

Niektoré metódy vyhľadávania zmien v krajine a aktualizácia CPD.

Autori: Ing. Marián Adamják¹, Mgr. Lucia Cachovanová,
Ing. Kamil Fako, PhD..

1. Úvod a ciele

Technologický pokrok, ktorý môžeme vnímať okolo nás, je spôsobený najmä rozvojom a možnosťami informačných technológií, ktoré ľuďom priniesli obrovské možnosti rozvoja spoločnosti a kvality života. Je prirodzené že samotný rozvoj informačných technológií je podmienený ich vlastnými možnosťami. Tieto trendy zrejme platia všeobecne a teda sa dotýkajú aj informačných systémov, ktoré obsahujú priestorovú zložku.

Hlavnou črtou rozvoja v oblasti spracovania priestorových informácií je v poslednom období významný posun k spracovaniu a využívaniu vektorových typov modelov území. Tieto údaje pre tieto vektorové modely sú získavané rôznymi metódami a postupmi od presných geodetických meraní, cez metódy založené na fotogrametrickom spracovaní až po vektorizácie mapových podkladov. Jednou zo spoločných vlastností údajov získaných týmito metódami je aj miera ich aktuálnosti.

Cieľom tohoto príspevku je na jednej strane rozobrať pojem „aktuálnosť“ a na strane druhej načrtnúť niektoré možnosti, ako je možné zisťovať výskyt zmien vlastností a výskytu objektov v krajine. Zisťovanie zmien vlastností a výskytu objektov v krajine má zásadný význam pre aktualizáciu údajov v informačných systémoch s priestorovou zložkou.

2. Vek údajov a čas poznania objektov

Aktuálnosť je možné v najjednoduchšej forme vyjadriť ako rozdiel času použitia informácie a času kedy bola informácia získaná, resp. zaznamenaná. Určenie času použitia informácie je obvykle triviálnou záležitosťou avšak určenie času kedy bola informácia skutočne získaná nemusí byť najmä pri nesprávnom interpretovaní údajov jednoduchou záležitosťou.

Za čas získania informácie o objekte je potrebné vždy považovať okamžik kedy bol naposledy objekt poznávaný v rozsahu, ktorý je určený parametrami vytváraného (aktualizovaného) modelu územia. V zásade je možné na účely vymedzenia času poznávania objektu vymedziť nasledovné veľké skupiny postupov:

- a) kontaktné metódy, ktorými sú najmä geodetické metódy a metódy mapovania, zisťovania a overovania vlastností priamo v teréne,
- b) bezkontaktné metódy, ktorými sú najmä metódy diaľkového prieskumu zeme a digitálnej fotogrametrie,
- c) preberanie informácií z iných informačných systémov a zdrojov.

Časom poznania objektu pri kontaktných metódach je okamžik kedy objekt bol týmito metódami naposledy meraný a poznávaný – tým sa však nemyslí len určenie geometrie objektu, ale najmä overenie samotnej existencie objektu či zisťovanie ďalších vlastností (napr. negeometrických).

Časom poznania objektu pri bezkontaktných metódach je okamžik vyhotovenia záznamu na nosič obrazu. Pri leteckom meračskom snímkaní je to okamžik exponovania (osvetlenia) filmu. Časom poznania objektu zjavne nie je čas vyvolania filmu, či interpretácie snímok v prostredí digitálnej fotogrametrie.

¹ Topografický ústav, Banská Bystrica

Určenie času pri preberaní informácií z iných zdrojov je závislé na tom či je možné získať z informačného zdroja údaj o tom kedy bol príslušný objekt naposledy poznávaný. Jedným zo spôsobov preberania informácií z iných zdrojov je získavanie informácií z máp vektorizovaním obsahu, alebo prebráním informácií o vlastnostiach objektov. Mapové podklady však obvykle obsahujú len údaje o redakčnej uzávierke, alebo údaj o dátume schválenia obsahu. Tieto časové údaje však neuvádzajú skutočný čas poznávania objektov a často tento nie je možné ani spoľahlivo odvodiť. Ďalšími zdrojmi môžu byť rôzne databázy, katalógy a ročenky. Tieto zdroje buď vôbec neuvádzajú čas poznávania objektov či získania informácií, alebo ho uvádzajú veľmi všeobecne.

3. Zložky informácie o priestorových objektoch a ich spoľahlivosť

Informácia o priestorovom objekte (entite ktorá patrí do určitej triedy objektov) je najčastejšie vyjadrená ako súbor hodnôt opisujúcich jeho vlastnosti (atribúty). Typicky existujú atribúty týchto typov – číselný, textový, logický, časový údaj, geometria a pod. Spoľahlivosť pravdivosti každého atribútu sa v čase mení a klesá. Pre informačné systémy s údajovými skladmi veľkého rozsahu, ktoré obsahujú značné množstvá informácií o objektoch by matematicky bolo možné vyjadriť mieru spoľahlivosti jednotlivých atribútov pravdepodobnosťou, že atribút je v čase ešte pravdivý. Vyšetrením vzťahov a závislostí pravdepodobností pravdivosti jednotlivých atribútov by bolo možné kvantifikovať mieru spoľahlivosti na úrovni jednotlivých objektov.

Závislosť atribútov je možné demonštrovať na závislosti medzi atribútmi typu geometria a ostatnými atribútmi. Pri hodnotení pravdepodobnosti geometrických atribútov je možné vyjadrovať dve pravdepodobnostné charakteristiky pravdivosti a to – pravdepodobnosť existencie objektu (resp. jeho neexistencie) a pravdepodobnosť zmeny tvaru (objekt existuje ale zmenil tvar). V prípade že nastane zmena v existencii objektu implikuje nepravdivosť všetkých ostatných atribútov, v prípade zmeny tvaru objektu môže dochádzať u niektorých tried objektov k významným zmenám v pravdepodobnosti pravdivosti len u niektorých vlastností. Zároveň je potrebné hodnotiť ako sa zmena existencie či vlastností jedného objektu prejaví na vlastnostiach iného objektu.

Príkladom takéhoto posudzovania nech sú triedy objektov „budova“, „cestný úsek“ a „mostná konštrukcia“. V prípade zmeny geometrie cestného úseku zjavne došlo k prestavbe celého objektu a jeho vlastnosti (napr. šírka, materiálové zloženie povrchu) je potrebné považovať za nespoľahlivé – pravdepodobnosť pravdivosti sa blíži nule. Naproti tomu zmena geometrie budovy bude mať zrejme odlišný dopad na spoľahlivosť atribútu „účel využitia budovy“ a odlišný dopad na spoľahlivosť atribútu „výška budovy“. Zmena existencie objektu triedy „mostná konštrukcia“ zásadne ovplyvní spoľahlivosť niektorých vlastností objektov v bezprostrednom okolí. Vybudovaním mostnej konštrukcie (premostenia) nad cestným úsekom spôsobí zníženie pravdepodobnosti pravdivosti jeho atribútu „prejazdna výška“ na nulu, keďže premostenie vybudované nad cestou môže obmedzovať (a často aj obmedzí) maximálnu výšku vozidla, ktoré môže popod mostnú konštrukciu prejsť.

Z vyššie uvedených príkladov sa dá usúdiť, že by bolo možné nájsť vzťahy medzi spoľahlivosťami jednotlivých atribútov v rámci jedného objektu i medzi viacerými objektmi a zároveň že by bolo možné kvantifikovať spoľahlivosť jednotlivých atribútov a aj objektov ako celok.

4. Riziko nespoľahlivej informácie a jeho obmedzovanie

4.1. Odhad nespoľahlivosti informácií

Informačný systém, jeho údajový sklad, databáza i informácie v tejto databáze „žijú svoj životný cyklus“. Tento životný cyklus je možné v prípade informácií o priestorových objektoch opísať nasledovne:

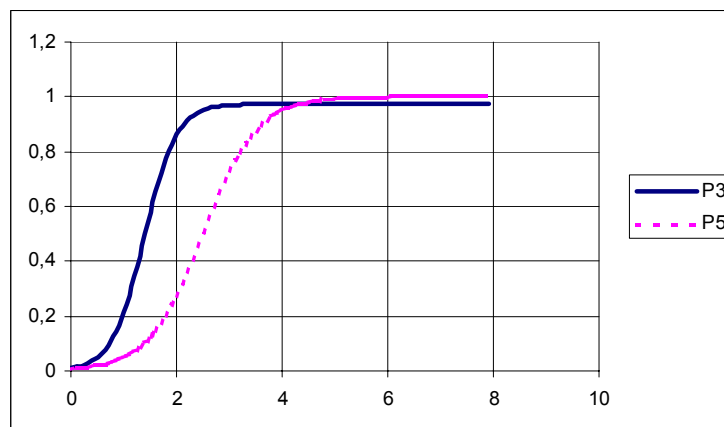
- prvotné poznávanie objektu a získanie prvotných informácií o objekte
- existencia informácie o objekte bez zmeny tejto informácie bez ohľadu na skutočný stav objektu – postupný pokles spoľahlivosti informácie o objekte
- pravidelné preverovanie vlastností objektu a ich prípadná aktualizácia
- existencia informácie o objekte bez zmeny tejto informácie bez ohľadu na skutočný stav objektu – postupný pokles spoľahlivosti informácie o objekte – po istom čase návrat k pravidelnému preverovaniu vlastností (predchádzajúci bod),
- zánik informácie o objekte spolu so zánikom objektu

Po každej etape životného cyklu v ktorej sa poznávajú vlastnosti objektu dochádza k zvýšeniu spoľahlivosti informácie o objekte – k zvýšeniu pravdepodobnosti pravdivosti informácie o objekte. Miera zvýšenia pravdepodobnosti pravdivosti je závislá od účinnosti metódy poznávania a overovania vlastností jednotlivých objektov, ktorú je možné tiež kvantifikovať pravdepodobnosťou že metóda poznávania zistí všetky zmeny existencie a vlastností objektu.

Pri použití nespoľahlivej informácie vzniká nebezpečenstvo, že výsledky a závery odvodené na základe takejto informácie budú nesprávne a ohrozia správnosť ďalších rozhodnutí a zvolených postupov. Toto nebezpečenstvo je možné kvantifikovať tak, že sa pokúsime určiť pravdepodobnosť že niektorý z údajov o priestorovom objekte je nepravdivý. Potom miera nebezpečenstva je priamo úmerná tejto pravdepodobnosti. Pre kvantifikáciu (pravdepodobnosti) sa navrhuje použiť napr. funkciu obmedzeného rastu:

$$P_t = \frac{ML}{M + (L - M)e^{-Lrt}}$$

kde P_t je pravdepodobnosť že údaj nie je spoľahlivý, M , L a r sú koeficienty charakterizujúce „starnutie“ príslušného údaju a t je čas.



obr. 1

Ak je údaj považovaný v horizonte 3 rokov za zastaralý a zároveň predpokladáme,

- že pravdepodobnosť zastarania údaju za 0,5 roku je menej ako 0,05
- a zároveň že údaj je zastaralý po 2,5 s pravdepodobnosťou viac ako 0,95 potom budú mať koeficienty približne nasledovné hodnoty:

$$M \approx 0,01 \quad L \approx 0,9737 \quad a \quad r \approx 3,3929$$

Ak je údaj považovaný v horizonte 5 rokov za zastaralý a zároveň predpokladáme,

1. že pravdepodobnosť zastarania údaje za 1 rok je menej ako 0,05
2. a zároveň že údaj je zastarany po 4 s pravdepodobnosťou viac ako 0,95 potom budú mať koeficienty približne nasledovné hodnoty:

$$M \approx 0,0073 \quad L \approx 1,00 \quad a \quad r \approx 1,9627$$

Na obr. 1 sú zobrazené grafy oboch funkcií. Funkcia P3 prislúcha prípadu kedy uvažujeme 3 ročnú spoľahlivosť a funkcia P5 päťročnú spoľahlivosť.

Autori pripúšťajú, že pre určenie pravdepodobnosti je možné použiť aj iné druhy funkcií prípadne metód. V ďalšom texte sa pridriavajú len toho faktu, že pravdepodobnosť že údaj je nespoľahlivý je určená dostatočne vhodnou metódou. Parametre týchto funkcií či metód by mali byť súčasťou technickej dokumentácie príslušného vektorového modelu územia (metaúdajov o tomto modeli).

4.2. Obmedzovanie nespoľahlivosti informácií

Model určovania pravdepodobnosti nespoľahlivosti údajov demonštrovaný v predchádzajúcej časti umožní vytvára nástroj na posudzovanie jednotlivých údajov, ktoré opisujú konkrétny objekt. Je teda možné v ľubovoľnom čase stanoviť aké spoľahlivé údaje o jednotlivých objektoch sú v modeli územia (databáze) zaznamenané.

Pri plánovaní aktualizácie bude teda možné sa oprieť o určenie pravdepodobnosti s ktorou údaje o jednotlivých objektoch zastarávajú v konkrétnom čase a vyberať len tie, ktorých miera nespoľahlivosti prekročila určenú medzu.

Obmedzovanie nespoľahlivosti údajov je možné zabezpečiť:

- a) pravidelnou aktualizáciou týchto údajov, alebo
- b) vytváraním takých údajových štruktúr, ktoré len odkazujú na zdroje údajov, tzn. na také údajové štruktúry do ktorých údaje primárne zaznamenávané (napr. kataster nehnuteľností, informačné systémy odbornej štátnej správy a pod.).

Riešenie uvedené v bode b) je už v súčasnosti známe aj keď ešte nie je rozšírené a zároveň nie je riešením pre všetky druhy údajov a modelov územia. Pre sa v ďalšom budeme zaoberať len aktualizáčnym procesom.

Aktualizáciu je možné vykonávať rôznymi metódami a postupmi. Tieto metódy a postupy budú mať rôznu úspešnosť pri vyhľadávaní (identifikácii) zmien vlastností objektov a zároveň budú mať rozdielnú schopnosť poskytovať aktuálne údaje v očakávanej kvalite. Príkladom môže byť na jednej strane porovnávanie materiálov diaľkového prieskumu zeme (vrátane leteckého meračského snímkovania) ktoré vznikli v rozdielnych časoch a na druhej strane preberanie údajov zo stavebného úradu. Metódy porovnania materiálov DPZ a LMS z rôznych časových období môžu indikovať rôzne zmeny vo vegetačnom kryte krajiny, rozsiahlu novú zástavbu atď. na veľkých častiach územia, avšak nemôžu podať spoľahlivú a úplnú informáciu o charaktere týchto zmien. Naproti tomu údaje zo stavebného úradu umožnia rýchlo nájsť nové budovy, opravované budovy či demolácie. Takto identifikované budovy sú spoľahlivo vymedzenou podmnožinou zo všetkých budov, ktoré sú zmenené. Informácie zo stavebného úradu môžu zároveň obsahovať aj ďalšie spoľahlivé údaje vhodné pre aktualizáciu informácií o týchto budovách. Oproti materiálom DPZ a LMS je však možno zo stavebného úradu získať len informácie týkajúce sa obmedzenej množiny objektov z niektorých tried.

Obdobne ako sme kvantifikovali mieru nespoľahlivosti údajov bude možné odhadnúť mieru účinnosti jednotlivých aktualizáčnych metód a postupov. Nech účinnosť U je vyjadrená ako pravdepodobnosť, že bude nespoľahlivý údaj identifikovaný a opravený, hodnota P_i nech je pravdepodobnosť nespoľahlivosti údaje, tak potom hodnota (rizikový index):

$$RI = P_i(1-U),$$

udáva mieru rizika s akým je možné údaj po aktualizácii použiť. Ak účinnosť U bude vyjadrená v intervale $(0,1)$ potom rizikový index RI nadobudne hodnoty taktiež v intervale $(0,1)$, kde 0 znamená žiadne riziko a 1 znamená maximálne riziko (údaj je nespoľahlivý).

Ak neaplikovanie žiadnej aktualizáčnej metódy má účinnosť $U = 0$ a zároveň ak teoretická metóda s účinnosťou $U = 1$ spoľahlivo zistí všetky zmeny a opraví údaje potom je možné hodnotu RI považovať za pravdepodobnosť že údaj aj po aplikovaní určitej metódy zostal nespoľahlivý a rovnicu pre určenie P_i je možné upraviť tak že koeficient M bude rovný hodnote tohoto rizikového indexu.

4.3. Závislosti a zovšeobecňovanie nespoľahlivostí údajov

Ako už bolo spomenuté existujú vzájomné väzby medzi zmenami vlastností v rámci jednotlivého objektu i medzi viacerými objektmi. Vyšetovanie týchto vzájomných závislostí bude zložité a náročné avšak potrebné.

Pre úspešné zvládnutie týchto závislostí bude potrebné dopracovať rozbor určenia pravdepodobnosti, že daný údaj je nespoľahlivý v kontexte zmien iných údajov a zároveň posúdiť vplyv zmien jednotlivých údajov podľa účelu ku ktorému sa údaje budú využívať.

Všetky tieto pravdepodobnosti a rizikové indexy by mali mať za následok a cieľ:

- a) spoľahlivo vyberať objekty u ktorých je potrebné preveriť spoľahlivosť údajov o nich (a prípadne ich opraviť)
- b) spoľahlivo zisťovať nové objekty, alebo zaniknuté objekty
- c) spoľahlivo určovať dôsledky zistených zmien v údajoch vybraných objektov, vznik nových objektov a zánik objektov na ostatné objekty.

5. Možnosti a metódy aktualizácie vektorových modelov územia typu Centrálnej priestorovej databázy

Aktualizácia rozsiahlych vektorových modelov územia, akými je aj Centrálna priestorová databáza Vojenského informačného systému o území, vyžaduje vzhľadom k množstvu registrovaných objektov, aby tento proces bol efektívne organizovaný a riadený.

V procese prípravy a plánovania aktualizácie bude potrebné spoľahlivo identifikovať zmeny, odhadnúť ďalšie rizikové údajové množiny a na základe týchto údajov optimalizovať celý aktualizáčny proces. V ďalšej časti poukážeme na niektoré metódy, ktoré by mohli byť v proces aktualizácie CPD použité.

5.1. Detekcia zmien porovnaním snímok

Prakticky v každom geografickom informačnom systéme, hlavne u tých „rastovo orientovaných“ existuje nástroj na automatickú detekciu zmien (change detection) – multitemporálna analýza. Tieto moduly pracujú prevažne na princípe porovnávania spektrálnych charakteristík obrazu pixel za pixelom. V prípade nájdenia väčšej odchýlky ako stanovený limit sa takéto pixle označia za miesta pravdepodobnej zmeny. Ďalším krokom je následne zisťovanie, ku akým zmenám na danom mieste došlo. Tu už sa vyžaduje vstup operátora, ktorý vizuálnym porovnaním obrazov eliminuje nepodstatné oblasti zmien. Dôležité predpoklady na správne fungovanie procesu:

- presne nageoreferencované obrazy. V prípade, že obrazy nie sú úplne identicky umiestnené v priestore, procedúra nájde zmeny v miestach, kde sa v skutočnosti nevyskytujú. Takáto chyba sa objavuje zriedka a dá sa zaradiť medzi hrubé chyby, ktoré spôsobujú, že výsledok analýzy je nesprávny.

- vhodný časový odstup obrazov. Hlavne z dôvodu rýchlosti a veľkosti zmien jednotlivých prvkov je správna voľba časového odstupu kritická. O optimálnom časovom odstupe sme písali v predchádzajúcich častiach príspevku.
- vhodná voľba typu materiálu a senzora. Nie každý prvok je dobre zreteľný v každom spektrálnom pásme. Pre detekciu zmien platí podobný princíp ako pre detekovanie prítomnosti/nepřítomnosti samotného prvku.

Metódy digitálnej identifikácie zmien:

- Metóda obrazových diferencií vyžaduje priestorovo identicky registrované obrazy v čase t_1 a t_2 . Odčítaním hodnôt DN identických obrazových prvkov vzniknú diferencie tvoriace nový obraz. Ten reprezentuje zmeny v sledovanom období. Prvky, ktoré neboli zmenené, budú mať diferencie blížiac sa nule, zmeny budú mať diferencie väčšie (kladné alebo záporne).
- Metóda obrazovej regresie je založená na predpoklade existencie lineárnej závislosti obrazových údajov v čase t_2 a t_1 . Táto metóda, bližšie popísaná v Žihlavník – Scheer², má výhodu v tom, že redukuje nepriaznivý vplyv rozdielnych atmosférických podmienok a uhol dopadajúcich slnečných lúčov.
- Metóda obrazových pomerov vypočítava podiely spektrálnych hodnôt jednotlivých prvkov. Ak je intenzita odraznosti približne rovnaká, v oboch časoch t_1 a t_2 , potom podiel sa blíži 1 a znamená to, že nedošlo k zmenám. Ak bude výsledok signifikantne väčší alebo menší ako 1 k zmenám došlo.

Vhodnosť ktorejkoľvek z uvedených metód pre aktualizáciu CPD je sporná. V prípadoch, kedy je potrebné identifikovať zmeny aj objektoch menších rozmerov je malá pravdepodobnosť, že obrazy budú mať zhodné základné charakteristiky a bude možné nájsť požadované zmeny. Doba získania materiálu, vzhľadom na ročné obdobia a vegetačné fázy bude zrejme limitujúcim faktorom.

Metóda manuálnej identifikácie zmien.

Táto metóda sa používa zriedka a iba na malé územia. Jej princíp spočíva v porovnávaní rastrového podkladu s už existujúcou databázou objektov operátorom. Dá sa konštatovať, že metóda je z hľadiska identifikácie zmien najlepšia. Navyše, operátor dokáže pri vhodnej rozlišovacej schopnosti detekovať aj menšie zmeny. Využitie metódy je však ekonomicky výhodnejšie iba v kombinácii s niektorou z predchádzajúcich digitálnych automatických metód detekcií.

5.2. Opakovaný zber údajov

V najbližšej budúcnosti bude potrebné rozpracovať metodiku obnovy CPD pre každý jednotlivý prvok. Z ekonomického hľadiska bude najdôležitejšie synchronizovať interval obnovy tak, aby sa detekovalo jednoduchou a rýchlou metódou čo najviac zmien. Ak dokážeme pre každý druh objektov stanoviť pravdepodobnosť nespoľahlivosti, nebude problém správne zvoliť interval jeho obnovy. Zároveň však bude potrebné určiť aj skupiny objektov, ktoré sa budú obnovovať spoločne. Tým sa dosiahne vhodný ekonomický prínos.

Na objektoch, kde sa zvolenou metódou detekovali zmeny bude potrebné pristúpiť v ďalšom kroku ku ich zberu, resp. oprave. Metóda získania geometrie objektu by mala byť podobná ako metóda, ktorou sa získavali objekty susedné. Tým sa zabezpečí kontinuita celej

² Žihlavník, Š. – Scheer, E.: Diaľkový prieskum Zeme v lesníctve, TU vo Zvolene, 1998.

databázy. Po získaní geometrie objektov bude zrejmé nutné pristúpiť aj k získaniu vlastností objektu. Opätovné miestne šetrenie je teda nevyhnuté. Jeho rozsah však bude značne menší a preto aj šetrenie v teréne bude lacnejšie.

5.3. Preberanie údajov

Perspektívnou metódou sa ukazuje metóda preberania údajov z databáz iných informačných systémov. Ak uvážime, že jednotlivé rezorty štátnej správy vedú svoje vlastné databázy objektov a vlastností, ktoré ich zaujímajú, môžu tieto predstavovať vhodný zdroj identifikovaných zmien oproti referenčnej (pôvodnej) databáze. V ideálnom prípade, ak objekty v rezortnej databáze sú metaúdajmi správne popísané nie je problémom selektovať iba objekty nové (dátum vzniku väčší ako špecifický). Takýto výber potom predstavuje základ preverovanej skupiny objektov. V prípade vhodnej metódy získania geometrie je možné prebrať objekt aj s jeho geometriou a následne ušetriť finančné prostriedky.

Výhodou tejto metódy je vysoká spoľahlivosť. Keďže rezortné databázy sú úzko špecializované na určité skupiny objektov je veľká pravdepodobnosť, že objekty sú udržiavané v aktuálnom stave. Naopak nevýhodou to, že sa jedná iba o určité triedy objektov a samozrejme, iba o tie vlastnosti objektov o ktoré má rezort záujem. To však nemusí korešpondovať s požiadavkami databázy, ktorú chceme obnovovať a teda miestnemu prešetrovaniu sa zrejme nevyhneme. Navyše, podmienkou je kompatibilita databáz a ochota rezortu takéto údaje poskytnúť.

6. Záver

V článku sa autori pokúsili naznačiť niektoré prístupy a možnosti aktualizácie údajov v informačných systém, ktoré obsahujú aj priestorovú zložku. Ukazuje sa, že kvalitný a spoľahlivý proces aktualizácie údajov je z pohľadu organizácie a riadenia náročnejší než proces prvotného zberu údajov o priestorových objektoch.

Je viac než isté, že je potrebné používať spoľahlivé metódy, ktoré budú indikovať zmeny v krajine a vlastností objektov v nej. Bez spoľahlivej indikácie zmien bude proces ekonomicky náročnejší a zároveň menej spoľahlivý.

Žiaľ v súčasnom čase nie je možné presne stanoviť, ktorá metóda je pre aktualizáciu údajov najlepšia. V prípadoch rozsiahlych databáz, akou bezpochyby CPD je, bude aktualizácia objektov zrejme kombinácia viacerých (ak nie všetkých) spomenutých metód.

