov/>).

<sup>2</sup> Města nad 20 tisíc obyvatel.



Obrázek 1: Schéma klasifikace

Vrstva zástavby se převedla z rastru na polygon (nástroj Raster to polygon) a upravil se souřadnicový systém<sup>3</sup>. Na vrstvě zástavby se vytvořila čtvercová síť 100 m x 100 m (nástroj Fishnet). Tyto dvě vrstvy se propojily (nástroj Identity) a smazaly se prázdné čtverce ("gridcode" = 0 AND "area" = 10 000). Nová vrstva se zcelila (nástroj Merge) a odstranily se drobné polygony i díry ve vrstvě (nástroj Eliminate Polygon Part) do velikosti rozlohy nad 40 000 m<sup>2</sup>. Vznikla tím výsledná vrstva vymezení urbánního území vybraných měst.

## Uplatnění metodiky v CENIA

Navržená metodika vymezení urbánního území je nutná pro zpracování pravidelných úkolů v CENIA. Navržená metodika představuje dlouhodobě používané metody zpracování prostorových dat, zahrávané jejich novým a aktuálním využitím v resortu životního prostředí.

CENIA pravidelně zpracovává a hodnotí různé indikátorové sady. Zejména se jedná o sadu indikátorů životního prostředí a sadu indikátorů zranitelnosti, dále se pak CENIA podílí na zpracování a naplňování indikátorové sady udržitelného rozvoje a indikátorů regionálního rozvoje. V rámci všech výše uvedených sad je plánováno využívat a uplatňovat navrženou metodiku. Využitím navržené metodiky je možné zamezit nesrovnatelné interpretaci obdobných fenoménů. Navržená metodika poslouží také k dalším rozvojovým aktivitám CENIA. V roce 2019 tak bude navržená metodika, kromě naplňování jednotlivých výzkumných a běžných úkolů, sloužit jako výchozí bod pro mapování dalších indikátorů zabývajících se městským prostředím.

Navržená metodika je uplatnitelná i v gesci ostatních resortů a v rámci územního a regionálního plánování.

<sup>3</sup> Jednotné souřadnice S-JTSK.



## Ekonomické aspekty metodiky

Navržená metodika pracuje s volně dostupnými prostorovými daty družice Sentinel-2 a Landsat. Výchozí finanční prostředky zahrnují tedy dostatečné softwarové vybavení a personální kvalifikaci. Uplatnění navržené metodiky je časově náročná agenda, nicméně ustanovený pracovní tým pracuje na tvorbě interního softwarového řešení v rámci architektury CENIA, které by zpracování dat urychlilo.

## Závěr

Navržená metodika je uplatňována na základě výsledků rešerše literatury a následné verifikace prostorových dat řešitelským kolektivem. Takto navržená metodika byla ověřována na příkladu měst ČR s populací nad 20 tisíc obyvatel, v případě potřeby je možné využít ji pro všechny obce ČR. Navržená metodika je využitelná pro potřeby resortu životního prostředí.

Na základě navržené metodiky je možné v budoucnu hodnotit ty fenomény, které se v urbánním území vyskytují, např. se jedná o vývoj a identifikaci sídelní zeleně a její potenciál a rizika, rozvoj zástavby, identifikaci ploch náchylných k přehřívání atd.

## Seznam literatury a zdrojů

CENIA – Technický list zastavěné území (2019): Vymezení urbánního území pro potřeby identifikace a vyhodnocení indikátorů zranitelnosti. 24 s.

Český statistický úřad [online]. Obyvatelstvo [cit. 4. 2. 2019]. Dostupné z: [https://www.czso.cz/csu/czso/obyvatelstvo\\_lide](https://www.czso.cz/csu/czso/obyvatelstvo_lide)

Dobrovolný, P. (1998): Dálkový průzkum Země, Digitální zpracování obrazu. 1. vyd. Brno: Masarykova univerzita v Brně, 210 s. ISBN: 80-210-1812-7.

Harris Geospatial solutions (2018): <http://www.harrisgeospatial.com/>. Available at: <http://www.harrisgeospatial.com/> (Accessed: 4. 2. 2019).

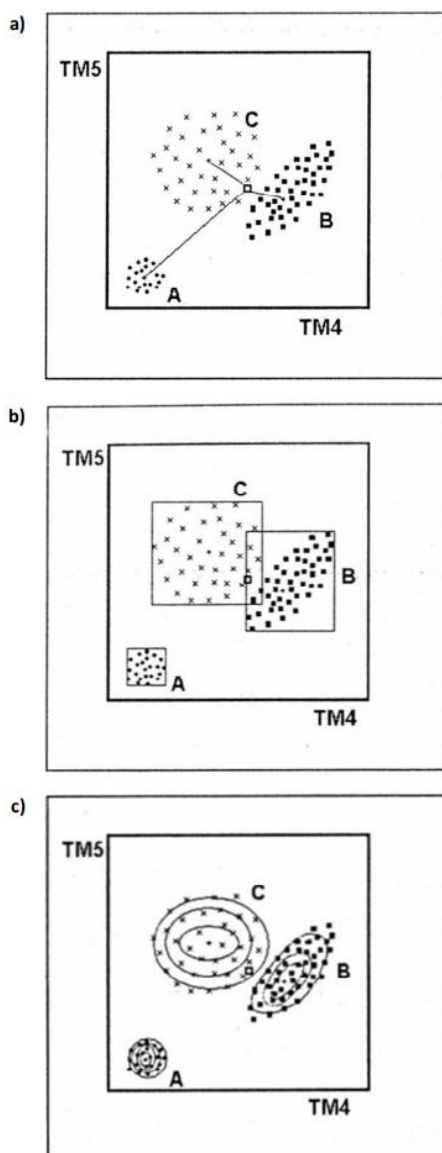
Kolář, J., Halounová, L., Pavelka, K. (1997): Dálkový průzkum Země 10. Praha: Vydavatelství ČVUT, 164 s. ISBN: 80-01-01567-X.

Lillesand, T. M., Kiefer, R. W. and Chipman, J. W. (2008): Remote sensing and image interpretation. John Wiley & Sons.

Pham, B. T., Prakash, I. and Tien Bui, D. (2018): 'Spatial prediction of landslides using a hybrid machine learning approach based on Random Subspace and Classification and Regression Trees', Geomorphology. Elsevier, 303, pp. 256–270. doi: 10.1016/J.GEOMORPH.2017.12.008

Ravindranath, N. H. and Ostwald, M. (2006): 'Remote Sensing and GIS Techniques', in Carbon Inventory Methods. Oxford University Press, pp. 181–199.

St-Louis, V. et al. (2014): 'Modelling avian biodiversity using raw, unclassified satellite imagery.', Philosophical transactions of the Royal Society of London. Series B, Biological sciences. The Royal Society, 369(1643), p. 20130197. doi: 10.1098/rstb.2013.0197.



Obrázek 2: Schéma zařazování pixelu (nevyplněný čtverec) k jednotlivým spektrálním třídám (A, B, C) pomocí tří různých klasifikátorů – a) minimální vzdálenosti, b) pravoúhelníků a c) maximální pravděpodobnosti.

Zdroj: Dobrovolný, 1998.