

Příloha 2

Statistické metody pro vyhodnocení metody nepřímého měření kontaminace ropnými látkami pomocí přirozeného stopovače Rn

Podrobný postup zpracování dat a použití programu RADOIL

Použité zkratky

OAR Objemová aktivita radonu

NAPL Non-Aqueous Phase Liquids (látky nerozpustné ve vodě, např. ropné látky)

1. Účel, rozsah a vazba na metodiku

Tato příloha podrobně vysvětluje statistické vyhodnocení OAR a prostorové interpretaci výsledků. Popisuje přípravu dat, kontrolu jejich kvality, matematické principy použitých metod, interpretaci výstupů a praktické použití programu RADOIL včetně exportu výsledků do PDF a Excelu.

Příloha je koncipována jako pracovní návod pro zpracovatele terénních měření. Předpokládá, že terénní měření bylo provedeno podle metodiky, včetně jednotné odběrové hloubky, zaměření bodů a průběžné kontroly kvality. Statistické metody zde popsané pracují s výsledky měření OAR v jednotlivých bodech a s jejich prostorovou polohou.

1.1 Hlavní výstup vyhodnocení

Hlavním výstupem vyhodnocení je vymezení bodů nebo souvislých oblastí s nízkou OAR, které mohou indikovat výskyt NAPL. Výsledek je screeningovou indikací. Sám o sobě neprokazuje přítomnost ropné fáze, ale slouží k vytipování prioritních míst pro přímé ověřovací metody, například vrtný průzkum a měření volné fáze na hladině podzemní vody.

1.2 Vztah k programu RADOIL

Program RADOIL zajišťuje jednotné zpracování tabulky měřených bodů OAR, volitelně také polygonu známého nebo ověřeného výskytu NAPL. Program vypočte vybrané statistické metody, vytvoří mapové vrstvy detekovaných bodů, vyhodnotí shodu více metod a při zadání polygonu provede validační porovnání vůči skutečné oblasti NAPL. Výsledky lze exportovat do PDF a Excelu.

1.3 Rozsah platnosti přílohy

- Příloha se vztahuje k plošným bodovým měřením OAR v půdním vzduchu, typicky v pravidelné síti 10 x 10 m nebo v podrobnější síti.
- Metody jsou určeny pro lokality, kde je možné statisticky vyhodnocovat relativní pokles OAR v rámci jednoho dostatečně homogenního měřeného souboru.
- Při významné geologické heterogenitě nebo při měření realizovaném za výrazně odlišných podmínek je nutné data rozdělit na dílčí homogenní soubory a vyhodnotit je samostatně.
- Percentilové prahy nejsou legislativním ani hygienickým limitem. Jsou relativním prahem odvozeným z naměřeného souboru konkrétní lokality.

2. Datový model, vstupní soubory a kontrola kvality

2.1 Základní datové vrstvy

V rámci programového zpracování se pracuje se třemi hlavními skupinami dat. První je tabulka měřených bodů, druhou je volitelný polygon obálky kontaminace a třetí jsou anomální body vybrané jednotlivými statistickými metodami během výpočtu. Všechny prostorové vrstvy musí být ve stejném souřadnicovém systému nebo musí být před výpočtem korektně transformovány.

Tabulka 1 - Základní datové vrstvy v programu RADOIL.

Datová skupina	Obsah	Použití ve vyhodnocení
Naměřená data OAR	ID bodu, souřadnice X, Y, hodnota OAR v kBq.m-3	Výpočet percentilů, nejbližších sousedů, mapové zobrazení a klasifikace anomálií
Obálka reálně ověřeného výskytu NAPL	Polygon vymezující měřením nebo jiným průzkumem ověřenou oblast výskytu NAPL	Volitelná validační vrstva pro určení správných, falešných a neidentifikovaných bodů
Vybrané anomální body	Podmnožina měřených bodů označená statistickou metodou jako podezřelá pro výskyt NAPL	Mapová interpretace, výpočet shody metod a vyhodnocení vůči polygonu NAPL

2.2 Doporučená struktura vstupní tabulky měřených bodů

Vstupní tabulka měřených bodů musí být jednoznačná, bez sloučených buněk a bez více tabulek na jednom listu. Zpracovatel musí uchovat vazbu na primární terénní protokol, aby bylo možné zpětně dohledat původ každé hodnoty.

Tabulka 2 - Doporučená struktura vstupních dat a návazného terénního protokolu.

Položka	Popis a požadavky
ID	Unikátní identifikátor měřicího bodu. U opakovaných kampaní zachovat stejná ID nebo uvést jednoznačnou vazbu na původní bod.
X	Souřadnice bodu. V ČR je pro terénní zaměření doporučen JTSK (metrický projektový souřadnicový systém je zásadní pro vyhodnocení metody)
Y	Souřadnice bodu. V ČR je pro terénní zaměření doporučen JTSK (metrický projektový souřadnicový systém je zásadní pro vyhodnocení metody)
OAR	Naměřená objemová aktivita radonu v půdním vzduchu. Hodnoty musí být numerické, jednotně v kBq.m-3.

2.3 Doporučená struktura polygonu hranice NAPL

Hranice polygonu NAPL je volitelný vstup. Používá se tehdy, když je na lokalitě známý skutečný nebo referenčně ověřený rozsah volné fáze ropných látek, například z vrtů změřený fázoměrem nebo jinou relevantní metodou. Program jej využívá pro validační porovnání, nikoli pro samotný výpočet percentilů měřených bodů.

Tabulka 3 - Vstupní tabulka hranice NAPL pro validační porovnání.

Položka	Požadavek
X	Souřadnice bodu obálky polygonu.
Y	Souřadnice bodu obálky polygonu.
Pozn.: Body musí tvořit logickou obálku bez křížení hran. Souřadnicový systém musí být totožný s měřeními body OAR.	

2.4 Kontrola kvality dat před importem do RADOIL

Kontrola dat musí proběhnout nejpozději bezprostředně po ukončení terénního měření. Doporučuje se provést ji ještě před importem do programu, aby statistika nepracovala s chybami, které lze odstranit opakovaným měřením nebo kontrolou terénního protokolu.

Tabulka 4 - Minimální QA/QC kontrola před statistickým vyhodnocením.

Kontrolní krok	Co zkontrolovat	Doporučený postup při zjištění problému
Úplnost	Chybějící ID, X, Y, OAR	Doplnit z terénních záznamů nebo bod opakovat; neodhadovat hodnoty OAR bez měření.
Odběrová hloubka	Všechny body jsou změřeny ve stejné hloubce	Opakovat měření

Kontrolní krok	Co zkontrolovat	Doporučený postup při zjištění problému
Duplicity	Duplicitní ID nebo shodné souřadnice bez vysvětlení	Ověřit, zda nejde o opakované měření nebo chybu zápisu; v datech ponechat pouze správně označený záznam.
Extrémní hodnoty	OAR ≤ 1 kBq.m-3 nebo hodnota o řády odlišná od ostatních	Provést kontrolu terénního postupu, těsnosti odběru a přístroje; podle možností opakovat měření v novém bodě.
Jednotky	Smísení kBq.m-3, Bq.l-1 nebo jiných jednotek	Převést na jednotné kBq.m-3 a převod zdokumentovat.
Souřadnice	Prohozené sloupce X/Y, smíšený JTSK a WGS84, chybné znaménko	Ověřit mapovým náhledem a porovnáním s podkladovou mapou; transformaci provést jednotně.
Vícedenní měření	Denní rozdíl kontrolního bodu > 30 %	Rozdělit lokalitu na samostatné jednodenní datové soubory a vyhodnotit je odděleně.
Geologická heterogenita	Výrazné litologické změny nebo nehomogenní navážky	Rozdělit území na homogennější celky nebo zvýšit nejistotu interpretace a ověřit výsledky přímou metodou.

3. Matematický základ vyhodnocení OAR

3.1 Datová množina

Měřený soubor obsahuje n bodů. Každý bod i je popsán identifikátorem ID_i , souřadnicemi X_i , Y_i a hodnotou OAR_i . V dalším textu se předpokládá, že nižší hodnota OAR_i znamená vyšší míru podezření na přítomnost NAPL, pokud nejsou přítomny jiné příčiny poklesu, například geologická změna, nízká plynopropustnost nebo chyba odběru.

Datová množina = $\{(ID_i, X_i, Y_i, OAR_i), i = 1, \dots, n\}$
 OAR_i ... objemová aktivita radonu v bodě i [kBq.m-3]

3.2 Percentilový práh

Percentilový práh P_q je hodnota, pod níž leží přibližně q % hodnot OAR v daném souboru. Protože je práh vždy počítán z konkrétní lokality nebo z homogenního dílčího souboru, jde o relativní lokální prahování. Program RADOIL používá pro hlavní bodové metody zejména 20. a 50. percentil.

$P_q = q$ -tý percentil hodnot $\{OAR_1, OAR_2, \dots, OAR_n\}$
 Bod i je bodově anomální, pokud $OAR_i \leq P_q$

Hodnoty rovné prahu se zařazují mezi anomální, protože v metodice i v popisu modelu je uvedeno pravidlo „nižší nebo rovno prahu“. Při větším počtu shodných hodnot okolo prahu proto může být počet označených bodů mírně odlišný od přesného procenta q %.

3.3 Eukleidovská vzdálenost a nejbližší sousedé

Prostorově kontextové metody pracují s nejbližšími sousedy měřeného bodu. Vzdálenost mezi bodem i a bodem j je v projektovaném souřadnicovém systému počítána eukleidovskými:

Eukleidovská vzdálenost $d_{ij} = \sqrt{(X_i - X_j)^2 + (Y_i - Y_j)^2}$
 $N_k(i)$ = množina k bodů s nejmenší vzdáleností od bodu i , obvykle včetně samotného bodu i

Zahrnutí samotného bodu odpovídá popisu testované metody „čtyři nejbližší sousedé včetně samotného bodu“. V pravidelné síti 10×10 m tak skupina reprezentuje lokální okolí bodu. U nepravidelné sítě je nutné zkontrolovat, zda nejbližší sousedé skutečně tvoří lokálně smysluplnou skupinu a nejsou ovlivněni okrajem plochy nebo velkou mezerou v síti.

3.4 Průměr nejbližších sousedů

U metod založených na průměru nejbližších sousedů se nejprve pro každý bod vybere skupina $N_k(i)$ a následně se vypočte průměrná OAR této skupiny. Skupina je označena jako podezřelá, pokud je její průměr nižší nebo roven stanovenému percentilovému prahu.

$\text{meanOAR}_k(i) = (1/k) * \sum_{j \in N_k(i)} OAR_j$
 Skupina $N_k(i)$ je anomální, pokud $\text{meanOAR}_k(i) \leq P_q$

3.5 Pravidlo „3 body menší než percentil“

Toto pravidlo neporovnává průměr celé skupiny, ale počet bodů ve skupině čtyř nejbližších bodů, které leží pod prahovou hodnotou. Skupina je označena jako anomální tehdy, když alespoň tři ze čtyř bodů

mají OAR nižší nebo rovnou zvolenému percentilu. Metoda tím omezuje vliv jednoho náhodně nízkého bodu a současně zachovává citlivost k souvislému lokálnímu poklesu.

```
s_i = sum_{j in N_4(i)} I(OAR_j <= P_q)
Skupina N_4(i) je anomální, pokud s_i >= 3
```

3.6 Shoda více metod

Shoda více metod je interpretační vrstva, která sčítá, kolika vybranými metodami byl bod označen jako anomální. Vyšší shoda zpravidla znamená vyšší robustnost indikace, nikoli automatické potvrzení kontaminace. Program RADOIL podle ukázkového exportu vytváří samostatné mapové výstupy pro shodu alespoň dvou metod a pro shodu tří metod.

```
M_i = počet metod, které označily bod i jako anomální
Shoda alespoň 2 metod: M_i >= 2
Shoda 3 metod: M_i = 3
```

4. Statistické metody testované v projektu RADOIL

V rámci tvorby modelu pro vyhodnocení terénních měření byly testovány percentilové, prostorové kontextové a kombinované metody. Jejich společným znakem je snaha nalézt lokální nebo plošné poklesy OAR, které mohou odpovídat přítomnosti volné fáze ropných uhlovodíků. Následující přehled shrnuje princip, přednosti a rizika každé metody.

Tabulka 5 - Statistické přístupy testované při tvorbě modelu RADOIL.

Metoda	Princip	Přednost	Hlavní omezení / riziko
Percentily z bodů	Vypočte se percentil z hodnot OAR celého souboru. Body s OAR <= práh jsou anomální.	Jednoduché, rychlé, vhodné pro orientační screening a jako základní srovnávací metoda.	Nezohledňuje prostorovou souvislost; izolovaná chyba měření může být označena jako anomálie.
Percentily z průměrů čtverců - 4 body	Pro každý bod se vyberou čtyři nejbližší body včetně samotného bodu; průměr OAR se porovná s prahem.	Robustnější vůči jednotlivým extrémům; lépe hledá souvislé oblasti nízkých hodnot.	Může vyhladit menší a ostré lokální anomálie.
Percentily z průměrů čtverců - 3 body	Analogicky jako 4 body, ale průměr je počítán ze tří nejbližších bodů.	Citlivější k menším lokálním anomáliím.	Vyšší náchylnost k šumu než u čtyř bodů.
Percentily z průměrů čtverců - 2 body	Průměr ze dvou nejbližších bodů se porovná s prahem.	Velmi citlivé na ostrý lokální pokles.	Vyšší riziko falešné detekce způsobené odlehlou hodnotou; vhodné spíše doplňkově.
Čtverce - 3 body menší než percentil	Ve skupině čtyř nejbližších bodů musí alespoň tři body ležet pod prahem.	Kompromis mezi citlivostí a robustností; omezuje vliv jednoho extrémně nízkého bodu.	Výsledek závisí na hustotě a pravidelnosti sítě a na volbě percentilu.

4.1 Percentily z bodů

Bodový percentilový přístup je základní metodou pro rychlou klasifikaci hodnot. Pokud je například zvolen 20. percentil, program identifikuje přibližně pětinu nejnižších hodnot OAR v daném souboru. Pokud je zvolen 50. percentil, program pracuje s mediánem a označuje nižší polovinu souboru. P20 je konzervativnější, P50 citlivější a agresivnější.

Výhodou je jednoduchost a srozumitelnost. Nevýhodou je absence prostorového kontextu. Bod může být nízký kvůli skutečné NAPL, ale také kvůli lokální změně plynopropustnosti nebo geologické heterogenitě.

4.2 Percentily z průměrů čtverců

Metody z průměrů nejbližších bodů rozšiřují bodové prahování o prostorový kontext. Pro každý bod se vytvoří lokální skupina sousedů a teprve průměr této skupiny se porovnává s prahovou hodnotou. Pokud je průměr nízký, nejde pravděpodobně o osamocený extrém, ale o prostorově souvislejší pokles OAR.

- Varianta 4 body: vhodná pro robustnější vymezení souvislých poklesů OAR; méně citlivá k osamoceným chybám.
- Varianta 3 body: zachytí menší a lokálnější anomálie, ale je citlivější na šum.
- Varianta 2 body: může upozornit na velmi malé lokální poklesy, ale má vyšší riziko falešných detekcí a je vhodná pouze jako doplňková analýza.

4.3 Čtverce - 3 body menší než percentil

Kombinovaná metoda „3 body menší než percentil“ pracuje se skupinou čtyř nejbližších bodů. Anomálie je potvrzena pouze tehdy, když nízká hodnota není izolovaná, ale vyskytuje se alespoň ve třech bodech lokální skupiny. Metoda proto dobře odpovídá cíli vymezit souvislé oblasti poklesu OAR a současně snížit citlivost na jeden mimořádně nízký bod.

Interpretační význam prostorově kontextových metod

Přítomnost NAPL se očekává spíše jako plošně nebo liniově souvislejší jev než jako jediný izolovaný bod. Prostorové metody proto obvykle poskytují vyšší interpretační hodnotu než samotné bodové prahování, zejména pokud je měřicí síť pravidelná a prostředí geologicky homogenní.

5. Doporučené metody implementované v programu RADOIL

Na základě porovnání testovaných přístupů byly pro praktické použití v programu RADOIL vybrány tři hlavní metody. Jejich kombinace umožňuje vyhodnocovat jak konzervativní detekci s nízkým počtem falešných pozitiv, tak citlivější záchyt a prostorově robustní lokální shodu.

Tabulka 6 - Hlavní metody používané v programu RADOIL.

Zkratka ve výstupech	Úplný název	Typ interpretace	Doporučené použití
P20	Percentily z bodů - 20. percentil	Konzervativní bodová detekce	Výchozí přísnější práh; vhodné pro omezení falešných detekcí mimo skutečnou oblast NAPL.
P50	Percentily z bodů - 50. percentil	Citlivější bodová detekce	Agresivnější screening; vhodné pro zachycení širší zóny podezření, pokud je akceptovatelná vyšší míra falešných detekcí.
Čtverce	3 body menší než percentil - 50. percentil	Prostorově kontextová detekce	Doporučený kompromis mezi citlivostí a robustností; vhodné pro hledání souvislých oblastí poklesu OAR.

5.1 Doporučená interpretační logika

Tabulka 7 - Doporučená interpretační logika kombinovaných výstupů.

Situace ve výstupu	Interpretační význam	Doporučený navazující krok
Bod/oblast je označena metodou P20	Silnější indikace nízké OAR, protože bod leží mezi nejnižšími hodnotami souboru.	Provéřit prostorovou vazbu na okolní body a zvážit přednostní přímé ověření.
Bod/oblast je označena metodou P50, nikoli P20	Slabší nebo širší screeningová indikace; může jít o okrajovou oblast poklesu OAR.	Nepřeceňovat izolované body; využít společně s mapou shody a s geologickými informacemi.
Bod/oblast je označena metodou Čtverce	Pokles OAR má lokální prostorovou oporu ve skupině sousedních bodů.	Považovat za důležitý kandidát pro vymezení anomální zóny.
Shoda alespoň 2 metod	Vyšší robustnost indikace, zejména pokud body tvoří souvislý celek.	Zařadit mezi prioritní zóny pro přímé ověření.
Shoda 3 metod	Nejvyšší statistická shoda v rámci programu RADOIL.	Považovat za jádro nejpravděpodobnější anomální zóny; ověřit přímou metodou.
Izolovaný bod bez shody v okolí	Možná chyba měření, lokální heterogenita nebo slabá okrajová indikace.	Zkontrolovat terénní záznam, plynopropustnost a případně bod opakovat.

5.2 Doporučené pořadí posuzování výstupů

1. Nejprve ověřit, zda datový soubor splňuje podmínky homogenity, rozsahu a kvality měření.
2. Zobrazit mapu všech měřených bodů a zkontrolovat, zda síť odpovídá terénní dokumentaci.
3. Vyhodnotit samostatně P20, P50 a Čtverce, aby bylo zřejmé, jak se mění rozsah indikovaných bodů.
4. Zobrazit shodu alespoň 2 metod a shodu 3 metod a posoudit, zda body tvoří prostorově souvislé zóny.
5. Pokud je k dispozici polygon skutečné NAPL, vyhodnotit validační metriky a posoudit, která metoda odpovídá nejlépe referenčnímu stavu.
6. V závěru uvést jak statistický výstup, tak geologické, hydrogeologické a technické faktory, které mohou interpretaci ovlivnit.

6. Validace výsledků vůči polygonu skutečného výskytu NAPL

Pokud je známý polygon skutečného nebo referenčně ověřeného výskytu NAPL, lze každou statistickou metodu porovnat s tímto polygonem. Cílem není pouze spočítat počet označených bodů, ale odlišit správné detekce uvnitř polygonu od falešných detekcí mimo polygon a od bodů uvnitř polygonu, které metoda nezachytila.

6.1 Definice stavů bodu

Tabulka 8 - Definice validačních stavů používaných při porovnání s polygonem NAPL.

Stav	Definice	Význam
Správně uvnitř	Bod je uvnitř polygonu NAPL a metoda jej označila jako anomální.	Správná detekce / true positive.
Falešně mimo	Bod je mimo polygon NAPL, ale metoda jej označila jako anomální.	Falešná detekce mimo ověřenou oblast / false positive.
Neidentifikované uvnitř	Bod je uvnitř polygonu NAPL, ale metoda jej neoznačila jako anomální.	Falešně negativní detekce / false negative.
Správně mimo	Bod je mimo polygon NAPL a metoda jej neoznačila jako anomální.	Správné odmítnutí / true negative.
Neidentifikované	Body, které metoda neoznačila jako anomální bez ohledu na polohu.	Doplňková informace o celkovém počtu neoznačených bodů.

6.2 Doporučené metriky

Programový výstup podle ukázkového exportu uvádí počet správně určených bodů uvnitř polygonu, procento správně uvnitř, počet falešných bodů mimo polygon, procento falešně mimo, počet neidentifikovaných bodů a počet neidentifikovaných bodů uvnitř polygonu. Doporučené vzorce jsou následující:

```

N_in  = počet měřených bodů uvnitř polygonu NAPL
N_out = počet měřených bodů mimo polygon NAPL
TP    = správně uvnitř
FP    = falešně mimo
FN    = neidentifikované uvnitř
TN    = správně mimo

% správně uvnitř = TP / N_in * 100
% falešně mimo  = FP / N_out * 100
Neidentifikované = n - (TP + FP)
Neidentifikované uvnitř = FN

```

Pro interpretaci je důležité sledovat obě procenta současně. Metoda s vysokým počtem správně identifikovaných bodů uvnitř polygonu může být prakticky méně vhodná, pokud současně vytváří příliš vysoký počet falešných detekcí mimo polygon. Naopak velmi konzervativní metoda může mít málo falešných detekcí, ale může nezachytit část skutečné oblasti NAPL.

6.3 Hodnocení kompromisu mezi citlivostí a robustností

Při výběru optimální interpretace se doporučuje hodnotit zejména:

- kolik bodů uvnitř skutečné oblasti NAPL bylo správně označeno,
- kolik bodů mimo skutečnou oblast NAPL bylo chybně označeno,
- kolik bodů uvnitř skutečné oblasti NAPL metoda nezachytila,
- zda označené body tvoří geologicky a hydrogeologicky smysluplnou prostorovou zónu,
- zda je výsledek stabilní při porovnání P20, P50, Čtverce a map shody.

7. Použití programu RADOIL - doporučený pracovní postup

Tato kapitola popisuje doporučený pracovní postup z hlediska zpracovatele dat. Nejde o popis konkrétního grafického rozhraní po jednotlivých tlačítkách, ale o metodický postup, který má být při použití programu RADOIL dodržen bez ohledu na technickou verzi nástroje.

7.1 Příprava pracovního balíčku vstupních dat

- Vytvořit pracovní složku lokality s jednoznačným názvem, datem kampaně a verzí zpracování.
- Uložit původní terénní protokol beze změn jako archivní zdroj.
- Připravit očištěnou vstupní tabulku měřených bodů ve formátu Excel s povinnými sloupci ID, X, Y a OAR.
- Pokud je k dispozici referenční hranice NAPL, připravit samostatnou tabulku polygonu se sloupci X a Y.

7.2 Import dat

Po importu musí být v programu zkontrolováno, že počet načtených bodů odpovídá tabulce, hodnoty OAR jsou numerické a mapa bodů odpovídá očekávané poloze lokality. Při importu polygonu je nutné ověřit, že polygon překrývá body ve správné poloze. Chybná transformace souřadnic patří mezi nejzávažnější chyby, protože může zcela znehodnotit validační metriky a mapovou interpretaci.

7.3 Spuštění výpočtu a kontrola výsledků

Tabulka 9 - Doporučený pracovní postup v programu RADOIL.

Krok	Akce	Kontrola
1	Načíst tabulku měřených bodů OAR.	Počet bodů, sloupce ID/X/Y/value, jednotky.
2	Volitelně načíst polygon hranice NAPL.	Poloha polygonu, uzavření obálky, shodný souřadnicový systém.
3	Spustit výpočet metod P20, P50 a Čtverce.	Přítomnost výsledků pro všechny metody, výpočet percentilů.
4	Zobrazit jednotlivé mapy metod.	Ověřit prostorovou logiku a souvislost anomálií.
5	Zobrazit mapy shody alespoň 2 metod a shody 3 metod.	Identifikovat prioritní zóny pro ověření.
6	Vyhodnotit tabulku validačních metrik.	Porovnat TP, FP, FN a relativní procenta.
7	Exportovat výsledky do PDF a Excelu.	Archivovat export spolu se vstupními soubory a verzí zpracování.

7.4 Programové prostředí testovaného modelu

Při vývoji a testování statistického modelu byly použity standardní balíčky jazyka R pro tabulkovou práci, prostorovou analýzu, mapové výstupy a export výsledků. Přehled je důležitý zejména pro dokumentaci výpočetní logiky a pro případnou odbornou reprodukovatelnost vyhodnocení.

Tabulka 10 - Programové balíčky použité při testování modelu RADOIL.

Balíček / skupina balíčků	Role ve zpracování modelu
dplyr	Zpracování, čištění, filtrování a sumarizace tabulkových dat.
sf, sp	Práce s prostorovými daty, body, polygony a prostorovým překryvem vůči hranici NAPL.
leaflet, htmlwidgets	Tvorba interaktivních mapových výstupů a jejich export do HTML.
writexl	Export výsledkových tabulek do formátu Excel.
readxl	Import naměřených dat a tabulkových vstupů z Excelu.
ggplot2	Alternativní grafické výstupy a kontrolní vizualizace.

8. Výstupy programu RADOIL do PDF a Excelu

Program RADOIL umožňuje export výsledků do PDF i Excelu. Oba výstupy mají odlišnou roli. PDF je vhodné jako statická výsledková zpráva pro sdílení a archivaci. Excel je vhodný pro další tabulkové zpracování, kontrolu hodnot, přenos do dalších systémů a uchování mapových listů v jednom souboru.

8.1 Export do PDF

PDF export obsahuje tabulku výsledků analýzy, mapu shody alespoň dvou metod, mapu shody tří metod, samostatné mapy metod Čtverce, percentil 20 a percentil 50, výpis datových bodů a výpis hranice. Struktura PDF je vhodná pro rychlé vložení do průzkumné zprávy nebo pro přílohu k internímu vyhodnocení.

Tabulka 12 - Doporučený obsah PDF exportu RADOIL podle ukázkového souboru.

Část PDF exportu	Obsah	Účel
Výsledky analýzy	Tabulka metrik pro metody Čtverce, P20 a P50.	Rychlé srovnání úspěšnosti a falešných detekcí.
Mapa - shoda alespoň 2 metod	Všechny body, polygon hranice a body splňující shodu nejméně dvou metod.	Prioritní zóny robustnější indikace.
Mapa - shoda 3 metod	Body označené všemi třemi metodami.	Jádro nejrobustnější indikace.
Mapa - čtverce	Výsledek prostorově kontextové metody.	Kontrola souvislých lokálních poklesů OAR.
Mapa - percentil 20	Konzervativní bodový výstup.	Kontrola nejnižších hodnot OAR.
Mapa - percentil 50	Citlivější bodový výstup.	Širší screening potenciální anomálie.
Data	Výpis měřených bodů ID, X, Y, value.	Auditní kontrola vstupu.
Hranice	Výpis souřadnic polygonu hranice.	Auditní kontrola referenčního polygonu.

8.2 Export do Excelu

Excel export obsahuje listy uvedené v následující tabulce. List Data uchovává vstupní body, list Výsledky shrnuje validační metriky a mapové listy obsahují statické mapové výstupy z programu. Příložený ukázkový soubor používá názvy listů: Data, Výsledky, Shoda_2_metod, Shoda_3_metod, Mapa_ctverce, Mapa_P20, Mapa_P50.

Tabulka 13 - Struktura ukázkového Excel exportu programu RADOIL.

List v ukázkovém Excelu	Obsah	Poznámka
Data	Tabulka bodů ID, X, Y, value.	Základní auditní stopa importovaných bodů.
Výsledky	Metriky Čtverce, P20, P50.	Hlavní tabulkový souhrn vyhodnocení.
Shoda_2_metod	Mapa bodů splňujících shodu alespoň dvou metod.	Prioritní interpretační vrstva.
Shoda_3_metod	Mapa bodů splňujících shodu tří metod.	Nejrobustnější shoda.
Mapa_ctverce	Mapa metody Čtverce.	Prostorově kontextová detekce.
Mapa_P20	Mapa metody P20.	Konzervativní bodová detekce.
Mapa_P50	Mapa metody P50.	Citlivá bodová detekce.

8.3 Využití exportů v dokumentaci lokality

- PDF export je vhodné přiložit jako neměnnou výsledkovou zprávu z konkrétního běhu programu.
- Excel export je vhodné archivovat jako pracovní tabulkovou stopu, která umožňuje kontrolu hodnot a případný import do dalších nástrojů.
- Oba exporty musí být svázány s konkrétní verzí vstupních dat a s datem zpracování.
- Do průzkumné zprávy se doporučuje přenášet zejména mapy shody, tabulku výsledků a stručné interpretační shrnutí.

9. Ukázkový export a interpretace výsledků

Tato kapitola představuje příklady exportů z programu RADOIL a jejich interpretaci. Slouží jako praktický návod k demonstraci struktury výstupů. Hodnoty uvedené v této kapitole jsou ukázkové.

9.1 Ukázková tabulka výsledků analýzy

Tabulka 14 - Ukázková tabulka výsledků analýzy z exportu RADOIL.

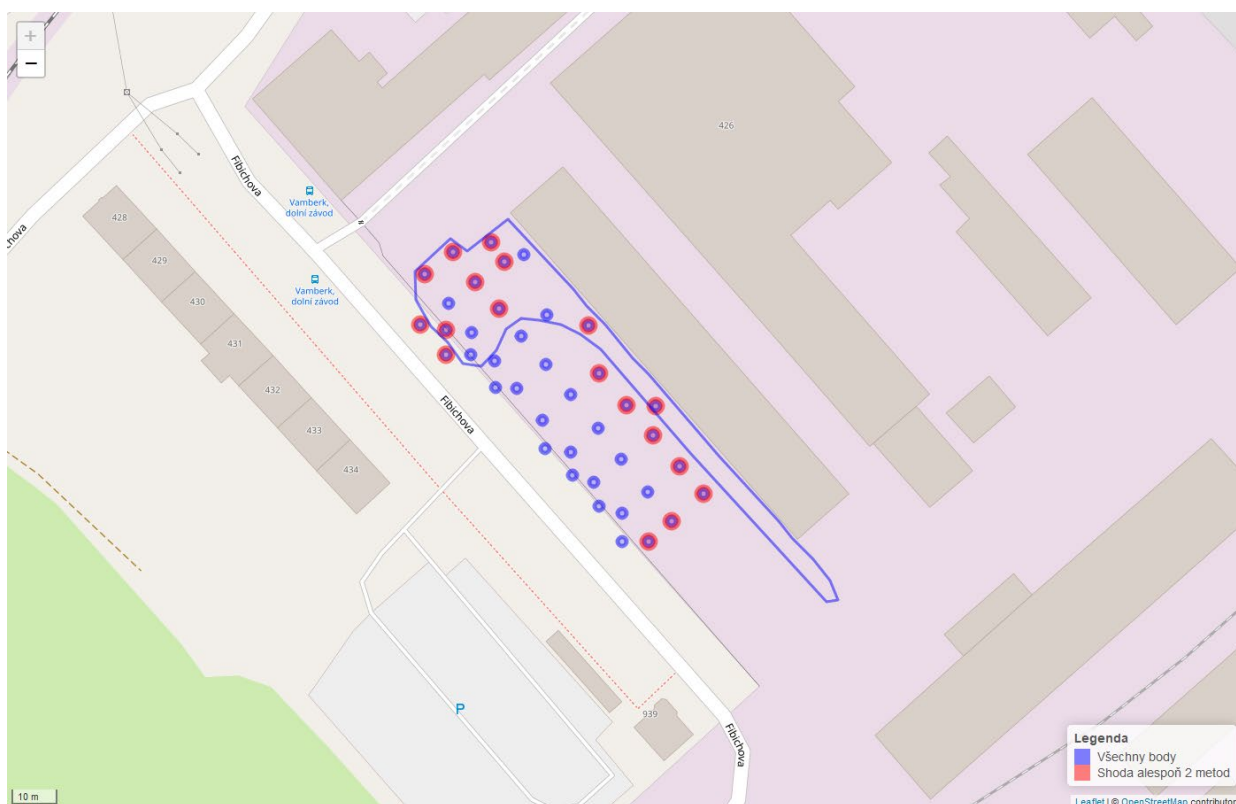
Metrika	Čtverce	P20	P50
Správně uvnitř	11	4	9
% správně uvnitř	78.57	28.57	64.29
Falešně mimo	13	4	11
% falešně mimo	50	15.38	42.31
Neidentifikované	16	32	20
Neidentifikované uvnitř	3	10	5
Interpretační poznámka	Nejvyšší záchyt uvnitř polygonu, ale zároveň nejvyšší zatížení falešnými detekcemi mimo polygon.	Nejpřísnější výběr; nízká falešná detekce mimo polygon, ale vysoký počet neidentifikovaných bodů uvnitř.	Citlivější bodová varianta; v ukázce zachycuje větší část polygonu než P20, ale s vyšší falešnou detekcí.

Z tabulky vyplývá, že ve validačním NAPL polygonu bylo v ukázkovém příkladu celkem 14 bodů (správně uvnitř + neidentifikované uvnitř). Celkový počet měřených bodů je 40 (správně uvnitř + falešné mimo + neidentifikované). Metoda Čtverce zachytila 11 z 14 bodů uvnitř polygonu, tedy 78,57 %, ale současně označila 13 bodů mimo polygon, což odpovídá 50,00 % bodů mimo referenční oblast. P20 je v tomto exportu nejkonzervativnější: zachytil pouze 4 body uvnitř polygonu, ale falešně označil jen 4 body mimo polygon. P50 zachytil 9 bodů uvnitř polygonu, avšak s 11 falešnými body mimo polygon.

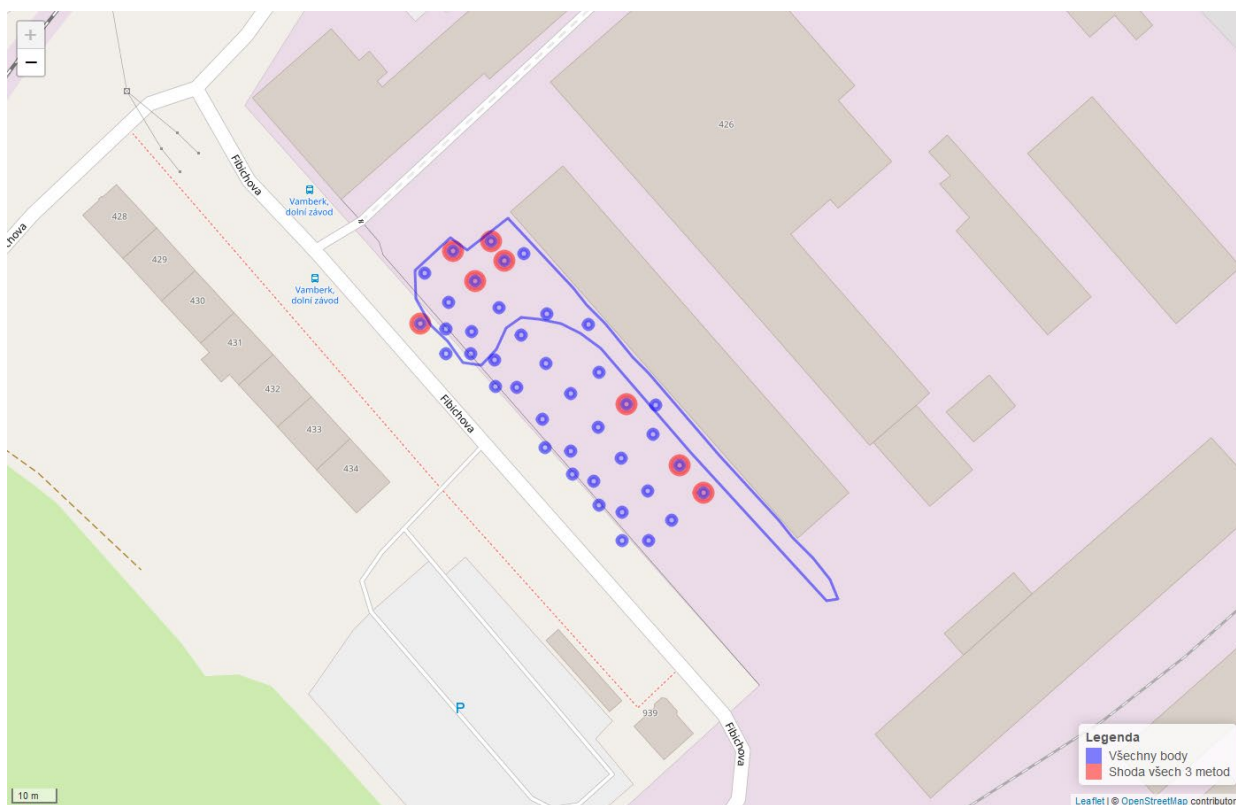
Interpretačně to znamená, že Čtverce v této ukázce maximalizují záchyt uvnitř referenční oblasti, zatímco P20 minimalizuje falešnou pozitivitu. P50 zůstává citlivější bodovou metodou vhodnou pro širší screening, ale musí být posuzována spolu s mapou shody a s prostorovým kontextem bodů.

9.2 Mapové výstupy

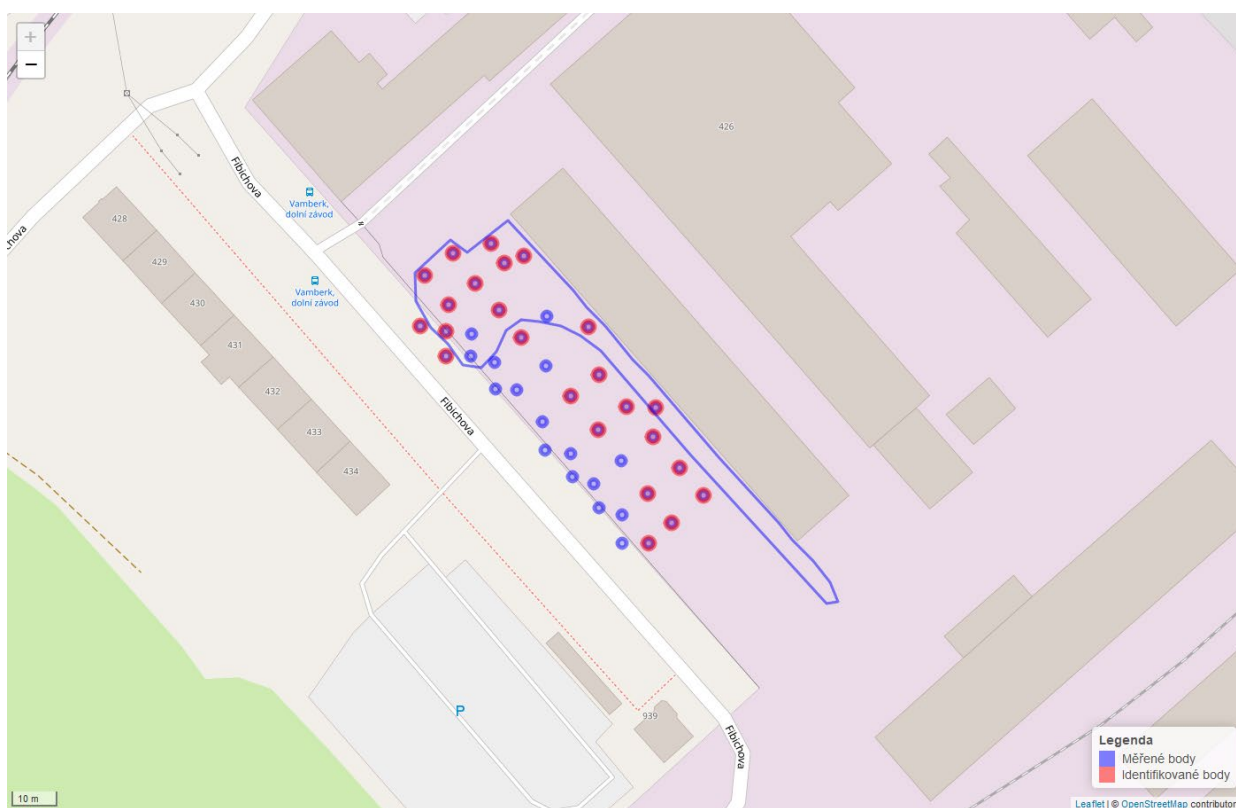
Mapové výstupy převádějí tabulkovou klasifikaci do prostorové interpretace. Modré body představují všechny měřené body, červené body body identifikované danou metodou nebo shodou metod a modrá linie vymezuje referenční hranici polygonu NAPL. Pro dokumentaci lokality se doporučuje ukládat minimálně mapu shody alespoň dvou metod, mapu shody všech tří metod a mapy jednotlivých metod Čtverce, P20 a P50.



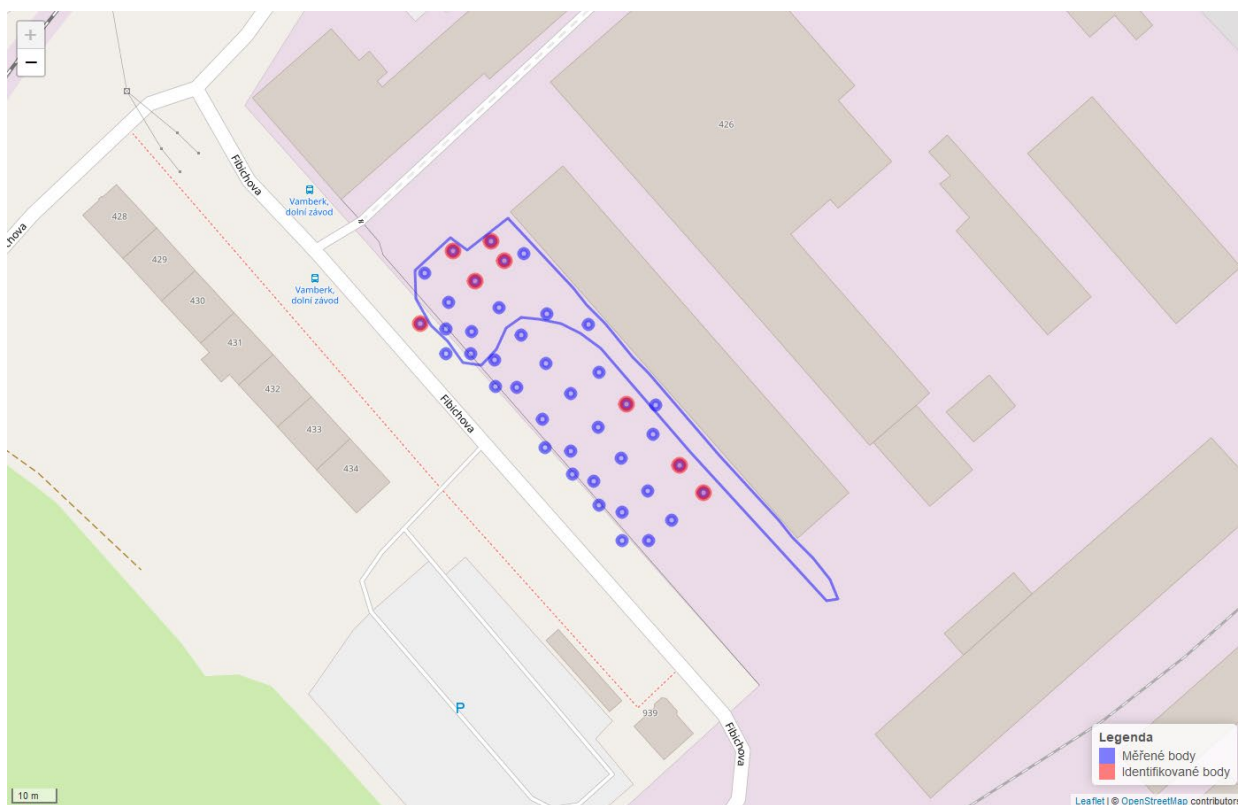
Obrázek 1 - Mapa shody alespoň 2 metod analýzy. Červeně jsou zvýrazněny body, u nichž se shodují minimálně dvě vyhodnocovací metody; modře jsou zobrazeny všechny měřené body a referenční hranice.



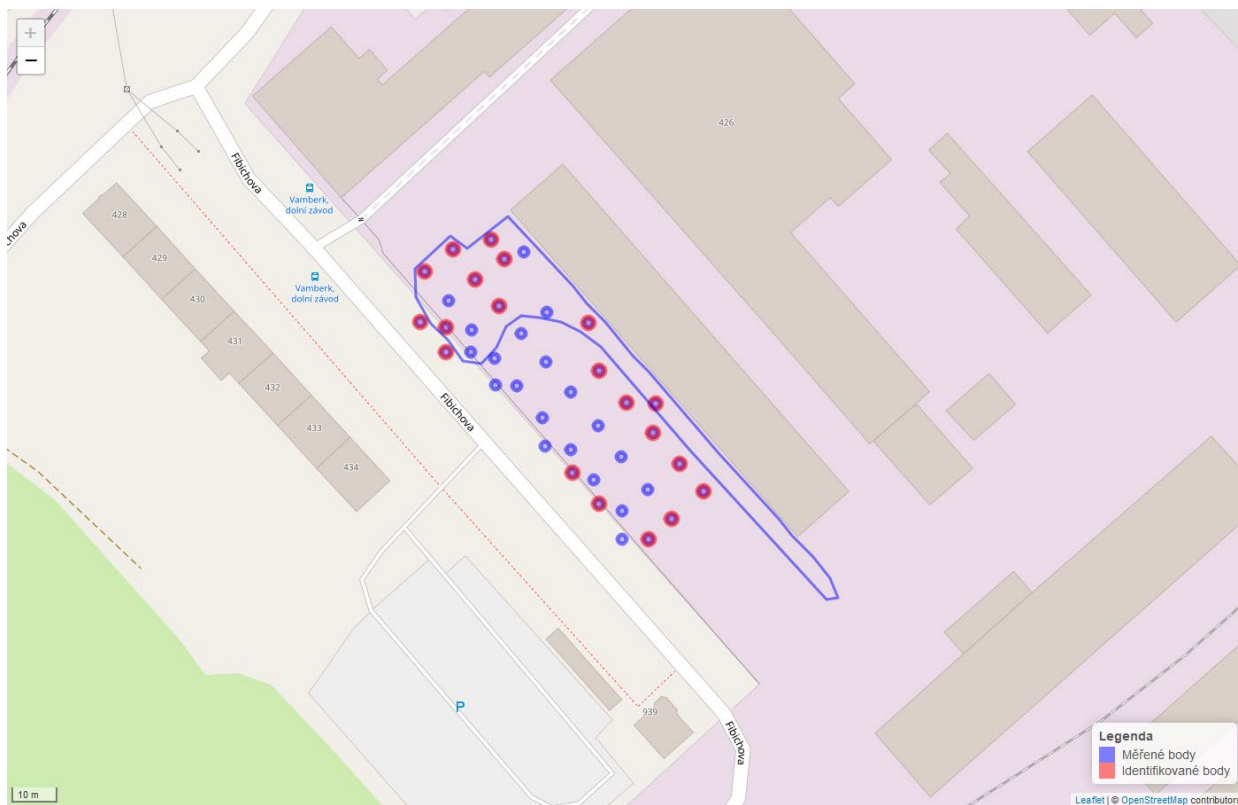
Obrázek 2 - Mapa shody 3 metod analýzy. Zobrazuje body označené současně všemi hlavními metodami programu RADOIL.



Obrázek 3 - Mapa metody Čtverce. Výstup prostorově kontextové metody určené pro zachycení souvislejších lokálních poklesů OAR.



Obrázek 4 - Mapa metody percentil 20. Konzervativní bodová detekce nejnižších hodnot OAR.



Obrázek 5 - Mapa metody percentil 50. Citlivější bodová detekce širšího poklesu OAR.

9.3 Ukázkový datový list Data

Datový list představuje primární auditní stopu bodových měření použitých pro výpočet. Obsahuje ID bodu, souřadnice X a Y a hodnotu OAR. V ukázkovém exportu je uvedeno 40 bodů. V praxi se kompletní datový list archivuje v Excelu. .

Tabulka 15 - Ukázkový datový list Data z exportu RADOIL.

ID	X	Y	OAR
1	16.28057	50.11489	3.90000
2	16.28065	50.11483	17.40000
3	16.28072	50.11483	58.50000
4	16.28080	50.11477	87.20000
5	16.28095	50.11465	68.50000
6	16.28103	50.11460	26.40000
7	16.28111	50.11454	16.80000
8	16.28118	50.11447	34.70000
9	16.28065	50.11488	18.50000
10	16.28078	50.11505	13.10000
11	16.28108	50.11489	23.00000
12	16.28128	50.11473	15.10000
13	16.28088	50.11502	30.30000
14	16.28126	50.11447	21.80000
15	16.28118	50.11452	33.76714
16	16.28109	50.11458	37.80885
17	16.28103	50.11464	40.05158
18	16.28094	50.11470	106.00000
19	16.28086	50.11476	107.00000
20	16.28079	50.11482	95.60000
21	16.28072	50.11487	44.70000
22	16.28066	50.11493	27.30000
23	16.28058	50.11499	26.10000
24	16.28067	50.11503	3.90000
25	16.28074	50.11497	2.60000
26	16.28081	50.11492	19.60370
27	16.28087	50.11487	43.70000
28	16.28095	50.11481	81.50000
29	16.28103	50.11475	65.60000
30	16.28111	50.11469	63.50000
31	16.28118	50.11463	58.50000
32	16.28126	50.11456	48.20000
33	16.28133	50.11451	18.10000
34	16.28143	50.11456	14.40000
35	16.28135	50.11461	9.90000
36	16.28127	50.11467	17.00000
37	16.28119	50.11473	14.40000
38	16.28111	50.11479	17.00000
39	16.28095	50.11491	79.10000
40	16.28082	50.11501	6.40000

9.4 Ukázkový výpis hranice

Výpis hranice slouží k auditní kontrole polygonu použitého pro validační porovnání.

Tabulka 16 - Ukázkový výpis souřadnic hranice NAPL z exportu RADOIL.

pořadí	X	Y
1	16.28066	50.11506
2	16.28060	50.11502
3	16.28055	50.11499
4	16.28055	50.11494
5	16.28060	50.11489
6	16.28065	50.11485
7	16.28070	50.11481
8	16.28075	50.11481
9	16.28080	50.11484
10	16.28083	50.11488
11	16.28087	50.11490
12	16.28093	50.11490

pořadí	X	Y
13	16.28099	50.11489
14	16.28105	50.11487
15	16.28112	50.11484
16	16.28128	50.11472
17	16.28133	50.11468
18	16.28139	50.11464
19	16.28144	50.11460
20	16.28156	50.11452
21	16.28164	50.11446
22	16.28180	50.11435
23	16.28184	50.11435
24	16.28181	50.11439
25	16.28176	50.11443
26	16.28170	50.11447
27	16.28165	50.11451
28	16.28159	50.11455
29	16.28159	50.11455
30	16.28148	50.11463
31	16.28126	50.11479
32	16.28121	50.11482
33	16.28116	50.11486
34	16.28113	50.11489
35	16.28107	50.11493
36	16.28103	50.11496
37	16.28099	50.11499
38	16.28094	50.11502
39	16.28090	50.11505
40	16.28087	50.11507
41	16.28084	50.11509
42	16.28076	50.11506
43	16.28084	50.11509
44	16.28071	50.11503
45	16.28066	50.11506

9.5 Stručná interpretační syntéza ukázky

Nejvyššího zachytu uvnitř referenčního polygonu je dosaženo metodou Čtverce. Tato metoda však zároveň přináší nejvíce falešných detekcí mimo polygon, a proto by neměla být použita izolovaně jako jediný podklad pro vymezení NAPL. P20 je naopak nejpřísnější a má nejnižší míru falešné detekce mimo polygon, ale v této ukázce ponechává většinu bodů uvnitř polygonu jako neidentifikované. P50 představuje širší bodový screening a může sloužit jako doplňková vrstva pro vymezení potenciální okrajové zóny poklesu OAR.

Pro praktickou interpretaci je proto vhodné vymezit jádrovou zónu podle shody 3 metod, případně podle shody alespoň 2 metod. Výsledek metody Čtverce lze využít pro posouzení prostorové souvislosti poklesu OAR, P20 pro konzervativní potvrzení nejvýraznějších anomálií a P50 pro širší screeningovou zónu. Body nebo skupiny bodů, které leží v jádrové zóně shody a současně mají hydrogeologicky smysluplnou polohu vůči známému směru proudění nebo zdrojové oblasti, mají nejvyšší prioritu pro následné přímé ověření.

Závěr z ukázkového exportu má být formulován jako indikace pravděpodobné zóny výskytu NAPL, nikoliv jako přímé analytické prokázání ropných uhlovodíků. K potvrzení interpretace se doporučuje využít vrtný průzkum a měření mocnosti volné fáze na hladině podzemní vody. V případě rozdílu mezi metodami je vždy nutné posoudit kvalitu terénních dat, plynopropustnost, geologickou homogenitu a možné lokální vlivy na transport půdního plynu.

10. Nejistoty, omezení a požadavky na dokumentaci

10.1 Hlavní zdroje nejistoty

Tabulka 17 - Hlavní zdroje nejistoty při statistickém vyhodnocení OAR.

Zdroj nejistoty	Možný dopad na výsledek	Doporučená reakce
Geologická heterogenita	Pokles OAR může souviset s litologií, plynopropustností nebo navážkami, nikoli s NAPL.	Dělit území na homogenní celky, uvést vyšší nejistotu, ověřit přímou metodou.
Nízká plynopropustnost	Odběr půdního vzduchu může být nespolehlivý a variabilita OAR vysoká.	Zaznamenat odpor sání, posoudit vhodnost lokality, případně metodu nepoužít.
Meteorologické změny	Mezidenní změny OAR mohou překrýt statistický signál.	Použít kontrolní bod; při rozdílu > 30 % vyhodnocovat dny odděleně.
Nerovnoměrná síť bodů	Metody nejbližších sousedů mohou vybírat nevhodné skupiny.	Před výpočtem zkontrolovat mapu bodů a případné mezery v síti.
Chybný souřadnicový systém	Nesprávné vzdálenosti, špatná poloha vůči polygonu a chybné validační metriky.	Ověřit transformaci a překryv s mapou před výpočtem.
Izolovaná chyba odběru	Jednotlivý bod může být falešně označen jako nízký.	Využít mapy shody a prostorové metody; zkontrolovat terénní poznámky.
Nesouvislá NAPL	Deficit radonu může být slabý nebo prostorově nejednoznačný.	Vyhodnotit výsledek jako indikativní a preferovat přímé ověření.

10.2 Pravidla pro opakované kampaně

Při dlouhodobém monitoringu je nutné zachovat maximální srovnatelnost kampaní. To znamená používat stejná ID bodů, stejnou síť, stejnou odběrovou hloubku, shodný typ přístroje, obdobné roční období a stejný způsob statistického vyhodnocení. Není vhodné měnit mezi kampaněmi percentilový práh ani interpretovat rozdíl výsledků, pokud je změna způsobena změnou metodiky výpočtu.

- Pro trendové hodnocení vždy používat stejnou kombinaci metod, například P20, P50, Čtverce a shoda alespoň 2 metod.
- Změny polohy nebo rozsahu anomální zóny interpretovat pouze společně s hydrogeologickými informacemi a s poznámkami o podmínkách měření.
- Pokles nebo nárůst počtu anomálních bodů nemusí sám o sobě znamenat změnu kontaminace, pokud se změnily podmínky odběru nebo počasí.
- Výraznější změny trendu ověřit přímou metodou.

10.3 Doporučený text do zprávy

Do průzkumné zprávy se doporučuje uvést minimálně následující informace:

- počet měřených bodů, datum měření, jednotky a souřadnicový systém,
- stručný výsledek QA/QC kontroly včetně případných vyřazených nebo opakovaných bodů,
- použité statistické metody a percentilové prahy,
- mapy jednotlivých metod a mapy shody,
- tabulku validačních metrik, pokud byl k dispozici polygon NAPL,
- vymezení prioritní zóny podezření a doporučení přímého ověření,
- popis nejistot a limitů interpretace.

Formulace závěru

Výsledky RADOIL je vhodné formulovat jako „indikaci pravděpodobné zóny výskytu NAPL podle poklesu OAR“, nikoli jako definitivní důkaz kontaminace. Přímé potvrzení kontaminace musí být provedeno referenčními průzkumnými metodami.

11. Kontrolní list pro zpracovatele měření

Tabulka 18 - Kontrolní list pro zpracovatele dat RADOIL.

Oblast	Kontrolní otázka	Splněno / poznámka
Vstupní data	Obsahuje tabulka jednoznačné ID, X, Y a OAR/value pro všechny body?	
Jednotky	Jsou hodnoty OAR v jednotkách kBq.m-3?	
Souřadnice	Jsou body a polygon ve stejném souřadnicovém systému?	

Oblast	Kontrolní otázka	Splněno / poznámka
Mapa bodů	Odpovídá poloha bodů skutečné lokalitě a terénní síti?	
Rozsah měření	Je počet bodů a hustota sítě dostatečná pro statistické vyhodnocení?	
Homogenita	Neobsahuje soubor výrazně odlišné geologické celky, které je nutné vyhodnotit odděleně?	
Odběrová hloubka	Byla zachována jednotná odběrová hloubka a postup odběru?	
Extrémy	Byly prověřeny hodnoty $\leq 1 \text{ kBq.m}^{-3}$ a extrémně vysoké hodnoty?	
Vícedenní měření	Byl vyhodnocen kontrolní bod a případný rozdíl $> 30 \%$?	
RADOIL výpočet	Byly vyhodnoceny metody P20, P50 a Čtverce?	
Shoda metod	Byly posouzeny mapy shody alespoň 2 metod a shody 3 metod?	
Validace	Byla při dostupném polygonu NAPL vyhodnocena tabulka metrik?	
Interpretace	Je závěr formulován jako indikace, nikoli jako definitivní důkaz NAPL?	
Export	Byl uložen PDF i Excel export včetně verze vstupních dat?	
Archivace	Jsou archivovány primární protokol, očištěná data, hranice, exporty a záznam zpracování?	