



TECHNICKÁ UNIVERZITA V LIBERCI
Fakulta mechatroniky, informatiky
a mezioborových studií ■

JEDNODUCHÉ ALTERNATIVNÍ MODELY PODZEMNÍ VODY

Autoreferát disertační práce

Studijní program: P 3901 Aplikované vědy v inženýrství
Studijní obor: 3901V055 Aplikované vědy v inženýrství

Autor práce: **Mgr. Kamil Nešetřil**
Školitel: doc. Ing. Jan Šembera, Ph.D.

Poděkování

Tato práce vznikla mj. za podpory Technologické agentury České republiky v rámci řešení projektů č. TA02020177 „Informační systém pro podporu rozhodování o využití krajiny po rekultivaci (MARE)“ a č. TA04020207 „Informační systém pro analýzy a predikce zásob podzemní vody v závislosti na lidské činnosti a klimatických změnách“ (ZÁVOD).

Abstrakt

Epistemologická nejistota modelů podzemní vody se projevuje v nejistotě koncepčního modelu, a to zejména ve struktuře modelu. S ní je možno se vyrovnat využitím více (i velmi jednoduchých) koncepčních modelů založených na různých přístupech či předpokladech. Práce shrnuje existující klasifikace nejistot (zejm. koncepčního modelu) a navrhuje novou klasifikaci využití modelů. Prezentované případové studie srovnávají jednoduché a komplexní modely, testují alternativní hypotézy sadou jednoduchých modelů či představují alternativní prediktivní modely. Taková strategie se ukazuje jako velmi účelná obzvláště při nedostatku vstupních dat, kdy není možné model validovat. Obhajitelnost modelů je semikvantitativně vyhodnocena. Aby takové modely bylo možno snadno formulovat, vytvářet a využívat, byl vyvinut informační systém využívající nástroje, přístupy a technologie business intelligence.

Klíčová slova

hydrogeologie, podzemní voda, matematický model, nejistota, epistemologická nejistota, koncepční model, struktura modelu, ekvifinalita, alternativní modely, jednoduchost software pro správu dat o životním prostředí (EDMS), business intelligence (BI), Pentaho, ETL, GeoBI, datový model

Abstract

Epistemological uncertainty of groundwater models is reflected in the conceptual model uncertainty, especially in the model structure. It is possible to cope with it by using more (even very simple) conceptual models based on different approaches or assumptions. The thesis summarizes the existing classifications of uncertainty (esp. that of conceptual model), and proposes a new classification of model usage. The presented case studies compare simple and complex models, test alternative hypotheses by a set of simple models or represent alternative predictive models. Such a strategy is very effective especially in the absence of adequate input data when it is not possible to validate the model. Tenability of the models is evaluated semi-quantitatively. For such models to be easy to formulate, develop and exploit, an information system using tools, approaches and technologies of business intelligence has been developed.

Keywords

hydrogeology, groundwater, mathematical model, uncertainty, epistemic uncertainty, conceptual model, model structure, equifinality, multiple model ensemble, parsimony, simplicity environmental data management software (EDMS), business intelligence (BI), Pentaho, ETL, GeoBI, data model

Obsah

Seznam zkratk.....	4
1 Úvod.....	5
1.1 Matematické modely podzemní vody	6
1.2 Klasifikace a hodnocení nejistoty (modelu)	6
1.3 Alternativní koncepční modely	8
1.3.1 Volba nejlepšího modelu.....	8
1.4 Jednoduchost a komplexnost modelů podzemní vody	9
1.5 Alternativní jednoduché modely	9
1.6 Princip ekvifinality.....	10
1.7 Formulace zvoleného přístupu	10
2 Případové studie.....	12
2.1 Ekvifinalita – interakce povrchové a podzemní vody.....	12
2.2 Srovnání jednoduchého a komplexního modelu – ovlivnění podzemních vod jezem	14
2.3 Kombinace principiálně odlišných modelů.....	16
2.3.1 Model ustálené délky kontaminačního mraku.....	16
2.3.2 Model dovoleného čerpání.....	17
2.4 Transport ropných uhlovodíků.....	18
2.5 Transport chlorovaných uhlovodíků.....	19
2.5.1 Advekční model – predikce.....	19
2.5.2 Bilanční model – testování hypotézy	20
2.5.3 Diskuse a závěr.....	20
2.6 Přirozená biodegradace ropných uhlovodíků	20
2.7 Hodnocení nejistoty koncepčního modelu podle Refsgaarda et al. 2006.....	21
3 Informační systém HgIS	21
3.1 Existující EDMS, výměnné formáty a datové modely	21
3.1.1 Existující EDMS	22
3.1.2 Existující výměnné formáty a datové modely	22
3.1.3 Hodnocení stávajících řešení.....	23
3.2 Typy informačních systémů	23
3.3 Business intelligence	25
3.4 Návrh informačního systému HgIS.....	25
3.4.1 Geologická vizualizace – EnviroInsite	26
3.4.2 Databáze	26
3.4.3 ETL	27
3.4.4 Reporting (tiskové sestavy).....	28
3.4.5 Mapový server.....	28
3.5 Implementace.....	29
3.5.1 ETL	29
3.5.2 Reporting (tiskové sestavy).....	29
3.5.3 Mapová online aplikace a správa nestrukturovaných dat a dokumentů	30
3.6 Aplikace HgIS.....	30
3.7 Diskuse a závěr	30
3.7.1 Plánované rozšíření	32
4 Shrnutí vědeckých výsledků.....	32
5 Shrnutí a závěr	33
Použitá literatura.....	34

Seznam zkratek

BI	<i>Business intelligence</i> – využívání dat o firmě pro podporu rozhodování manažerů
CIU	Chlorované uhlovodíky
DNAPL	<i>Dense non-aqueous phase liquids</i> – Kapalná fáze těžší než voda
DS	Distribučním sklad pohonných hmot
DSS	<i>Decision support system</i> – Systém pro podporu rozhodování
EDMS	<i>Environmental data management software</i> – Sw pro správu dat o životním prostředí
DWH	<i>Data warehouse</i> – Datový sklad: databáze aktuálních i historických dat pro analýzu
EI	EnviroInsite – software pro vizualizaci hydrogeologických dat
ETL	<i>Extract, transform, load</i> – Datová pumpa
FOSS	<i>Free and open source software</i> (téže FLOSS) – software zdarma s otevřenou licencí
GeoBI	<i>Geospatial business intelligence</i>
GIS	Geografický informační systém
GUI	<i>Graphical user interface</i> – Grafické uživatelské rozhraní
HgIS	Hydrogeologický informační systém – významný výsledek této práce
IF	Impakt faktor
IS	Informační systém
JÚ	Vodárenské jímací území
LIMS	<i>Laboratory information management system</i> – Laboratorní informační systém
n. s.	Nerostné suroviny
ODBC	<i>Open Database Connectivity</i>
OGS	<i>Open Geospatial Consortium</i> – organizace mj. vydávající geoinformační standardy
OLAP	<i>Online Analytical Processing</i>
OLTP	<i>Online Transaction Processing</i>
PBA	<i>Pentaho Business Analytics</i> – nasazený server firmy Pentaho
PCE	Tetrachlorethylen
PDI	<i>Pentaho Data Integration</i> – neboli Kettle: ETL od firmy Pentaho
QA/QC	<i>Quality Assurance / Quality Control</i> – Jištění jakosti a kontrola kvality
RU	Ropné uhlovodíky
SQL	<i>Structured Query Language</i> – dotazovací jazyk relačních databází
SŘBD	Systém řízení báze dat = DBMS – <i>Database Management System</i>
SW, sw	Software
TCE	Trichlorethylen
TUL	Technická univerzita v Liberci
WFS	<i>Web Feature Service</i> – standard webové služby pro poskytování vektorových map
WMS	<i>Web Map Service</i> – standard webové služby pro poskytování rastrových map
WMTS	<i>Web Map Tile Service</i>
XML	<i>Extensible Markup Language</i>

1 Úvod

Celý profesní život se autor předkládané práce zabývá modelováním podzemní vody a po počáteční fascinaci komplexními modely zjišťoval, že složité modely s vysokým počtem stupňů volnosti (parametrů) často nejsou užitečným podkladem pro porozumění probíhajícím procesům ani pro rozhodování – mimo jiné proto, že často nejsou adekvátně podpořeny daty. Navíc je obtížné je sestavit a výsledky vyhodnotit. Předkládaná práce proto demonstruje užitečnost využití více jednoduchých modelů. Pro usnadnění jejich tvorby byl vyvinut informační systém HgIS.

V modelování podzemní vody čelíme nejistotě ve výsledcích modelu. Ta je dána nejistotou vstupních dat a nejistotou naší interpretace těchto dat (tzv. epistemologická nejistota – tzn. nedostatečné pochopení fungování systému). Při modelování podzemní vody je dobrou praxí hodnotit nejistotu predikcí na základě shody modelu s validačními daty. Skutečná nejistota je však často hlubší – podzemí totiž není přístupné přímému pozorování, a proto často bývá důležitější nejistota interpretací. Tyto úvahy vyžadují jednoznačné vymezení pojmů, které však nejsou v literatuře používány jednoznačně – v první kapitole je proto představeno shrnutí existujících klasifikací nejistot. Je zde definována epistemologická nejistota a její projevy: nejistota koncepčního modelu a nejistota struktury modelu. Pro vypořádání se s takovou nejistotou je v práci obhajována a využita strategie více jednoduchých modelů, které jsou založeny na různých přístupech či předpokladech. Tyto jednoduché modely mohou být formulovány jako testování hypotéz a jednoznačně snižovat nejistotu porozumění fungování studovaného systému. Využití modelů pro testování hypotéz je spolu s dalšími uvedeno v navržené klasifikaci využití modelů (kap. 1.7). V práci je diskutováno použití principu ekvifinality a tohoto pojmu v oblasti modelování podzemní vody. První kapitola vychází z článků zabývajících se zejm. hydrogeologií, i když principy využití alternativních jednoduchých modelů jsou platné pro různé obory.

V kap. 2 je prezentováno využití alternativních jednoduchých modelů v modelování různých hydrogeologických úloh. Jedna úloha je řešena více modely, které jsou založeny na různých přístupech a předpokladech. Vysoká nejistota modelů je dána nedostatkem adekvátních vstupních dat. Proto také nebylo možno výsledky modelů kalibrovat ani validovat, ale případně pouze srovnávat výsledky jednotlivých modelů mezi sebou. Prezentované případové studie srovnávají jednoduché a komplexní modely, testují alternativní hypotézy sadou jednoduchých modelů či představují alternativní prediktivní modely. Obhajitelnost modelů je semikvantitativně vyhodnocena s využitím metodiky nalezené v literatuře (kap. 2.7).

Postupně stále více vyvstávala potřeba efektivní správy dat, která je předpokladem pro tvorbu alternativních koncepčních modelů a o pro vývoj znovu využitelných matematických modelů. K tomu je možno využít existující software pro správu dat o podzemní vodě a životním prostředí, které se však ukázaly jako nevyhovující. Velmi se však osvědčilo využití nástrojů, přístupů a technologií z naprosto jiné oblasti – manažerské informatiky. Úlohy řešené při zpracování dat o podzemní vodě a životním prostředí jsou do značné míry podobné úlohám business intelligence. Business intelligence spočívá ve sběru a analýze dat (např. o firmě) pro podporu manažerského rozhodování. Manažer i hydrogeolog potřebují načítat archivní data z různorodých zdrojů (databáze, soubory atd.). Není ani tak důležité interaktivně editovat data, ale především je zobrazovat a analyzovat. Oba potřebují tiskové sestavy, které jsou automaticky generovány, ale přitom je možno je ručně upravovat. Je proto vhodné, aby byly mj. v editovatelných formátech, jako je MS Word a MS Excel. Proto byl vyvinut informační systém HgIS (kap. 3). Pro vizualizaci a analýzu dat a pro matematické modelování podzemní vody existuje mnoho užitečných nástrojů. Jen v některých případech je rozumné implementovat jejich funkcionalitu přímo do

HgIS. Důležitou součástí HgIS je proto možnost, aby si uživatel mohl interaktivně vytvořit export dat pro využití v libovolném nástroji. Při vývoji HgIS se uplatnily přístupy, nástroje a technologie business intelligence jako je např. datová pumpa (tzv. ETL), která umožňuje automatizovaně importovat data z různých zdrojů a to i málo strukturovaných, jako je MS Word. ETL byl též využit pro export dat.

Pro rychlou orientaci je tedy možno říci, že první kapitola představuje teoretický úvod do problematiky nejistoty, alternativních modelů a jednoduchosti v modelování. Druhá kapitola přináší ukázky praktického využití principů z první kapitoly na modelech různých lokalit. Třetí kapitola prezentuje systém, který usnadňuje práci s hydrogeologickými daty včetně jejich interpretace a formulace adekvátních koncepčních modelů a implementaci jednoduchých matematických modelů. Výsledkem práce není formální metodika – uvedené principy je však možno uplatnit při zpracování dalších případových studií. Předkládaná práce má do jisté míry syntetický charakter a může být tudíž obtížněji uchopitelná.

1.1 Matematické modely podzemní vody

Každý model je abstrakcí, zjednodušením a interpretací reality. Podle Bevena a Younga (2013) neexistují (alespoň zatím) modely v hydrologii plně založené na fyzikálních principech. Autor předkládané práce se domnívá, že se v hydrogeologické praxi často přeceňuje fyzikální pojetí úlohy, ačkoliv v důsledku heterogenit a měřítko není toto pojetí plně možné. Alternativou k těmto redukcionistickým modelům (*deterministic reductionism*) jsou techniky založené na datech. Přístupy prezentované v předkládané práci jsou většinou fyzikálně založené modely, i když jejich matematická formulace není vždy plně korektní. Většímu využití alternativních strategií modelování však může přispět informační systém HgIS prezentovaný v kap. 3.

V literatuře je používána rozličná terminologie ohledně fází tvorby modelu. Dále je používána terminologie Bevena (2001, 2012), podle níž percepční model představuje subjektivní porozumění (vnímání) probíhajícím procesům, které může být čistě kvalitativní. Koncepční model je jeho konkrétní zjednodušené vyjádření (např. řídicí rovnice, doména, okrajové podmínky, parametry), které se týká konkrétní lokality. Procedurální model je pak implementace koncepčního modelu v konkrétním software (sw) bez ohledu na to, zda jde o numerické, či analytické řešení.

Ověřování modelu má tři stupně – mohli bychom říci s klesající silou pojmu: verifikace, validace a potvrzení. Validace nespočívá v testování úplnosti modelu či jeho absolutní pravdivosti (jako u vědecké teorie), ale ověřuje, zda je model přijatelný pro zvolený účel. Koncepční model není možné verifikovat či validovat, nýbrž je možné jej pouze potvrdit (*confirmation*). Každý koncepční model je zjednodušením probíhajících procesů a může být proto vyvrácen (falzifikován), pokud jej prozkoumáme do dostatečného detailu a uplatníme na něj dostatečně vysoká kritéria.

1.2 Klasifikace a hodnocení nejistoty (modelu)

Výsledky matematického modelu nemají význam, pokud není zřejmá jejich důvěryhodnost. Proto je vhodné ověřovat model a hodnotit nejistotu modelu. Kreye et al. (2011) uvádí klasifikaci nejistot včetně nejistot modelu. Z klasifikace plyne rozdíl mezi nejistotou koncepčního modelu a strukturální nejistotou modelu, což jsou pojmy, které jsou v literatuře hojně používány. Z článku je převzata tabulka 1, jež byla doplněna dalšími údaji z literatury. Hvězdička naznačuje druhy nejistot důležité z hlediska této disertace.

Tabulka 1: Klasifikace nejistoty podle pěti vrstev

Vrstva	Popis
Povaha (Nature)	Všeobecná charakteristika nejistoty: <ul style="list-style-type: none"> • stochastická (<i>aleatory</i>) čili ontologická (daná inherentní variabilitou) <ul style="list-style-type: none"> ○ variabilita (nahodilé chování systému) ○ nejednoznačnost (rozpory v informacích) • *epistemologická čili systematická (daná nedokonalou znalostí)
Příčina (Cause)	Důvod či zdroj nejistoty: <ul style="list-style-type: none"> • *nedostatek porozumění • nejednoznačnost • lidská činnost
Úroveň (Level)	Podle Warminka et al. (2010) a Walkera et al. (2003): <ul style="list-style-type: none"> • statistická (nejistota může být kvantifikována) • *nejistota scénářů (nejistota může být popsána alternativami chování systému) • kvalitativní (nejistota může být popsána) • rozpoznaná neznalost (nejistota nemůže být popsána) • naprostá neznalost
Projev (Manifestation)	Místo v procesu, kde se nejistota projeví: <ul style="list-style-type: none"> • nejistota kontextu (z vnějšku modelu – co model nepostihuje – např. účel) <ul style="list-style-type: none"> ○ endogenní (z vnitřku organizace) – přírodní, technologické, ekonomické ○ exogenní (z vnějšku organizace) • nejistota dat – vstupů, pozorování <ul style="list-style-type: none"> ○ neúplnost dat (mezery v dostupných datech) ○ nepřesnost dat (nepřesnost či nespolehlivost dostupných dat) ○ *variabilita dat (různé možné alternativy) • nejistota modelu <ul style="list-style-type: none"> ○ *koncepční (zjednodušení v koncepčním modelu) <ul style="list-style-type: none"> ▪ *nejistota struktury modelu (neadekvátnost či diskrepance modelu) – celkový nedostatek porozumění <ul style="list-style-type: none"> • opomenutí procesů (řídící rovnice) • (příliš malý) rozsah modelu • geologická nejistota • ... ▪ nejistota parametrů modelu <ul style="list-style-type: none"> • přesné parametry – např. π • fixní parametry – např. gravitační zrychlení • a priori zvolené parametry – obtížné identifikovat • kalibrované parametry – např. transmisivita <ul style="list-style-type: none"> ○ nejistota kalibrace z důvodu chyb v datech ○ nejednoznačnost z důvodu necitlivosti parametrů ○ nejednoznačnost z důvodu korelace parametrů ○ technická stránka modelu: <ul style="list-style-type: none"> ▪ matematická čili algoritmická (zjednodušení v matemat. vyjádř.) ▪ výpočetní čili numerická (zjednodušení ve výpočetní metodě) • fenomenologická nejistota (neznámé události v budoucnosti)
Umístění (Location) (Walker et al. 2003; Warmink et al. 2010)	
Vyjádření (Expression)	Způsob jak je nejistota vyjádřena či sdělena. Může být: <ul style="list-style-type: none"> • kvantitativní (měřitelná) • kvalitativní (neměřitelná).

S nejistotou v modelech je možno se vyrovnat pomocí metod jako je Monte Carlo, stochastických / pravděpodobnostních modelů, variantních výpočtů – alternativních scénářů, alternativních koncepčních modelů a dalších metod. Metody pro hodnocení (*evaluating*) nejistoty jsou zároveň jedním z přístupů pro vypořádání se (*coping*) s nejistotou. V případě nedostatku dat pro kalibraci a validaci modelu je třeba využívat semikvantitativní a kvalitativní hodnocení jako matice rodokmenu (*pedigree*) – jednodušeji řečeno matice nejistot. Matice rodokmenu Refsgaarda et al. (2006) představuje kvalitativní či semikvantitativní hodnocení nejistoty koncepčního modelu. Kvůli jisté vágnosti se zde nehovoří o nejistotě, ale o obhajitelnosti (*tenability*). Matice je využita pro hodnocení nejistoty modelů v případových studiích (kap. 2.7). Uusitalo et al. (2015) uvádí přístupy pro hodnocení nejistoty v modelech pro podporu rozhodování. Strategie zvolená v předkládané práci v podstatě odpovídá kombinaci „emulace modelu“ a „více modelů“.

Jak ukazuje tabulka 1, dělíme nejistotu podle povahy (*nature*) na stochastickou (*aleatory*) a epistemologickou (*epistemic*). Epistemologická nejistota se může týkat toho, které jevy určují chování studovaného systému (omezené porozumění komplexnosti přírodních dějů). Její důsledky proto mohou být často závažnější než důsledky stochastické nejistoty. Podle Bevena a Younga (2013) je nejistota povahy spíše epistemologické než stochastické. Často jsou využívány statistické metody hodnocení nejistoty, které jsou vhodné pro odhad stochastické nejistoty, ale není snadné hodnotit těmito metodami epistemologickou nejistotu. Epistemologická nejistota se v modelu projevuje jako nejistota koncepčního modelu (*uncertainty in conceptual model*), a to konkrétně jako strukturální nejistota (nejistota ohledně struktury modelu – *structural uncertainty*). Neadekvátní struktura modelu je pro předpovědi modelu často mnohem důležitější než nepřesnost parametrů modelu.

1.3 Alternativní koncepční modely

Jako jedna z úrovní (*level*) nejistoty (tabulka 1) je uváděna úroveň scénářů (*scenario*) – nejistota nemůže být kvantifikována statisticky, ale může být vyjádřena pomocí alternativních popisů (koncepčních modelů). Tato úroveň nejistoty je běžná při řešení praktických hydrogeologických úloh, jako jsou mj. předkládané případové studie (kap. 2). Alternativní koncepční modely snižují dopady faktu, že tvorba koncepčního modelu je zatížena subjektivním vnímáním autora. Bredehoeft (2005) uvádí, že mnozí autoři navrhnou využití alternativních koncepčních modelů, které jsou testovány, či je z nich vybrán ten nejlepší. Uvádí však, že nikdy nezaznamenal tento přístup uskutečněný v praxi, protože modeláři pracují s jedním koncepčním modelem, který mění až v případě, kdy jej není možné nakalibrovat. Předkládaná práce naopak demonstruje, jak je možno uplatnit alternativní koncepční modely v praxi – totiž používat (alespoň v první fázi) více jednoduchých modelů. S pomocí informačního systému HgIS (kap. 3) či jiných vhodných nástrojů je tvorba koncepčních modelů snazší, a je tak možno jich prakticky vytvořit více.

1.3.1 Volba nejlepšího modelu

Alternativní modely je možno srovnávat a hodnotit. Jedním z kvantitativních přístupů je **průměrování modelů** (*model averaging*). Jednou z takových metod je GLUE – *Generalized Likelihood Uncertainty Estimation* (Beven). Uusitalo et al. (2015) uvádí obecně pojem „*ensemble modelling*“ a pro modely s odlišnou strukturou „*multiple model ensemble*“. Publikace, které by používaly pojem *ensemble modeling* pro modelování podzemní vody, nebyly nalezeny.

Již geolog Chamberlin (1890) vyzýval k tvorbě více pracovních hypotéz jakožto strategii pro rychlý pokrok v porozumění praktickým i teoretickým problémům. **Testování hypotéz** je nesmírně

důležité, protože představuje jádro metodologie vědy. Vědecká metoda spočívá ve formulaci hypotéz, které se pokoušíme vyvrátit – falzifikovat (Popper 1997). To, že pozorování odpovídají teorii, ještě nedokazuje pravdivost teorie. Stejně tak shoda modelu s pozorováním nepotvrzuje, že předpoklady či další předpovědi modelu jsou správné. Vyvrácení chybné hypotézy je však prokazatelný výsledek. Nadneseně lze říci, že vyvrácení koncepčního modelu, vyvrácení vědecké hypotézy a změna paradigmatu – vědecká revoluce (Kuhn 1970) – jsou podobné události na různých měřítcích vědy.

Výše uvedené metody nejsou však v hydrogeologické komunitě zcela samozřejmé. Článek zabývající se modelováním podzemních vod (Bredehoeft 2005) nepracuje s pojmy jako hypotéza a jejich vyvrácení. Místo toho změnu koncepčního modelu označuje jako překvapení (*surprise*). K němu dojde podle zkušeností autora článku u 20–30 % případů. Autor se nesnaží koncepční modely vyvrátit. K jejich nahrazení dojde až ve chvíli, kdy se mu nedaří stávající koncepční model ani podpořit.

Počáteční koncepční model obsahuje strukturu a počáteční hodnoty parametrů. Při kalibraci jsou upraveny parametry. Takový model může být dále validován konfrontací s nezávislými pozorováními – pokud možno jiné veličiny, než na kterou byl model kalibrován. Takový postup je možno chápat jako testování hypotézy, protože ve fázi kalibrace i validace se může ukázat, že ani model s optimalizovanými parametry není schopen adekvátně reprodukovat chování systému. Je tedy možno tvrdit, že testování hypotéz je součástí standardní modelářské praxe. Autor předkládané práce se domnívá, že je to pravda jen částečně. Pokud se nesnažíme úlohu zformulovat aktivně a přímo jako testování hypotéz (spolu s jasnými kritérii), budeme samovolně směřovat k přidávání dalších parametrů do modelu tak, aby byl model nakalibrován a obhájen – nikoliv vyvrácen. Validace je skutečně metoda aktivně navržená k testování hypotéz. Bývá však nedostatečné, pokud je jejím důsledkem další kalibrace s přidáváním parametrů bez přehodnocení struktury modelu. V praxi modely navíc často nebývají vůbec validovány – u případových studiích (kap. 2) nebyla dostupná data pro validaci. Není třeba doplňovat data pro vyvrácení hypotéz, které byly již vyvráceny jednoduchým modelem za nedostatku dat. Je tak možno se soustředit na získání dat potřebných pro vyvrácení zbylých hypotéz.

1.4 Jednoduchost a komplexnost modelů podzemní vody

V praxi je často účelné použít jednodušší (např. analytické) modely. Je u nich jasnější vztah mezi předpoklady a výsledky modelu a jsou také vhodné pro testování hypotéz. Dobře umožňují získat cit pro simulované procesy a jsou pochopitelné např. pro řešitele komplexního projektu, kterému jsou výsledky určeny. Komplexnost koncepčního modelu může vycházet z komplexnosti struktury modelu (nelineární) či parametrů (velký počet časově a prostorově proměnných parametrů) vyžadující numerické řešení. Jednoduchý model může být naopak lineární a celistvý (*lumped*) s analytickým řešením.

Přes vehementní chválu jednoduchých modelů autor předkládané práce samozřejmě netvrdí, že jednoduché modely jsou vždy lepší. Na alternativní jednoduché modely má v ideálním případě navázat komplexnější model.

1.5 Alternativní jednoduché modely

Použití většího množství jednodušších modelů umožňuje zadávat hodnoty vstupních parametrů tak, aby odpovídaly cíli konkrétního dílčího modelu. Vhodně využití matematické modely pomáhají navrhnout, která data je třeba doplnit, abychom model omezili daty a zmenšili nejednoznačnost. Jednodušší modely je snadnější testovat (*refutability*), posoudit adekvátnost jejich struktury (*transparency*) a je u nich také

snadnější vyhodnotit výsledky (Oreskes 2000). Pokud alternativní koncepční modely vznikají nezávislým zjednodušováním komplexní reality, je možné je využít pro hodnocení strukturní nejistoty modelu.

Jednoduchý model (např. kap. 2.5) je jednodušší a proto průkaznější, v mnoha případech by však takový model neumožnil vyvrátit hypotézu, protože model musí být na straně bezpečnosti kvůli nekorektní kvantifikaci procesů. Pokud je však hypotéza zamítnuta, je toto zamítnutí průkaznější, protože u jednoduchého modelu je nižší pravděpodobnost, že obsahuje skrytou vadu či jinou „zálužnost“. To se podobá statistickému usuzování, kdy při nastavení nízké hladiny významnosti (α) se často nepodaří hypotézu zamítnout. Přitom je vhodné využívat test s největší silou ($1-\beta$).

1.6 Princip ekvifinality

Formálnějším důvodem pro využití více modelů je princip ekvifinality (*equifinality thesis*). Princip ekvifinality říká, že v systému existuje řada rozdílných cest vedoucích k dosažení požadovaného stavu. V kontextu modelování pak více modelů (lišících se strukturou či jen parametry) simulují srovnatelně chování systému. Modelové řešení daného problému proto není jednoznačné. Princip demonstruje, že často není možné nalézt správný model. Průkopníkem vnímání principu ekvifinality v hydrologickém modelování je prof. Keith J. Beven.

V dostupné literatuře se často diskutuje účelnost jednoduchých modelů či princip ekvifinality, ale nikoliv obě zároveň. Někteří simulovali proudění podzemní vody a transport rozpuštěných látek v puklinovém prostředí za pomoci tří přístupů (stochastické kontinuum, síť diskretních puklin a síť 1D potrubí). V jiné práci byly alternativní koncepční modely realizovány sadami parametrů téhož modelu. Další používali alternativní koncepční modely (geologické, výpočet evapotranspirace atd.), které však byly implementovány v MODFLOW či v podobných programech jako třeba MIKE SHE. Gupta et al. (2012) se zaměřil na strukturální adekvátnost a možnost jejího hodnocení. Ta podle něj pro modely podzemní vody odpovídá 3D hydrostratigrafii, což je z pohledu autora předkládané práce pojetí zúžené na komplexnější 2D/3D modely.

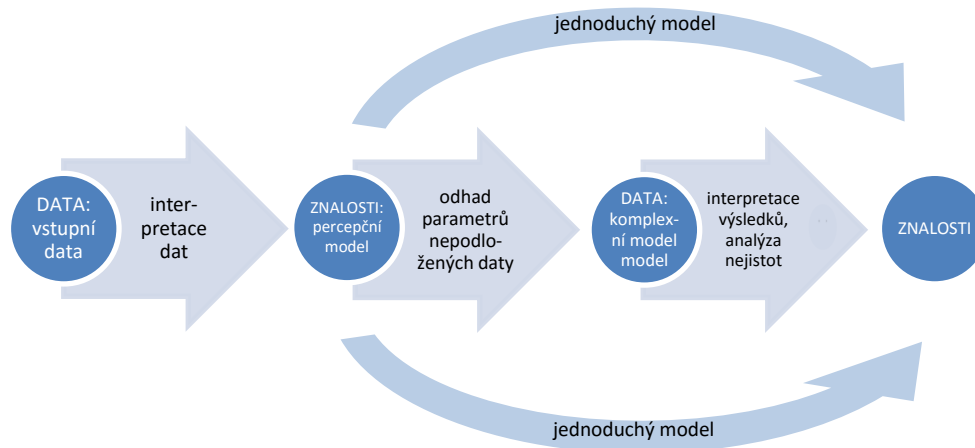
1.7 Formulace zvoleného přístupu

Studie zmiňující princip ekvifinality v podstatě vždy využívají sadu parametrů pro jeden model. To je praktické v případě hydrologických modelů, kde jsou dostupné dlouhé časové řady významných veličin (průtok ve vodním toku je soustředěný – integrovaný – odtok z celého dílčího povodí) a mnohé parametry jsou známy se značnou přesností (např. terén). V případě podzemní vody může být epistemologická nejistota větší. Neznáme směr proudění a preferenční cesty; někdy neznáme všechna ohniska kontaminace, a zda se kontaminace šíří horninovým prostředím či například umělým drénem apod. U povrchových vod alespoň vždy víme, že vodní tok existuje a kde se nachází. V případě řešení praktických hydrogeologických úloh často nemáme dostupná dostatečná data.

Proto jde předkládaná práce v aplikaci principu ekvifinality ještě dál a nevytváří alternativní sady parametrů (varianty, scénáře) ani alternativní geologické modely, ale jednoduché alternativní modely, které simulují jiné procesy, používají jinou metodu či představují odlišné koncepční přístupy (jsou principiálně odlišné). Případové studie (kap. 2) často vychází z nedostatečných vstupních dat a není účelné a často ani možné modely kvantitativně srovnávat. Modely tak nejsou kalibrované ani validované.

V některých případových studiích se neuplatňuje princip ekvifinality, protože modely se doplňují (např. kap. 2.3).

Schéma na obrázku 1 představuje dvojnásobnou transformaci dat ve znalosti. Jednoduchý model (šipky nahoře a dole) umožňuje tuto transformaci zjednodušit mj. tím, že nezavádí do koncepčního modelu nepodložené předpoklady (neznámé parametry), jejichž vliv na výsledek by bylo třeba vyhodnotit. Je srozumitelné, z jakých předpokladů vycházejí. Strategii je možno nadneseně vyjádřit takto: „Je lepší vytvářet jednoduché modely nad měřenými daty než komplexní modely nad vymyšlenými daty.“



Obrázek 1: Data a znalosti – jednoduchý model má o jednu transformaci méně

Alternativní jednoduché modely mohou být účelně využívány různým způsobem. Jednoduchost tvorby modelů dává větší prostor pro využití modelu rozmanitým způsobem, aby tak byl vhodným nástrojem pro podporu rozhodování, a to i v případě nedostatku dat. Navrhuji klasifikaci využití modelů se zaměřením na jednoduché modely sestavené při nedostatku dat. Je to tedy souhrn otázek, které může zadavatel klást (zpracovateli) modelu:

- Interpretací model – Porozumění probíhajícím procesům / tvorba koncepčního modelu – pomocí modelu, který odpovídá měřením prokázat, že naše porozumění procesům odpovídá měřeným výsledkům a případně, že data jsou vnitřně konzistentní.
- Testování hypotéz – Příklad: Ani nejdůležitější fyzikálně možná kombinace parametrů modelu není schopna generovat výsledek, který odpovídá pozorování – proto je struktura modelu chybná, a tedy procesy, na kterých je model založen, nejsou dostatečné pro simulování skutečného chování systému. Příklad využití tohoto přístupu je v kap. 2.1 a 2.5.
- „Proof of concept“ – Měření či alespoň fyzikálně možné hodnoty modelu jsou schopny popsat vzorec (*pattern*) chování systému. Struktura modelu tak může být správná a simulované procesy mohou být schopny vysvětlit chování systému.
- Hodnocení vlivu očekávaných změn:
 - Analýza citlivosti – ukazuje, které parametry výrazně ovlivňují výsledky modelu.
 - Scénáře – typové varianty budoucího vývoje.
- Předpovědi neznámé budoucnosti:
 - *Prediction* – předpokládáme konkrétní časový vývoj parametrů modelu. Příkladem je model v kap. 2.2. Při nejistotě vstupních dat je možné použít variantní výpočty (scénáře): např. optimistický, pesimistický či realistický.

- *Forecasting* – model počítá vývoj celého systému.
- „Inverzní model“ – zjišťování parametrů modelu, které odpovídají meznímu požadovanému (nežádoucímu) stavu (např. nárůst látkového toku kontaminace v recipientu nad stanovenou hodnotu; výtopa chráněného území apod.). Výsledkem je parametr (například hydraulická vodivost), který je třeba ověřit podrobným průzkumem. Přístup je využit v kap. 2.4.
- Analýza nejistoty – Jak spolehlivé jsou výsledky modelu.
- Analýza citlivosti – Jaká data jsou pro hodnověrnost modelu nejdůležitější?
- Neparаметrický výpočet – Např. porovnání doby do běhu radionuklidů do recipientu pro různé lokality jaderného úložiště při stejné hodnotě hydraulické vodivosti pro všechny lokality. Je tak hodnocen vliv geometrie problému, nikoliv jeho parametrů. Jedná se o relativní srovnání lokalit. Takováto klasifikace strategií modelování či formulování zadání, na které může model odpovědět, nebyla v literatuře nalezena. Uvedená klasifikace demonstruje, že výsledkem modelů nemusejí být předpovědi týkající se budoucnosti, byť by byl zprvu zadavatelem požadavek právě tak formulován.

Alternativní modely jsou v praxi používány, ačkoli se v tomto kontextu nemluví o ekvifinalitě. Toto téma je však natolik závažné, že stojí za hlubší rozbor, a je tak přínosem k debatě ve vědeckých časopisech, jež v současnosti intenzivně probíhá. Níže představené případové studie originálním způsobem demonstrují výše uvedené principy. Nepředstavují schematický přístup, ke kterému vede řešení problematiky za pomoci existujících softwarových balíčků pro modelování podzemní vody, ale nespoutané přemýšlení, které je ve shodě se současnou filosofií vědy.

2 Případové studie

Využití alternativních jednoduchých modelů je demonstrováno na případových studiích. Případové studie jsou v terénním měřítku a jsou sadou jednoduchých modelů, či srovnáním jednoduchého modelu s modelem komplexním. Případové studie byly zpracovány většinou s malým množstvím dat, a umožňují tak testovat „hloubku naší neznalosti“ – epistemologickou nejistotu. Vzhledem k tomu, že modely jsou často prediktivní a není nedostatek dat pro srovnání s výsledky modelu, jednotlivé modely nejsou mezi sebou srovnávány matematickými a statistickými metodami, ale nejistotu vyjadřují víceméně rozmezím hodnot. Modely byly většinou sestaveny pro analýzu rizika v rámci komerčních zakázek.

Předkládané případové studie řeší běžné hydrogeologické úlohy s běžným (tj. malým) množstvím dat za běžných časových podmínek pro zpracování modelu. Jedinečný je však tvůrčí přístup k řešení, který je zasazen do teoretického rámce (kap. 1). Všechny případové studie byly zpracovány samostatně autorem předkládané práce. V textu této práce není vždy uváděno umístění lokalit ani konkrétní číselné výsledky.

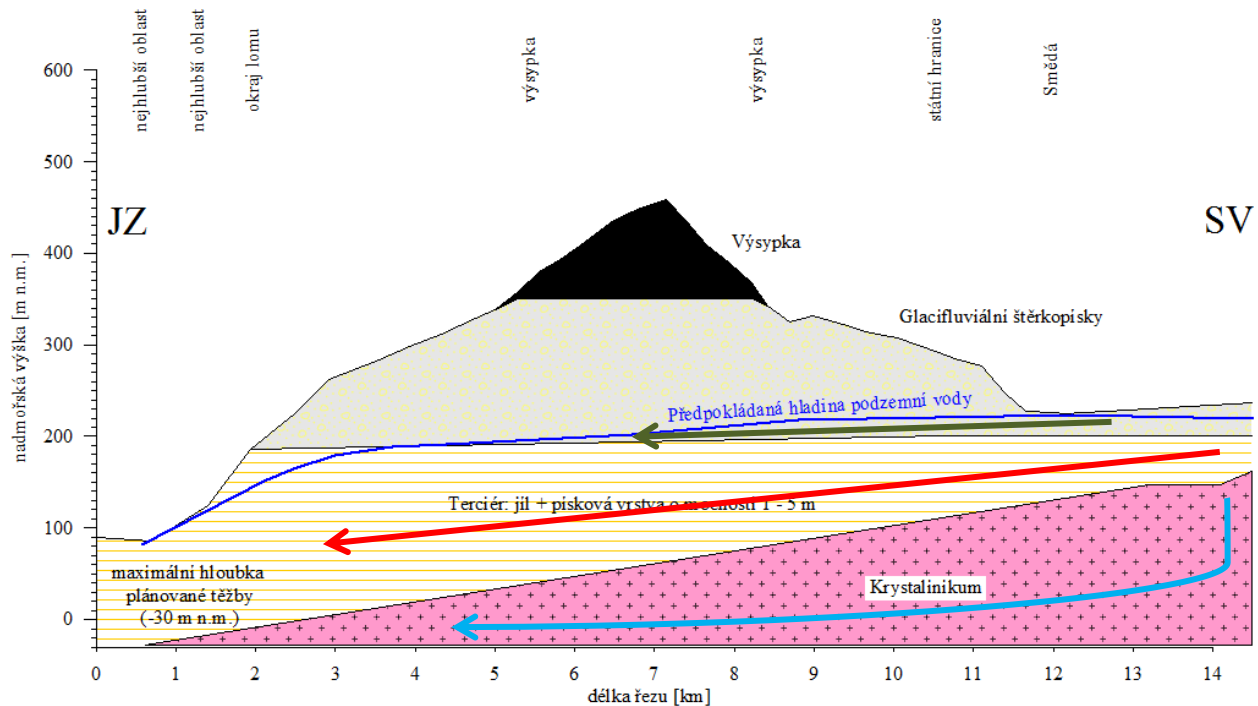
2.1 Ekvifinalita – interakce povrchové a podzemní vody

V této kapitole je prezentována pravděpodobně nejnázornější aplikace principu ekvifinality a testování hypotéz ze všech případových studií (Skořepa et al. 2009). Nejedná se totiž o predikce, ale o vysvětlení pozorovaného stavu.

V toku Smědá ve Frýdlantském výběžku dochází ke ztrátám průtoku (m^3/s). Za léta 1995–2008 byla ve 24 měsících (40 % ztrátových měsíců) ztráta statisticky významná (větší než 10 %). Průměrná ztráta (pokud se projeví) je cca 240 l/s. Ztráta se projevuje zejména při nízkých průtocích. Ke ztrátě

může docházet i v měsících, kdy po směru toku průtok narůstá. Hodnocení totiž nepracuje s ostatními přítoky Smědé (drobné vodoteče, povrchový a podpovrchový odtok).

Je možno zformulovat několik hypotéz (obrázek 2), kam se voda ze Smědé ztrácí. Některé hypotézy jsou dále testovány matematickým modelem. Všechny modely byly představovány jednoduchými vzorci, které jsou založeny na Darcyho zákoně a rovnici kontinuity. Modely byly vypočítány v MS Excel.



Obrázek 2: Ideový řez zájmovým územím – koncepční model

Tři hypotézy uvažují ovlivnění činností dolu Turów a jedna uvažuje pouze přirozené příčiny. Přirozenou možností, kam se může voda ztrácet, je *subglaciální koryto*. Tato hypotéza není schopna vysvětlit pokles hladin podzemní vody, částečně však vysvětluje ztráty průtoku ve Smědě. Při přirozeném gradientu Smědé proudí podzemní voda nivou paralelně s tokem. Subglaciální koryto má možná podstatně větší transmisivitu než holocénní štěrkopisky. Je totiž cca 50 m hluboké, zatímco holocénní štěrkopisky v údolní terase jsou pouze cca 5 m mocné. Ztráty povrchové vody tak dotují poříční vodu v údolní nivě. Pokud by korytem měla proudit celá ztráta ze Smědé (240 l/s), bylo by třeba, aby hydraulická vodivost výplně glacifluviálního koryta byla $3 \cdot 10^{-3}$ m/s. To je hodnota velmi vysoká, proto ne příliš pravděpodobná, avšak rozhodně nikoli nemožná. *Hypotéza ztráty vodnosti Smědé pouze v důsledku přetoku do subglaciálního koryta tedy nebyla vyvrácena.*

Na J a Z od Filipovky se nachází podle hydrogeologické mapy akumulace *glacifluviálních sedimentů*. Vzhledem k jejich možné propustnosti lze uvažovat, že přes ně dochází k proudění podzemní vody do dolu Turów. Pro tuto hypotézu však nemáme žádný přímý důkaz. Odtok vody z toku do dolu byl vypočten jako součin gradientu, šířky průtočného profilu a transmisivity. Z výpočtu vyplývá, že hydraulický spád mezi Smědou a dolem Turów je dostatečný, aby za předpokladu neověřených hydrogeologických podmínek způsobil prokázané ztráty průtoku ve Smědě. *Hypotéza ztráty vodnosti pouze v důsledku přetoku přes glacifluviální sedimenty tedy nebyla vyvrácena.*

K přetoku podzemní vody do dolu Turów může také docházet přes *písčité polohy v terciérních sedimentech*. Byl pozorován pokles hladiny ve studni, ale byl zachován směr proudění (hydraulický gradient) z Polska do ČR. Výpočet kvantifikuje snížení přítoku z Polska do ČR v důsledku poklesu

hladiny podzemní vody. Snížení přítoku vody z Polska byl vypočten jako součin rozdílu současného a původního gradientu, šířky průtočného profilu (\check{S}) a transmisivity (T): $\text{rozdíl_gradientu_hladin} \times \check{S} \times T$. V důsledku poklesu hladiny ve vrtu o cca 40 cm došlo ke snížení přítoku podzemní vody do ČR ve Višňové nejvýše o 10 l/s. Takové množství se neprojeví výrazně na přítoku Smědé. *Hypotéza ztráty vodnosti pouze v důsledku přetoku přes písčité vločky v terciérních sedimentech tedy byla vyvrácena.*

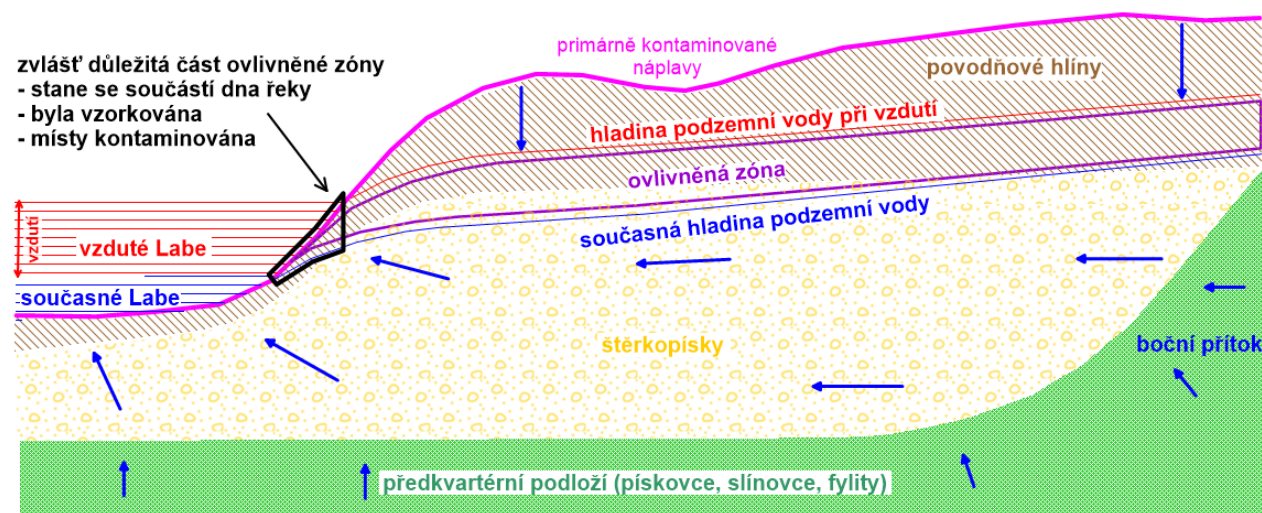
K přetoku může taktéž docházet přes tektonické poruchy. Ztráty vody vlivem poruch by vysvětlovaly, proč k poklesu hladiny podzemní vody dochází na relativně rozsáhlém území a pokles není větší. Tato hypotéza nebyla pro neexistenci relevantních dat vůbec testována matematickým modelem. *Hypotéza přetoku přes diskontinuity tedy nebyla testována a tedy ani vyvrácena.*

Problematika ovlivnění Smědé dolem Turów je velmi obtížná a dále je ztížena nedostatkem relevantních dat. Ve skutečnosti se na ztrátách průtoku Smědé pravděpodobně podílejí všechna uvedená vysvětlení. Matematické modely umožnily maximálně vytěžit existující data a testovat hypotézy o příčině ztrát vodnosti Smědé. Jedna hypotéza byla vyvrácena. U jedné hypotézy byla vypočtena mezní hodnota hydraulické vodivosti pro vyvrácení hypotézy.

2.2 Srovnání jednoduchého a komplexního modelu – ovlivnění podzemních vod jezem

Kvantitativní hodnocení transportu kontaminantů z náplavů Labe při vybudování jezu u Děčína bylo prováděno ve dvou etapách pomocí analytického modelu (Nešetřil 2008c) a 2D vertikálního numerického modelu (Nešetřil 2009a), na který navázal detailní model proudění dílčí lokality (Matula et al. 2009, 2014). Jedná se tedy o srovnání jednoduchého a komplexního modelu.

Zájmové území je tvořeno fluviálními sedimenty Labe a podložními, zejm. křídovými horninami (obrázek 3), ze kterých přitéká do kvartéru značné množství vody – cca 45 l/s/km². Cílem prací bylo vymezit možné transportní cesty kontaminantů při napouštění jezu i při jeho provozu a kvantifikovat potenciální ovlivnění kvality podzemní vody. Studie se zabývá šířením kontaminantů obsažených v náplavech Labe, které budou zatopeny v důsledku vzduť jezu.



Obrázek 3: Schematický hydrogeologický řez a koncepční model

Byl sestaven matematický model (obrázek 4) transportu kontaminantů při napouštění jezu (kontaminanty pronikají do štěrkopískového kolektoru) a při návratu kontaminantů do Labe.

Matematický model je odvozen z rovnice kontinuity. Nepočítá s Darcyho zákonem (neuvažuje hydraulickou vodivost). Uvažujeme tak nepříznivou variantu. Hydraulická vodivost štěrkopískového kolektoru je tak vysoká (až $9 \cdot 10^{-3}$ m/s), že při zatápní jezů (min. 1 měsíc) je dost času na ustálení proudění.

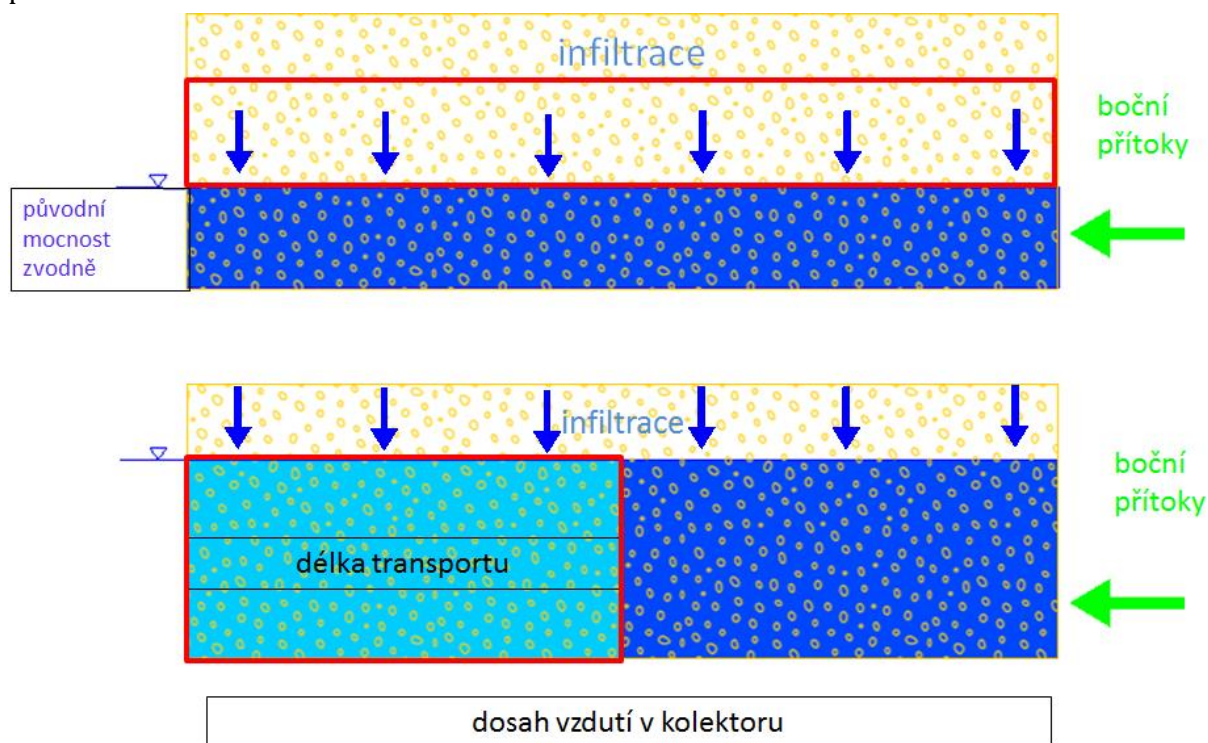
Modelové řešení průniku kontaminantů do fluvialních sedimentů při zatápní jezů je odvozeno z úvahy, že objem kontaminované vody, která vnikne do kolektoru, je roven objemu suchých pórů, které jsou při zatápní zaplaveny. Takže platí:

$$\text{délka_transportu} = \text{dosah_vzdutí_v_kolektoru} \times \text{vzdutí} / (\text{původní_mocnost_zvodně} + \text{vzdutí})$$

Dosah vzdutí v kolektoru odpovídá šířce štěrkopískového kolektoru (nivy). V některých úsecích je šířka nivy značná (až cca 450 m). Protože dochází k přítokům z křídly na bocích štěrkopískového kolektoru, byl maximální dosah stanoven na základě archivního hydrogeologického průzkumu na 450 m. Jedná se o vzdálenost od řeky, ve které se současná hladina podzemní vody rovná jezem vzduté hladině v Labi.

Modelové řešení následného vytlačení kontaminantů v důsledku přirozených přítoků do fluvialních sedimentů počítá s tím, že těleso kontaminované vody je zatlačováno do Labe přítokem z boků (svahů), infiltrací a přítokem z křídly do části kolektoru, který není kontaminován. Samostatně byla vypočítána doba návratu v důsledku infiltrace a v důsledku bočních přítoků. Vzorec pro dobu návratu z důvodu infiltrace srážkových vod byl odvozen pomocí integrace, protože délka kolektoru nezasaženého kontaminací se v průběhu výpočtu mění. Celková doba návratu byla vypočtena podobně jako odpor dvou paralelně zapojených rezistorů.

Zpomalení šíření kontaminantů ve štěrkopískovém kolektoru vlivem sorpce (retardace) byla zanedbána (nepříznivá varianta). Model předpokládá, že dno řeky je kontaminováno stejně jako vzorkované náplavy (nepříznivá varianta). Není uvažováno míšení vod v kolektoru. Modelové řešení bylo provedeno v tabulkovém kalkulátoru MS Excel.



Obrázek 4: Schéma bilančního výpočtu: Objem pístu vody z řeky odpovídá zvětšení zvodně

Výsledky modelu ukazují, že v extrémním případě by se kontaminace dostala cca 60 m od břehu, byla by však do cca devadesáti dní od začátku zvyšování hladiny v daném místě zatlačena zpět. Skutečnost bude však příznivější, protože model neuvažuje kolmataci dna, hydraulickou vodivost štěrkopísků, sorpci kontaminantů v štěrkopískovém kolektoru, kinetická omezení desorpce z náplavů ani hydrodynamickou disperzi a míšení vod v kolektoru. Studie prokázala, že vybudování jezu nebude mít dlouhodobý vliv na kvalitu podzemní ani povrchové vody.

Tato případová studie je jediná, která se zabývá rozsáhlým územím, které je z hydraulického hlediska velmi dobře prozkoumané. Nicméně znalosti o kontaminaci náplavů a jejich mobilitě nejsou dostatečné pro to, aby bylo opodstatněné využití komplexnějšího modelu. Již jednoduchý model prokázal, že vybudování jezu nebude mít dlouhodobý vliv na kvalitu podzemní ani povrchové vody a tak by komplexnější model nebyl lepším podkladem pro rozhodování.

Na základě požadavku oponenta bylo v další etapě (Nešetřil 2009a) navázáno na tento jednoduchý model sadou 2D vertikálních modelů provedených v MODFLOW a MT3D-MS. Byla tak vytvořena predikce relativních koncentrací ve 2D řezech, která však nepřispěla k lepšímu pochopení procesů. Výsledky nebyly podpořeny adekvátními daty. V důsledku vysoké hydraulické vodivosti nemá její konkrétní hodnota v modelu praktický význam, protože k průniku povrchové vody dochází téměř okamžitě. Do numerického modelu bylo třeba zadat storativitu a efektivní porozitu, které nejsou známy. V analytickém modelu však stačí zadat jejich poměr. Protože zjednodušeně jsou si tyto dvě veličiny pro zvedeň s volnou hladinou rovny, došlo k jejich vykrácení, a do analytického modelu se vůbec nezadávají.

Případová studie uplatňuje první bod návodu – Guidelines for effective modelling (Hill a Tiedeman 2007, s. 268) – Start simple and add complexity...

V následujících případových studiích se nemusí nutně uplatňovat ekvifinalita, ale vhodně volená kombinace jednoduchých modelů (přístupů) pro dosažení zvoleného cíle.

2.3 Kombinace principiálně odlišných modelů – model délky kontaminačního mraku a dovoleného čerpání

Modelová studie (Nešetřil 2008a) byla zpracována pro analýzu rizik kontaminace ropnými uhlovodíky (RU). Bylo vypočítáno a hodnoceno:

- maximální rozsah oblasti zastižené v budoucnu kontaminací (na základě délky ustáleného kontaminačního mraku RU vypočítané pomocí modelu CoronaScreen) a
- dovolené čerpání, které nebude přitahovat kontaminaci z potenciálně kontaminované oblasti.

Šíření kontaminace je myslitelné v turonském kolektoru (dále jen „kolektor“) a to zejména v jeho svrchní části, která je přípovrchově rozpukaná, a proto propustnější.

2.3.1 Model ustálené délky kontaminačního mraku

Obecně lze říci, že se růst kontaminačního mraku vlivem biodegradace zpomalí. Po jisté době se již délka kontaminačního mraku nezvětšuje. Dojde totiž k dostatečnému promíšení vod a všechny transportované RU jsou postupně spotřebovávány biodegradacími reakcemi. Látkový tok RU z ohniska se tak rovná množství RU spotřebovaných biodegradacími reakcemi. Toto stádium se nazývá ustálený kontaminační mrak (*steady state plume*). Délka ustáleného kontaminačního mraku byla vypočítána pomocí programu CoronaScreen. Je to soubor tří matematických modelů vyvinutých v rámci evropského projektu

CORONA, na kterém se autor předkládané práce podílel. Výsledkem každého modelu je odhad délky ustáleného kontaminačního mraku.

CoronaScreen obsahuje následující modely:

1. Model bilance elektronů
2. Analytický model
3. Pohyblivý 1D numerický model – využívá PHREEQC (Parkhurst a Appelo 1999)

CoronaScreen používá předpoklad, že ohnisko kontaminace má konstantní koncentraci a konstantní mocnost a že vstupní parametry jsou distribuovány uniformně. Rychlost biodegradace je do značné míry určena mírou promíšení kontaminované vody a čisté vody obsahující elektronové akceptory v reaktivní zóně na okraji mraku. Aby bylo možno modelovat šíření RU včetně biodegradace pro celou lokalitu, musí být mřížka transportního modelu dostatečně jemná, aby nedocházelo k nadhodnocení míšení a tím nadhodnocení biodegradace. CoronaScreen tento nedostatek nemá. Velkou výhodou CoronaScreen ve srovnání s univerzálními 3D simulátory (např. RT3D) je, že CoronaScreen dobře pracuje s disperzí (dostatečná diskretizace, pokud je vůbec potřeba), používá více metod a byl vyvinut právě pro přirozenou atenuaci.

Vstupními daty modelů CoronaScreen jsou: hydrogeologické charakteristiky (rychlost proudění, pórovitost, koeficient retardace), transportní charakteristiky (disperzivity či mocnost reaktivní zóny), šířka a mocnost kontaminačního mraku, koncentrace kontaminantů a elektronových akceptorů v pozadí a v ohnisku.

2.3.2 Model dovoleného čerpání

Na základě modelu bylo stanoveno, jaké množství podzemní vody je možno čerpat v obci. Čerpání nemá ovlivnit přirozený směr proudění podzemní vody v oblasti maximálního možného rozšíření kontaminace. Kontaminace se tak nebude šířit do nekontaminované oblasti na sever a směrem k využívaným studnám v obci. Dovolené čerpané množství Q bylo stanoveno tak, aby hydraulická deprese vyvolaná tímto čerpáním nedosáhla k oblasti maximálního možného rozšíření kontaminace.

Hydraulickou depresí pro tento účel rozumíme oblast, pro kterou platí současně, že se zde:

- projevuje snížení hladiny podzemní vody vyvolané čerpáním Q z využívané studny,
- veškerá podzemní voda (z části kolektoru vertikálně zastiženého využívanou studnou) je vyčerpána využívanou studnou.

Dovolené čerpané množství Q je dáno součtem dílčích přítoků Q_1 a Q_2 .

Přítok Q_1 je přítok z oblasti mimo hydraulickou depresí, která byla vyvolána čerpáním Q .

$Q_1 = T \times i \times 2L$, kde i je hydraulický gradient, T je transmisivita a L je vzdálenost k oblasti maximálního možného rozšíření kontaminace tj. polygonu na mapě (obrázek 5), který je vytvořen na základě možného směru proudění podzemní vody a délky ustáleného kontaminačního mraku.

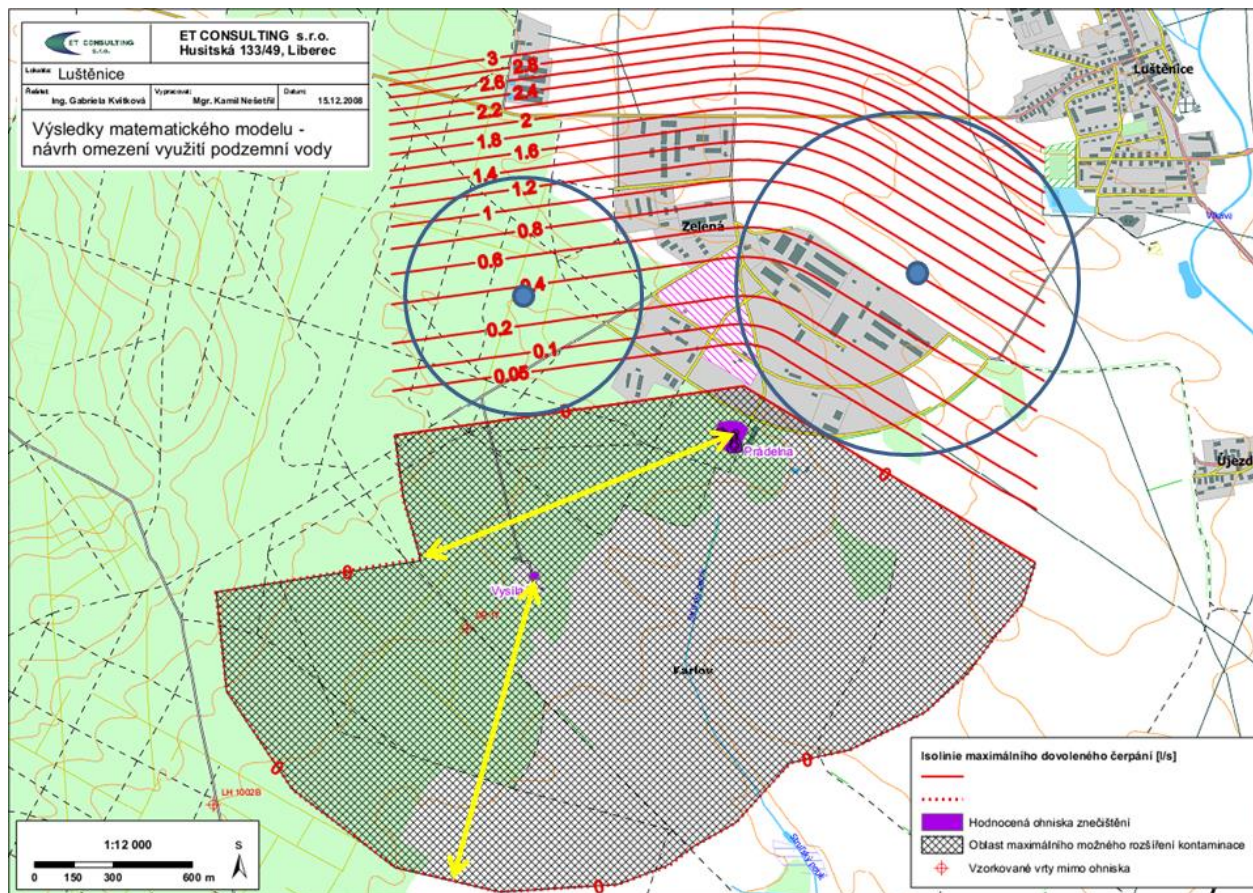
Přítok Q_2 je dán infiltrací srážkových vod do oblasti hydraulické deprese.

$Q_2 = \pi \times L^2 \times q$, kde q je infiltrace

a tedy $Q = Q_1 + Q_2$.

Za zadaných podmínek je $Q_1 \ll Q_2$.

Výsledky modelů jsou zobrazeny níže (obrázek 5).



Obrázek 5: Výsledky modelu ustálené délky kontaminačního mraku (oblast max. možného rozšíření kontaminace) a modelu čerpání (isolinie). Žluté šipky představují maximální délku kontaminačního mraku. Kruhy mají poloměr L.

Pro specifický účel byla využita unikátní kombinace modelů, která dala praktický návrh na využívání podzemních vod v okolí kontaminovaného území. V případě modelu délky ustáleného kontaminačního mraku se uplatnil princip ekvifinality (tři metody výpočtu tétož). Ve standardním numerickém transportním modelu by docházelo k numerické disperzi, která by mohla vést k rychlejší biodegradaci kontaminantu – na druhou stranu by bylo možno zpomalovat odbourávání zahrnutím kinetiky. Model čerpání by bylo nutno realizovat jen pro explicitně zadané plánované studny.

Aktualizované analýze rizik (v rámci které vznikl prezentovaný model) předcházela analýza rizik. V rámci ní byl zpracován numerický model šíření kontaminace (MODFLOW, MT3D). Jeho využití pro předkládanou studii by nebylo účelné.

2.4 Transport ropných uhlovodíků

Případová studie demonstruje využití většího množství relativně jednoduchých (zejm. analytických) modelů pro hodnocení přirozené atenuace ropných uhlovodíků. Jsou využity modely založené na Darcyho zákoně, vodní bilanci, látkových tocích i aplikaci simulačního softwaru. Model je podrobně dokumentován ve zprávě (Nešetřil 2008b), jež je přílohou analýzy rizika (Kubricht 2008). Výsledky byly publikovány v konferenčním sborníku (Nešetřil 2009b).

Pro potřeby analýzy rizik (stanovení sanačních limitů) bylo hodnoceno šíření RU a jejich přirozená atenuace. Ohnisko kontaminace RU se nachází v distribučním skladu pohonných hmot (DS). Podzemní voda proudí směrem k vodárenskému jímacímu území (JÚ). Úbytek elektronových akceptorů, nárůst

produktů biodegradace, rozborů pevné fáze a mikrobiologické rozborů prokázaly, že přirozená biodegradace probíhá. Matematický model se pak pokusil tyto procesy kvantifikovat. Dílčí modely shrnuje tabulka 2.

Tabulka 2: Shrnutí případové studie „Transport ropných uhlovodíků“

Výsledek	Způsob výpočtu
Doba transportu	Darcyho zákon
	Bilanční výpočet (konzervativní stopovač se dostane do JÚ za dobu, kdy je vyčerpán objem vody mezi zdrojem a jímacím územím)
Látkové toky kontaminantů	Ze zdroje vs. vyčerpáno vodárenskými vrty
	Zpětná úloha (jaká je bezpečná koncentrace kontaminace ve zdroji tak, aby v jímacím území byly dosaženy limity)
Bilance biodegradace	Pokles elektronových akceptorů při průtoku ohniskem
Jaká může být koncentrace v ohnisku, aby kontaminační mrak nedosáhl do jímacího území?	Model bilance elektronů (CoronaScreen)
	Analytický model (CoronaScreen)
	Pohyblivý 1D numerický model (CoronaScreen)

Byly provedeny modelové výpočty od jednoduchých analytických modelů až k použití screeningového modelu CoronaScreen. Modely vycházely z různých předpokladů. Pro určení oxidační kapacity pevné fáze byla využita metodika Herona et al. (1994a, 1994b), která byla využita taktéž v článku (Topinkova et al. 2007), jehož spoluautorem je autor předkládané práce.

2.5 Transport chlorovaných uhlovodíků – advekční a bilanční model

Na lokalitě byla řešena (Nešetřil 2012b) problematika transportu chlorovaných uhlovodíků do studní pro individuální zásobení podzemní vodou a do vodárenských vrtů. Ohnisko kontaminace je v málo propustném kvartérním kolektoru, ale k čerpání ve vodárenských vrtech dochází v propustném cenomanském kolektoru. V důsledku čerpání je hladina v cenomanu níže.

2.5.1 Advekční model – predikce

První způsob spočíval ve výpočtu doby, za kterou by se mohla kontaminace dostat kvartérním kolektorem ke studnám individuálního zásobování a cenomanským kolektorem k jímacím vrtům. Advekční model je založen na Darcyho zákonu a na rovnici kontinuity. Proudění v cenomanském kolektoru je určeno především čerpáním ve vodárenských vrtech (radiální proudění) a je proto nutno uvažovat 2D proudění. Model je založen na metodě analytických elementů (AEM) – simulační kód Bluebird / Cardinal. Byly vypočítány hydroizopiezy cenomanské zvodně a *particle tracking*. Výsledkem obou advekčních modelů je odhad doby doběhu kontaminace. Nejkratší doběhové doby vycházejí u cenomanského kolektoru. Koncentracemi CIU v jímacích vrtech v důsledku šíření v cenomanské zvodni se proto zabývá navazující bilanční model.

2.5.2 Bilanční model – testování hypotézy

Druhý přístup kvantifikuje přetok podzemí vody a kontaminace mezi kolektory. Bylo vypočítáno, jaká koncentrace CIU v jímacím vrtu odpovídá tomuto toku CIU.

Byl vypočítán hmotnostní tok CIU do cenomanské zvodně na základě rozdílů hladin v jednotlivých zvodních, mocnosti a propustnosti izolátorů a plochy kontaminované zvodně o dané koncentraci CIU. Byly vypočítány alternativy toku přímo z kvartérní zvodně a z turonské zvodně. Bylo uvažováno, že veškeré CIU, které přetečou do cenomanské zvodně, budou rovnoměrně naředěny ve vodě čerpané z vodárenského vrtu. Tak byla vypočítána koncentrace CIU ve vodárenském vrtu L2 – 0,032 µg/l odpovídá látkovým tokům CIU z kvartérního kolektoru a 0,025 µg/l tokům z turonu. Tyto hodnoty jsou neočekávaně nízké. Jsou výrazně menší než měřené koncentrace ve vodárenském vrtu (0,83 µg/l TCE a 0,55 µg/l PCE). Když se obdobný výpočet provede pro koncentrace CIU v turonské zvodni před začátkem sanace, vyjdou koncentrace CIU ve vodárenském vrtu přibližně 12krát větší (0,30 µg/l), což je však stále výrazně méně než měřené koncentrace CIU. To naznačuje, že cenomanský kolektor byl pravděpodobně kontaminován CIU v důsledku gravitačního proudění volnou fází CIU (DNAPL). Zatímco rozpuštěné CIU migrující puklinou v izolátoru difundují do bloku horniny, tak DNAPL – podobně jako koloidy (např. nanočástice) – difúzi nepodléhají, a mohou migrovat mnohem snáze než konzervativní stopovač (Cherry 2007). Model uvažuje šíření rozpuštěných CIU advekcí, nikoliv fáze DNAPL CIU. Fáze se může svisle šířit i proti směru proudění podzemní vody. Šíření fáze je obtížně kvantifikovatelné a nejsou k němu pro tuto lokalitu dostupná data.

2.5.3 Diskuse a závěr

Vypočítané intenzity přetoku podzemní vody mezi kolektory jsou nerealisticky veliké. Ve výpočtu se totiž nejedná o bilancování vody pro celou lokalitu, ale o výpočet, který prokázal, že ani tak absurdně veliké toky nejsou schopny generovat měřené koncentrace CIU ve vodárenském vrtu. Úvaha ukazuje na to, že kontaminace se vertikálně šířila především jako DNAPL. Jedná se myšlenkový experiment, kterým testujeme hypotézu, zda jsou měřené koncentrace rozpuštěných CIU schopny ovlivnit kvalitu vody ve vodárenských vrtech. Jedná se o nejpesimističtější myslitelný scénář.

Bylo by možno sestavit distribuovaný numerický model. Aby byl takový model smysluplný, bylo by třeba shromáždit a interpretovat data z širšího okolí ohniska kontaminace a jímacích vrtů. Celistvý analytický model umožňuje přímou interakci s uživatelem (ne-modelářem) a snadné porozumění implementaci modelovaných procesů. Je transparentní a snadno reprodukovatelný. Není přeparametrizovaný, což znamená, že se do modelu nezadávají parametry, jež nejsou pro výpočet skutečně významné. Výhodou takového jednoduchého modelu je mj. možnost snadno zpracovávat variantní výpočty.

Model vyvrátil implicitní hypotézu, že je rozhodující transport rozpuštěných CIU. Složitější model by pouze zamlžil vztah mezi předpoklady a důsledky. Vyskytovaly by se v něm totiž další veličiny a procesy (např. čas, disperzivita).

2.6 Přirozená biodegradace ropných uhlovodíků

V rámci práce na článku Topinkové et al. (2007) autor této práce provedl geochemické výpočty (speciace, saturační indexy a inverzní model) v programu PHREEQC. Jednalo se o kombinaci jednoduchých modelů, které pomohly identifikovat probíhající procesy.

2.7 Hodnocení nejistoty koncepčního modelu podle Refsgaarda et al. 2006

Nejistota koncepčního modelu předložených případových studií byla posouzena podle metodiky (Refsgaard et al. 2006). Všechny případové studie je možno charakterizovat následně (tabulka 3):

Tabulka 3: Matice rodokmenu (pedigree) podle Refsgaarda et al. (2006) pro prezentované případové studie

Skóre	Podložení empirickými důkazy		Teoretické porozumění	Reprezentace porozumění procesům	Věrohodnost	Shoda názorů odborníků
	Reprezentativnost	Kvalita a kvantita				
15	Bez přímé vazby (0)	Archivní/terénní data, neřízené experimenty, malé vzorky, přímá měření (3)	Prověřená teorie (4)	Jednoduchý model s agregovanými parametry (2)	Vysoce věrohodný (4)	Konkurenční školy (2)

Modely vycházejí z archivních dat, která jsou většinou doplněna terénními měřeními. Nejsou dostupné dlouhodobé časové řady. Nejsou dostupná data pro kalibraci modelu ani jeho verifikaci. Model je postaven na prověřených teoriích, jež jsou v modelu implementovány v zjednodušené formě. Jednoduchost modelu umožňuje jeho pochopení a posouzení hydrogeology ne-modeláři. Nemusí však existovat shoda mezi odborníky, zda dostupná data jsou dostatečná a zda zjednodušení modelu není příliš velké.

3 Informační systém HgIS

Pro tvorbu alternativních jednoduchých modelů nemá smysl zpracovávat formální metodiku. Aby však bylo možno snadno vyhodnocovat data, vytvářet analýzy, formulovat koncepční modely a vytvářet sady různorodých alternativních matematických modelů, je vhodné využívat informační systém (IS), který to technicky snadno umožňuje. Je tak možno lépe se vyrovnat se strukturální nejistotou koncepčního modelu. V předkládané práci je představen informační systém HgIS, který tyto požadavky splňuje, a je tak schopen řešit většinu úloh, jež byly představeny v kap. 2. HgIS byl vyvinut pod vedením autora předkládané práce.

Pro správu dat, jejich vizualizaci, vyhodnocení a tvorbu modelů je možno používat existující software. Na specifika mj. hydrogeologických dat se specializuje tzv. EMDS (*Environmental data management software*). Mnohé funkcionality jsou však dostupné v nástrojích business intelligence (BI). V této práci jsou proto představeny existující EDMS a prezentovaný informační systém HgIS, jenž využívá software vyvinutý pro řešení úloh BI.

3.1 Existující EDMS, výměnné formáty a datové modely

Předkládaná práce je výsledkem důkladné rešerše a testování existujících produktů ve snaze vyhnout se, pokud možno, vývoji nového IS. Existující sw se ukázal jako nevyhovující, a proto byl vyvinut nový IS. Tato kapitola popisuje existující EDMS, datové modely, formáty pro uložení a výměnu dat, standardy a normy. Předkládaná práce se vůči těmto vymezuje a přináší řešení, jež by mělo mít výhodu oproti níže uvedeným systémům a standardům.

3.1.1 Existující EDMS

EDMS čili *Environmental data management software* (sw pro správu dat o životním prostředí) je systém, který spravuje bázi dat o životním prostředí a provádí nad ní operace specifické pro environmentální data (import, export, validace dat, QA/QC, vizualizace: reporty, grafy, vrtné profily). Nebyl nalezen zdarma dostupný EDMS. Byly nalezeny následující EDMS (v závorce je uveden původce) EQuIS (EarthSoft), SiteFX (EarthFX), Enviro Data (Geotech Computer Systems), ESdat, GW-Base (ribeka), WISKI (Kisters), Hydstra (Kisters), GeODin (Fugro), Oasis-montaj (GeoSoft), HydroManager (Waterloo Hydrogeologic), GT-Web – dříve Waiora (Groundswell Technologies), Project Portal (ddms), gINT (Bentley), EPIPHINY (Summit Envirosolutions), Virtual Observatory and Ecological Informatics System (VOEIS), gdBase (GD Software,), SIS (vyvíjí adiNET pro AQUATEST), SED (Progeo Consulting), ETRA GIS (VŠB – TU Ostrava), H+ (Réseau National de Sites Hydrogéologiques), National Groundwater Information System (The Australian National University), GWSDAT (Wayne Jones), IMS – Information Management System (DHI), HydroDaVE, DataSight (Seveno), Geological Data Management (Software Datamine), PINE (USGS), QUIMET (GHS Barcelona), Visual Site Manager (GeoAnalysis), GIM Suite (acquire), GEMS (GEOVIA), Enterprise Integration (MineRP), BoreIS (McCarthy a Graniero), DigiFract (Hardebol a Bertotti) a MGeobase (Deltares).

Bylo provedeno důkladné srovnání uvedených EDMS. Z nich byly vybrány systémy EQuIS od firmy EarthSoft Inc. („světová jednička“) a Enviro Data od firmy Geotech Computer Systems, Inc., které mají zdrojový kód dostupný za podmínek dohody o mlčenlivosti a jsou zaměřeny na interoperabilitu se sw třetích stran. Tyto systémy byly otestovány. S majiteli firem, které tyto dva produkty vyvíjejí, autor vedl jednání. Bylo plánováno zakoupení systému EQuIS. Nákup se nepodařilo realizovat, což se ze zpětného pohledu jeví jako šťastná okolnost. Nabyté zkušenosti s uvedenými systémy byly využity pro návrh vlastního řešení, které není nepřiměřeně zatíženo autorskými právy třetích stran.

3.1.2 Existující výměnné formáty a datové modely

Databáze (datový sklad) je klíčovou komponentou HgIS, a je proto důležité navrhnout datový model (schéma databáze), který umožní efektivní uložení všech potřebných typů dat. I v případě nevyužití existujícího EDMS by mohlo být vhodné využít existující standardní datový model.

Celosvětovým mezinárodním standardem pro uchování a výměnu geografických dat je Geography Markup Language (GML) a jeho aplikační schémata, jako je GeoSciML (Geoscience Markup Language – pro geologická data), WaterML (pro hydrologická data) či Ground Water Markup Language (GWML – pro data o podzemní vodě). Na tyto standardy navazuje velmi důležitá evropská směrnice INSPIRE (INfrastructure for SPatial InfoRmation in Europe). Pojmenování položek definuje např. ČSN EN 14968 Sémantika pro výměnu dat o podzemní vodě. Dalšími zdarma dostupnými datovými modely jsou ty od firmy ESRI. Jedná se o Arc Hydro pro hydrologická data a Arc Hydro Groundwater pro hydrogeologická data. Další jsou Open Geoscience Data Models od Britské geologické služby.

Další prostudované standardy a datové modely jsou (původce: standard): Open Geospatial Consortium (OGC): WaterML-WQ, O&M, HY_Features: a common hydrologic feature model; CUAHSI (The Consortium of Universities for the Advancement of Hydrologic Science, Inc.): Observations Data Model (ODM); BoreholeML; SEP3 ; Bundesanstalt für Gewässerkunde: XHydro; Wojda: HydroCube, Hg20; Réseau National de Sites Hydrogéologiques: H+; Basin of Mexico hydrogeological database (BMHDB); Australian Government: Bureau of Meteorology Australian national groundwater data transfer standard, National Groundwater Information System: Data Model; CSIRO and Bureau of Meteorology: Water data transfer format; Wilson Engineering : Water Resources Database (WRDB); Geological Survey of Den-

mark and Greenland (GEUS): Borehole and Water Works Database; USGS: National geochemical survey database; Bavorský spolkový úřad pro životní prostředí: Hydrogeological Map of Bavaria; BRGM: SANDRE; HydrIS a Environmental Database Model (EnviroDB).

Uvedené datové modely jsou z hlediska požadovaných typů dat neúplné. Mezinárodní standardy se v čase poměrně rychle mění, protože se teprve vytvářejí. O sjednocení standardů pro data o podzemní vodě (GWML1, GeoSciML, INSPIRE) usiluje Groundwater Interoperability Experiment 2. Jeho výsledkem má být v červnu 2016 druhá verze Ground Water Markup Language (GWML2).

3.1.3 Hodnocení stávajících řešení

Existující sw, formáty pro výměnu a uložení dat a standardy se ukázaly jako nevhodné pro požadované praktické využití. Datové modely jsou často nedokonale dokumentované, což omezuje jejich využití v HgIS. Níže jsou stručně uvedeny problémy existujících řešení.

EDMS:

- Některé jsou technologicky zastaralé (např. MS Access jako běhové prostředí).
- Některé jsou uzavřené a neumožňují nezávislý rozvoj.
- Je třeba je upravit pro české prostředí (formáty pro import a export vč. ohlašovacích povinností, lokalizace, legislativní limity atd.).

Standardy, datové modely:

- Standardů je mnoho a mění se.
- Neobsahují šíři dat. Pro vytvoření struktury pro všechny typy dat by bylo třeba zkombinovat různé standardy, a výsledek by tedy byl nestandardní. Ty relativně úplné (GML, INSPIRE) jsou nepřiměřeně komplexní a zaměřené pro jiný účel.
- Jejich efektivní využití vyžaduje EDMS. Existující EDMS však většinou nepodporují mezinárodní standardy.

Všechny uvedené EDMS a standardy byly excerpovány a znalosti byly využity pro tvorbu vlastního datového modelu (kap. 3.4.2.2). Navíc byla excerpována i data a datové formáty, s nimiž se autor setkal a jež má k dispozici.

3.2 Typy informačních systémů

Volba architektury a komponent IS závisí na pojetí úlohy. Pokud budeme považovat za nejdůležitější požadavky na IS správu prostorových dat a zobrazování dat v mapách, zvolíme pravděpodobně geografický informační systém (GIS). Pokud požadavky zformulujeme jako správu dat, bude klíčovou komponentou databáze (DBMS, OLTP). Pokud chápeme jako klíčovou interakci s uživatelem, snadnou editaci a dostupnost přes internet, budeme se soustředit na vývoj GUI či online aplikace. Pokud však zadání chápeme jako načítání archivních i aktuálních dat z rozmanitých (strukturovaných i semi-strukturovaných) zdrojů, které se již dále nemění; zobrazení dat v tabulkách a grafech (včetně možnosti stažení v běžně užívaných formátech – MS Word a MS Excel), analýzu dat a tvorbu modelů; tak se jedná o typické zadání pro řešení BI.

Klíčovou komponentou HgIS je ETL a databáze (datový sklad). HgIS neobsahuje transakční databázi pro správu dat – OLTP, ale analytickou (OLAP). Transakční databáze je vhodná pro laboratorní

informační systém (LIMS) či provozní IS (úprava dat z hladinoměřů či výpočet průtoků z hladin atd.). HgIS může snadno integrovat data ze všech takových zdrojů.

HgIS kombinuje nástroje BI a GIS. HgIS pracuje především s bodovými objekty (vrty, studny) a prostorovou operací je v současnosti v podstatě pouze převod souřadnicových systémů. Využití jiných komponent GIS než online mapové aplikace je proto minimalistické. Využití komponent GIS v HgIS je uvedeno níže (tabulka 4).

Tabulka 4: Běžné komponenty GIS

Komponenta	Účel	Využito v HgIS
Prostorová databáze	Úložiště	Částečně (datový model HgIS není založen na standardu prostorových dat)
Online aplikace	Lehký klient	Ano
Desktop GIS	Těžký klient	Částečně (editace dat v PostGIS)
Mapový server	Server	Částečně (jen zobrazení podkladových map)
Prostorový ETL	Načítání dat	Částečně (běžně se využije jen převod souřadnic)

Na rozdíl od běžných úloh BI pracuje HgIS s malým množstvím dat, a proto komponenty jako DSA, ODS nejsou potřeba. Databáze HgIS plní z důvodu malého množství dat funkci datového skladu i OLAP kostky. Využití komponent BI v HgIS je uvedeno níže (tabulka 5). BI podrobněji rozebírá následující kapitola 3.3.

Tabulka 5: Komponenty BI

Komponenta	Anglicky	Účel	Využito v HgIS
Produkční (zdrojové) systémy	OLTP, legacy	Zdroj dat	Geofond, Lab-systém, Geobanka
Dočasné úložiště dat	DSA: <i>Data staging area</i>	Dočasné uložení extrahovaných dat	Ne (málo dat)
Operativní úložiště dat	ODS: <i>Operational data store</i>	Operativní úložiště pro analýzy	Ne (málo dat)
Datová pumpa	ETL: <i>Extract, transform and load</i>	Integrace dat	Ano: PDI
Integrační nástroje	EAI: <i>Enterprise Application Integration</i>	Integrovat primární podnikové systémy	Ne (nevyužíváme primární systémy)
Datový sklad	DWH: <i>Data warehouse</i>	Hlavní úložiště dat	Ano: PostgreSQL
Datové tržiště	<i>Data mart</i>	Problémově orientovaný DWH	Ne
Datová kostka	<i>OLAP cube: On-line Analytical Processing Cube</i>	Uložení dat pro analýzy	V HgIS dostačuje DWH
Tiskové sestavy	<i>Reporting</i>	Jednoúčelové zobrazení dat určené k tisku	Ano: Pentaho Reporting
Pracovní panely	<i>Dashboards, scorecard</i>	Přehledné a interaktivní zobrazení dat	Zatím ne (Pentaho Business Analytics)
Dolování dat	Data mining		Weka – zatím ne

3.3 Business intelligence

Problematika správy dat, vizualizace, analýz a modelování podzemní vody je v mnoha ohledech podobná business intelligence (BI). BI je soubor dovedností, znalostí, technologií, aplikací a postupů používaných v podnikání pro získání lepšího pochopení fungování obchodní společnosti pro potřeby manažerského rozhodování. Za tímto účelem provádí sběr, integraci, analýzu, interpretaci a prezentaci firemních dat. Běžné funkce aplikací BI zahrnují ETL, uložení všech dat v datovém skladu (*data warehouse*), OLAP (*OnLine Analytical Processing*), reporting, přehledové zobrazení (*dashboard*), analýzy a dolování dat (*data mining*) či automatizovaná hlášení (alerty). Významnou roli hraje zpracování semistrukturovaných dat. S geodaty pracuje *geospatial BI* čili GeoBI. Obdobné pojmy jsou *location intelligence* a *spatial intelligence*.

Cílem předkládané práce bylo získat (často semistrukturovaná) data, uložit je do jedné databáze a využít je pro reporty, analýzy a modely. Technicky se tedy jedná o podobnou úlohu, kterou řeší BI (DWH). Metodika a především nástroje jsou do jisté míry přenositelné. Naše problematika se týká prostorových dat (řešeny GeoBI) a geologických dat (BI systémy se geologií nezabývají). Nebyla nalezena práce, která by využívala nástroje BI v hydrogeologii.

Pro BI byly vytvořeny sady nástrojů (*BI tools*). Pentaho (community.pentaho.com) je aktivně vyvíjený FOSS s permissivní licencí. Odvozená verze (*fork*) těchto produktů – např. GeoKettle – pracují s prostorovými daty a vyvíjí je firma Spatialytics. Ekosystém Pentaho obsahuje mj. datovou pumpu (ETL) Kettle (čili Pentaho Data Integration – PDI), nástroj Weka pro dobývání znalostí z databází (*data mining*) a Pentaho Report Designer pro návrh reportů, jež poté mohou být zobrazovány na aplikačním serveru PBA (Pentaho Business Analytics Platform – dříve BI Server). Na PBA lze také vytvářet a spouštět přehledová zobrazení (*dashboard*) a je možné je integrovat do dalších aplikací. Datový model lze při tvorbě reportů lokalizovat pomocí Pentaho Metadata Editor – vytvoří se abstraktní vrstva nad datovým modelem a zjednodušuje se tak tvorba tiskových sestav a přehledových zobrazení.

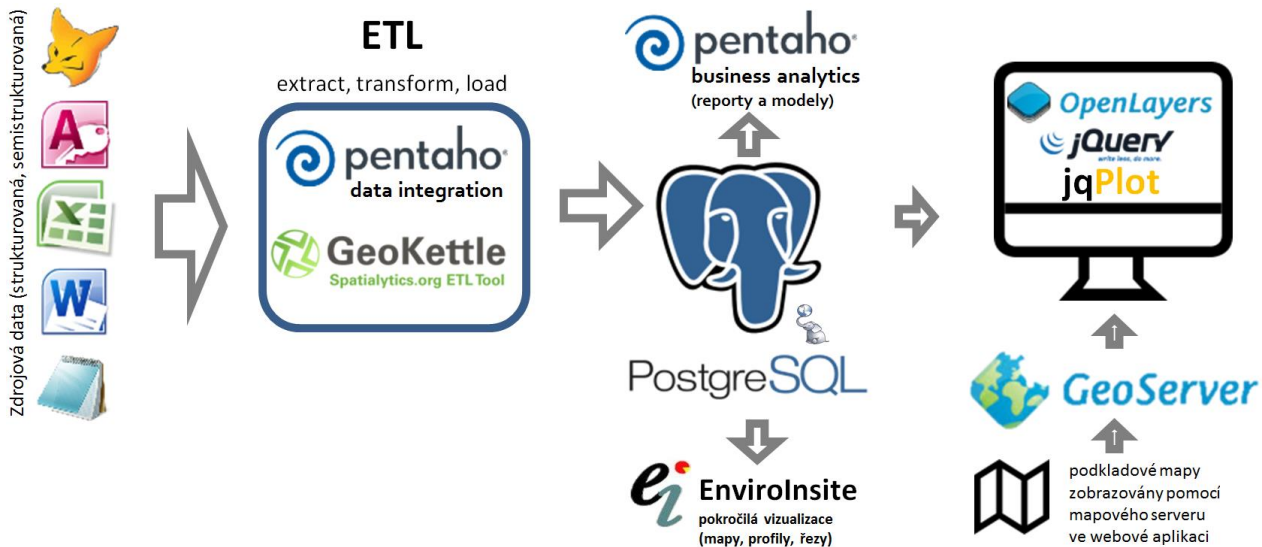
V HgIS jsou využívány některé nástroje Pentaho. Důvodem volby je dobrá integrace jednotlivých komponent, dostatečná funkcionality již v komunitní (FOSS) verzi, existence verze ETL pro práci s prostorovými daty (GeoKettle) a uživatelská přívětivost GUI. Alternativou k Pentaho je například SpagoBI. Komerční alternativou jsou např. nástroje od firmy Microsoft.

3.4 Návrh informačního systému HgIS

V této kapitole se podrobně rozebírá volba komponent HgIS.

Klíčovou komponentou HgIS je databáze (datový sklad), která slouží k uložení strukturovaných dat. Data jsou načítána ze zdrojových souborů pomocí datové pumpy (ETL) PDI a případně GeoKettle. Data jsou zobrazována v reportech na BI serveru (PBA) a interaktivně v mapové online aplikaci (mapy, grafy, tabulky). Mapové poklady mohou být využity mj. v online aplikaci díky mapovému serveru. Další zobrazení je možno provádět pomocí exportu do specializovaného sw (3D program pro zobrazování environmentálních dat EnviroInsight). Celkové schéma HgIS je uvedeno níže (obrázek 6).

HgIS využívá především komponenty BI (produkty firmy Pentaho) a GIS (PostGIS a GeoServer jsou produkty firmy Refractions Research).



Obrázek 6: Schéma architektury HgIS

3.4.1 Geologická vizualizace – EnviroInsite

Pro pokročilou vizualizaci hydrogeologických dat (která je velmi užitečná pro tvorbu koncepčního modelu) byl zvolen sw EnviroInsite – dále EI. Jedná se o cenově dostupný komerční program implementovaný v .NET. Software je velmi flexibilní a dokáže zobrazovat veškerá hydrogeologická data. Práce s ním je intuitivní. Jeho cílem je, aby potřebnou vizualizaci dat mohl běžně provádět sám hydrogeolog a nepotřeboval k tomu specialistu GIS. Program zobrazuje data z databáze s danou strukturou. Databáze může být implementována v programu MS Access či MS Excel. Software zobrazuje dokumentaci jednotlivých vrtů, geologické řezy, 3D vizualizaci geologie, mapy, chemické interpretační grafy (Piper, Stiff, Schoeller) umístěné na mapě nebo na samostatném listu. Obdobně zobrazuje souhrnné i detailní tabulky; grafy a tabulky časových řad. Software interpoluje data ve 2D i 3D a je možno do něj načíst běžné formáty podkladových map. Alternativou by byl například HydroGeoAnalyst či RockWorks.

3.4.2 Databáze

Datový model HgIS vychází z datového modelu programu EnviroInsite. To umožnilo jednoduše vyvinout export dat pro zobrazení v EnviroInsite. Navíc struktura souborů pro načítání dat a celková logika dat zůstává stejná. Nejlepším alternativním datovým modelem, z něhož by bylo možno vyjít, je datový model EDMS SiteFX, který lze zdarma používat a šířit. V HgIS je využíván DBMS PostgreSQL, protože má prostorové rozšíření PostGIS a jedná se o silný FOSS vyvíjený komunitou.

3.4.2.1 Datový model EnviroInsite

Datový model EI obsahuje data o objektech (vrty, studny), a to geologický popis, stratigrafii a technické provedení vrtů. Dále obsahuje pozorování veličin (např. analytů) v hloubkových intervalech (např. perforace vrtu či interval odběru zeminy) a bodech v hloubce (např. karotáž). Databáze implementovaná v MS Excel či MS Access musí obsahovat následující strukturu odpovídající relační databázi, aby program EnviroInsite mohl uložená data zobrazovat). Některé sloupce nejsou povinné. EI může zobrazovat v dokumentaci vrtu (vrtném profilu) data z uživatelem definovaných polí. EI umožňuje, aby soubor obsahoval další tabulky a sloupce. EI přistupuje k datům přes ODBC.

3.4.2.2 Datový model HgIS

Datový model EI definuje pole, která je třeba zadat pro zobrazení v EI. Pro uložení všech dat, s nimiž HgIS pracuje, byl datový model EI rozšířen tak, aby umožnil uložení popisných dat o sondách (vrty), podmínkách vzorkování podzemní vody a časových intervalech. Dále aby obsahoval zejména číselníky, kódovníky a pomocná data pro načítání dat do systému, pro převod jednotek a veličin a pro přejmenování. Datový model HgIS existuje ve dvou verzích.

1. První verze slouží jako výměnný formát a je málo normalizovaná. Obsahuje vícepoložkové (složené) primární klíče (*composite primary keys*), které mají význam v reálném světě (název vrtu, název veličiny atd) – tzv. přirozené klíče (*natural keys*). Pro návrh byl použit MS Access z důvodu zpětné kompatibility s EI.
2. Druhá verze slouží pro nasazení na serverové databázi a je více normalizovaná. Obsahuje umělé klíče (*surrogate keys, ID*), z tabulky *constituents* je vyčleněna tabulka *standards* a obsahuje číselníky *kvalita* a *well_construction_material*. Navíc obsahuje další tabulky pro data (sestavy analýz, plán vzorkování), metadata (obecná struktura pro zaznamenání interpretací k uloženým datům) a lokalizaci, které budou moci být v budoucnosti případně využity v HgIS. Struktura databáze odpovídá do jisté míry datovému skladu ve schématu sněhové vločky (*snowflake*) či souhvězdí (*fact constellation*), kde tabulky *observations* a *point_values* hrají roli faktových tabulek a například *screens* a *constituents* roli dimenzí (*observations.date_* je pak degenerovaná dimenze). Velký počet atributů (denormalizace) např. v tabulkách *well* či *vzorky* a málo číselníků odpovídá běžné struktuře datových skladů. To odpovídá koncepci HgIS, protože taková struktura je vhodná pro analýzy, resp. reporting a nikoliv pro editaci dat více uživateli s udržením konzistence všech dat, jak je tomu v případě transakčních databází (OLTP).

Datový model byl vytvořen tak, aby maximálně využíval přístupy, pojmy atd. existujících standardů, EDMS a výměnných formátů a umožňoval uložení všech relevantních dat.

V rámci předkládané práce byly na základě české terminologie (Labsystém, karotáž AQUATEST a.s., zdroje dat atd.) vytvořeny číselníky veličin, další číselníky, kódovníky a řízené slovníky.

3.4.2.3 Prostorové databáze

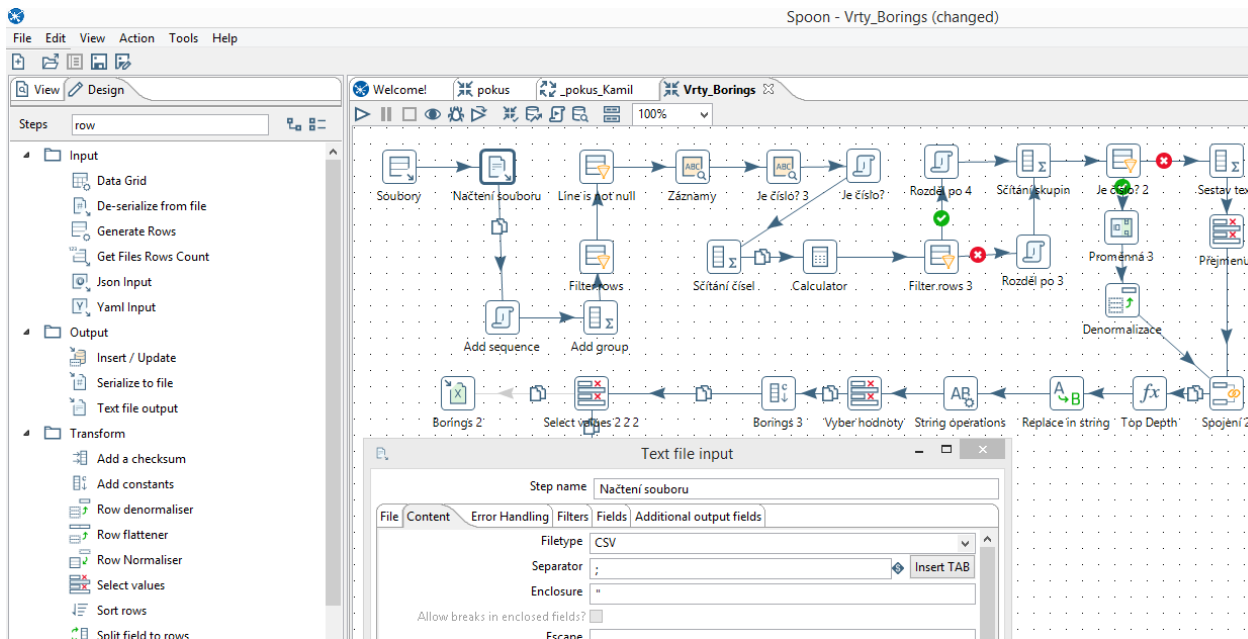
Datový model HgIS nepodporuje přímo žádný standard prostorové databáze, protože je určen zejména pro správu bodových dat (vč. vrtů), nikoliv např. polylinií a polygonů, jež jsou do značné míry prostorovou interpretací bodových měření. V HgIS je využíván PostGIS (*postgis.net* – prostorové rozšíření databáze PostgreSQL) v odděleném schématu – bez jinak striktně definovaného datového modelu. Pro tvorbu a editaci dat v prostorové databázi je možno využívat desktop GIS (např. ArcGIS či QGIS).

3.4.3 ETL

Pro načtení rozmanitých strukturovaných i semistrukturovaných dat je možno použít nástroje ETL (*Extract, Transform, Load*). *Extract* neboli extrakce znamená získání dat ze zdrojů. *Transform* (česky transformace) představuje zpracování dat, a to např. konverzi, filtrování, normalizaci, denormalizaci, matematické operace či kontrolu údajů. Třetí složkou ETL je *load*, česky naplnění. V této fázi se jedná o nahrání zpracovaných dat do cílového úložiště (soubor, databáze). (Geo)prostorový ETL čili (*geo*)*spatial ETL* je ETL (geo)prostorových dat. Jedná se o systémy, které umožňují pracovat systémově s prostorovými daty (geografické souřadnice) a řešit prostorové úlohy (podobně jako GIS).

V HgIS je využíván Pentaho Data Integration (PDI), známý též jako Kettle. K PDI existuje verze, jež je primárně zaměřena na práci s prostorovými daty (GeoKettle). PDI je interpretem procedur ETL ve

XML. PDI obsahuje dílčí programy. Spoon (obrázek 7) je integrované prostředí pro vývoj, testování a monitorování transformací – využívá tzv. vizuální programování. Z hlediska praktické práce s vývoje-
ným prostředím PDI (Spoon) hrají hlavní roli tzv. kroky, pomocí kterých se sestavují transformace.
Jedná se o prvky, jež mají specifické funkce. Při tvorbě transformací uživatel využívá kroků, které spojuje
šipkami (*boxes and arrows*). Ty určují směr, jakým transformace probíhá. K dispozici jsou i kroky, které
umožňují použít v transformaci vlastní kód Java nebo příkaz SQL.



Obrázek 7: Uživatelské rozhraní Spoon – nabídka kroků (levý sloupec) po zadání výrazu „row“

Jádro Data Integration je obsaženo i v PBA. Uživatel PBA tak může snadno spouštět transformace či *jobs* PDI přímo či pomocí tzv. řešení (*solutions*). Řešení je sled akcí, který je definován v souboru *.xaction*.

GeoKettle, je prostorový *fork* obecného ETL nástroje PDI (Kettle). Nevýhodou je, že aktuální verze je založena na starší verzi PDI (3.2). Využití GeoKettle je plánováno pro převody souřadnic objektů a další prostorové úlohy, které je vhodné automatizovat (např. kap. 2.3.2). Využil jej ve svém bakalářském projektu Hadač (2014) pro převody souřadnic mezi prostorovými referenčními systémy S-JTSK a WGS 84.

3.4.4 Reporting (tiskové sestavy)

Reportingem rozumíme tiskové sestavy, analytické tabulky, grafy a přehledy realizované na základě dotazů do databází. Tyto sestavy je možno zobrazovat ve webovém prohlížeči i stahovat v různých formátech (PDF, RTF, MS Excel). Reporty je možno vytvářet pomocí GUI Pentaho Report Designer. Reportingové vykreslovací jádro je součástí PDI i PBA. Proto je možno zobrazovat reporty online na základě uživatelských voleb (výběr objektu, veličiny atd.) na serveru PBA a reporty v definovaných formátech je možno generovat a distribuovat pomocí PDI.

3.4.5 Mapový server

Pro publikaci prostorových dat kromě samotných bodů je možné využít mapový server. Zobrazení dat z bodových objektů (např. vrty, studny) je možné v programu EnviroInsight či ve webové aplikaci. V testovacím režimu byl nasazen GeoServer (geoserver.org), který by měl být využíván i nadále. Výstupy

z mapového serveru mohou být publikovány pomocí standardizovaných služeb, jako jsou např. WMS či WFS. Ty mohou být zobrazeny v desktopovém GISu, EnviroInsite či začleněny do webové aplikace. Publikace dat pomocí webových služeb je jednou z forem naplnění směrnice INSPIRE. Využití mapového serveru v HgIS je v současném stavu pouze okrajové.

3.5 Implementace

V této kapitole je popsáno konkrétní uplatnění výše uvedených nástrojů v HgIS. Vlastní implementace byla provedena částečně studenty TUL v rámci jejich školních prací, dohod a v rámci zakázek. Autor předkládané práce tyto dílčí úkoly navrhl, zadal a metodicky vedl.

3.5.1 ETL

Byly vytvořeny transformace pro načítání dat do databáze. Jsou rozděleny do tří navazujících skupin: A – Převod dat ze zdrojových formátů do výměnného formátu. Jedná se konkrétně o:

- Laboratorní informační systém (LIMS) Labsystém (www.labsys.cz): sada 2 souborů DBF.
- Česká geologická služba (Geofond) – soubory MS Access.
- Inklinometrie (textové soubory – strojově generované exporty od AZ Consult).
- Data o chemismu z Palivového kombinátu Ústí, s. p. v Excelu (ručně vytvořené dokumenty).
- Archivní průzkumné vrty ve Wordu (2413 ks) – sw Geobanka firmy Data-PC Sokolov.
- gdBase – načítání některých polí.
- Srážky a teploty ve formě textového souboru na FTP serveru Povodí Ohře s. p.
- Kontingenční tabulka – obecný formát.
- Data (v Excelu a v textových souborech) v různých formátech různých dodavatelů.

B – Úprava, čištění, validace, sjednocení užití terminologie, výpočty veličin, agregace atd.

- Výpočet hydrochemického typu vody (např. Ca-Mg-HCO₃).
- Agregace dat (např. roční úhrn srážek z denních srážek, minimální měsíční průtok v roce atd.).

C – Načtení do databáze (datového skladu).

D – Export pro další analýzu v sw třetích stran (plánované exporty jsou v kap. 3.7.1):

- Průměrný hydraulický gradient z hladin vybraných vrtů a z toho odhad rychlosti proudění a doby doběhu – export do sešitu MS Excel, kde je výpočet prováděn maticovými vzorci.
- Export dat do kontingenční tabulky.

Pro vizualizaci hydrogeologických dat je používán mj. EnviroInsite. Data pro něj jsou výsledkem transformace B či D.

3.5.2 Reporting (tiskové sestavy)

Pomocí Pentaho Report Designer a PDI byly vytvořeny následující analýzy a reporty:

- Graf a tabulka časového průběhu libovolných veličin v libovolných objektech.
- Identifikace redukčně-oxidačních podmínek ve vodách z chemického složení.
- Multikriteriální analýza hodnotící, zda dojde k eutrofizaci ve vodní nádrži.
- Profil vrtu – report dokumentace geologicky dokumentovaných objektů.
- Výpočet hydraulického gradientu z libovolného počtu objektů.

3.5.3 Mapová online aplikace a správa nestructurovaných dat a dokumentů

V rámci bakalářské práce byla vyvinuta online aplikace (Jodas 2014). Práci vedl a funkcionalitu kompletně navrhl autor předkládané práce. Cílem mapové aplikace je provázat mapy, tabulky a grafy a názorně prezentovat většinu dat HgIS. Aplikace obsahuje interaktivní podkladovou mapu s objekty. Pro jednotlivé objekty (vrty atd.) zobrazuje grafy a tabulky libovolných veličin (nejvýše dvou zároveň). Jsou to grafy časových řad (tabulka *observations*) či karotáže (tabulka *point_values*). V grafu se zobrazují i meze (sanační limity, legislativní omezení atd.) z tabulky *standards*. K jednotlivým objektům se zobrazí *popup* okno s informacemi o objektu z tabulky *wells*. Je možno nastavit různé podkladové mapy (např. WMS). V jednoduché textové formě je možno zobrazovat geologii, stratigrafii a technické provedení objektů (vrtů, studen). Aplikace využívá PHP, mapový framework OpenLayers, jQuery, AJAX a jqPlot.

Data sbíraná v rámci projektů řešených HgIS jsou v současnosti ukládána na souborový server dostupný přes FTP. Vybraná data jsou načítána do databáze (datového skladu). Pro systematické uložení zejména publikací (a dokumentů) je využíván bibliografický manažer Zotero (zotero.org).

3.6 Aplikace HgIS

HgIS byl doposud prakticky využit následujícím způsobem: Data, se kterými se pracuje v rámci projektů MARE a ZÁVOD, byla načtena do databáze, jsou zobrazována ve webové aplikaci a analyzována s pomocí multikriteriální analýzy (PDI a reporty v PBA). Jakub Říha nad databází naprogramoval srážko-odtokový model jezera Milada – Chabařovice. Některé komponenty HgIS jsou součástí IS MARE, který je na komerční bázi rozvíjen a nasazován na Palivovém kombinátu Ústí, s. p. Načítání dat z *gdBase* bylo využito pro transformace dat pro projekt Povodeň. Raná vývojová verze systému byla využita pro zpracování dat pro případovou studii (kap. 2.5.) a pro vyhodnocení a modelování geochemické odezvy sanačního zásahu na lokalitě Kuřívody.

3.7 Diskuse a závěr

Byl vyvinut informační systém, který využívá nástroje BI (Pentaho) a GIS (OpenLayers atd.). Souhrnně je tak možno HgIS charakterizovat jako „*(geo)spatial business intelligence (GeoBI) tool for hydrogeology*“. Takový systém doposud neexistoval. V ESdat jsou integrovány komerční nástroje pro reporting – SSRS, Telerik, Crystal Reports. Telerik je využit také v EQUIS. Na nástrojích BI od firmy Microsoft je postaven MineRP (minerp.com), který však zpracovává data týkající se agend v souvislosti s těžbou nerostných surovin. Boulil et al. (2014) prováděli OLAP analýzu kvality povrchové vody s využitím mj. Talend a PostgreSQL.

Pro tvorbu koncepčních modelů podzemní vody je vhodné integrovaně zobrazovat data různého typu. Jedná se o geologii, technické provedení objektů, hladiny, chemismus, karotáž, a to přímo měřené hodnoty i interpolované. To umožňuje sw EI, jehož datový model byl použit jako základ pro datový model HgIS. Logika uspořádání dat v EI je tak pro uživatele HgIS snadno pochopitelná, což zjednodušuje využití EI pro tvorbu koncepčních modelů. HgIS navíc umožňuje snadno vytvářet procedurální modely. Jednoduché modely je možné spouštět v PBA, pro náročnější simulace je možno data snadno exportovat pomocí PDI do formátu pro načtení do specifického simulačního software. Výpočty je možno provádět v PDI například za pomoci kroků *Formula*, *Analytic query* či *Calculator*. V reportech je možno používat vzorce, jejichž syntaxe (*LibFormula*) je založena na standardu (*OpenFormula*), který využívá například

LibreOffice/OpenOffice. HgIS je navržen tak, aby umožňoval vytvářet dílčí (geo)prostorové analýzy v (*geo*)spatial ETL GeoKettle či nad databází např. pomocí prostorového rozšíření jazyka SQL. Těmito vlastnostmi je HgIS unikátní. Zvolená kombinace nástrojů a přístupů umožnila navrhnout a částečně implementovat systém tak, aby odpovídal požadavkům z hlediska funkčnosti i architektury (tabulka 6).

Tabulka 6: Splnění požadavků na systém z hlediska architektury

Požadavek	Způsob splnění
Zahrnutí všech dat při zachování jednoduchosti. Struktura i pro doposud neznámé veličiny (<i>data agnostic</i>).	Datový model, ve kterém jsou v podstatě všechny časové řady v jedné faktové tabulce. Veličiny jsou definovány pomocí záznamů v tabulce, nikoliv samostatnými sloupci (normalizace).
Škálovatelnost (lokální i serverové nasazení, paralelizace transformací, <i>big data</i>).	Využití nástrojů BI, paralelizace transformací pomocí ETL serveru Carte (součást PDI).
Přenositelnost – využitelnost v rámci jiných IS.	Pentaho je možno integrovat do jiných IS.
Automatizované zpracování dat včetně netriviálních operací s prostorovými daty (nejen body).	Využití PDI, GeoKettle a PostGIS.
Rozvoj může provádět pokročilý zaškolený uživatel (<i>power user</i>).	Využití nástrojů BI s GUI.
Snadná implementace systému včasného varování a automatických hlášení.	PBA a PDI umožňují např. odeslání reportu e-mailem v pravidelných intervalech či při definované události (.xaction).
Snadné propojení s nástroji pro analýzy dat a data mining	Využití PDI (<i>steps</i> : Weka scoring, ARFF output, Tableau data extract, Execute R script).
Snadná tvorba exportu do specializovaného modelovacího sw či implementace speciální analýzy.	Využití PDI, Report Designer, PBA.

Architektura HgIS by mohla být kompaktnější – tedy nikoliv poskládaná z velkého množství existujících sw nástrojů. Při zvolení jiné architektury (*Service Oriented Architecture*) by bylo možné jednotlivé komponenty operativněji kombinovat. Bylo by elegantní začít datové modelování na vyšší míře abstrakce – na konceptuální úrovni, nikoliv vyjít z výměnného formátu (relační databáze), ten rozšířit a normalizovat. Využití ontologií či objektových databází by umožňovalo využít sofistikovanější datovou strukturu. Komunitní vrze Pentaho má mnohá omezení – např. neumožňuje šifrované spojení a PBA není přizpůsoben pro zobrazení v mobilních zařízeních. Je však třeba zdůraznit, že novost předkládané práce není v oblasti informatiky, ale v představení řešení GeoBI, které vyhovuje zadaným cílům a potřebám hydrogeologa – modeláře.

Nevýhodou HgIS je, že na rozdíl od některých EDMS se nejedná o kompaktní sw, ale o soubor více modulů. HgIS jako celek nemá jednotné GUI, které by uživatel mohl ihned začít intuitivně používat. Srovnání s některými EDMS je tak může být poněkud zavádějící. Nicméně platí, že nebylo nalezeno využití BI (včetně DWH, OLAP či ETL) pro problematiku podzemní vody. V práci se blíže nerozebírají principy DWH a OLAP, protože už datový model EI v podstatě odpovídal schématu sněhové vločky či souhvězdí (*snowflake* či *constellation*).

Použitý přístup se velmi osvědčil, jelikož umožňuje snadnou udržitelnost a rozšiřitelnost. Je plánováno využití pro další projekty na TUL. Firmy zabývající se hydrogeologií v ČR EDMS nepoužívají

(s výjimkou gBase), případně využívají velmi specifický sw, který odpovídá dílčím požadavkům. HgIS má tedy dobré předpoklady, aby mohl být široce používán.

Přínosem předkládané práce je však především to, že identifikovala potřeby hydrogeologů a navrhla koncepční přístup pro práci s daty. Hydrogeologové po desetiletí běžně používají nástroje GIS, ale nikoliv BI. Tato práce představuje BI hydrogeologické komunitě. HgIS prokazuje funkčnost zvoleného přístupu a otevírá příležitosti pro využití dílčích komponent od jiných dodavatelů.

3.7.1 Plánované rozšíření

Dále je plánováno vyvíjet další analytické a simulační nástroje a exporty do speciálních nástrojů pro vyhodnocování, které naplní strategii využívání více jednoduchých modelů. Mezi takové patří např.: výpočet geochemického pozadí či export měřených dat do CoronaScreen pro výpočet ustálené délky kontaminačního mraku. Dále je možno přidávat další funkcionalitu jako např. exporty dat pro plnění ohlašovací povinností, atd. V současnosti je v rámci studentských prací vyvíjena:

- Online aplikace pro zadávání a editaci dat a metadat
- Přehledová zobrazení (*dashboard*) za pomoci CTools.
- Řešení (*solutions*) – automatizace procesů za pomoci posloupností událostí.
- Abstraktní vrstva nad datovým modelem pro lokalizaci a zjednodušení tvorby reportů a přehledových zobrazení koncovým uživatelem (využíván Pentaho Metadata Editor).
- Import dat z Databáze geologicky dokumentovaných objektů ČGS (XML ve formátu aplikace eEarth).

4 Shrnutí vědeckých výsledků

Práce shrnuje existující klasifikace nejistot (zejm. koncepčního modelu) a navrhuje novou klasifikaci využití modelů. Tato práce je unikátní tím, že se zabývá ekvifinálními a alternativními jednoduchými modely podzemní vody, jež jsou založeny na principálně odlišných přístupech či předpokladech. Princip je demonstrován na unikátních případových studiích, které vycházejí z malého množství vstupních dat. Tím se tato práce vymezuje vůči článku, jehož autor (Bredehoeft 2005) se nesetkal v praxi s využitím alternativních koncepčních modelů.

Pro tvorbu takových modelů byl vyvinut informační systém, jenž je založen na původním datovém modelu a využívá kombinaci typů sw (GeoBI = GIS + BI), která pro problematiku podzemní vody dosud nebyla použita. Největší vědecký přínos má práce chápána jako celek. Práce má vědecký přínos v oblasti hydrogeologie (modelování podzemní vody), případně environmentální informatika. Nemá vědecký přínos v oblasti matematika, matematická teorie rozhodování, či informatika.

Autor zpracoval práci samostatně s využitím citované literatury a konzultací. Dílčí úlohy řešili studenti Krejch (2013, 2015), Jodas (2014), Hadač (2014), Štírek (2014) a Láska (2011) v rámci svých projektů, bakalářských a diplomových prací, dohod o provedení práce či byly řešeny dodavatelsky. Studenti pracovali pod vedením autora této práce. Většina práce vznikala v rámci výzkumných projektů a zakázek, z nich je však prezentována pouze ta část, již autor zpracoval samostatně.

Tematika jednoduchých ekvifinálních modelů, tj. kap. 1 a 2, byla prezentována na kongresu a publikována v konferenčním sborníku (Nešetřil 2009b) – kap. 2.4. Další případové studie (kap. 2.1, 2.2 a 2.3) byly prezentovány na mezinárodním semináři, přičemž abstrakt byl publikován ve sborníku (Nešetřil 2012a). Výsledky kap. 2.6 byly publikovány v časopise (Topinkova et al. 2007), který je zařazen

v Science Citation Index ($IF_{2013} = 1,71$). Článek byl citován dalšími 5 články z téhož indexu. Tematika navazující na kap. 2.2 byla publikována (Matula et al. 2014) v časopise, který je zařazen v Science Citation Index Expanded (SciSearch) – $IF_{2013} = 1,68$, kde byl článek jednou citován. Tematika informačního systému byla publikována v recenzovaných sbornících konference EnviroInfo 2014 (Nešetřil a Šembera 2014) a českého hydrogeologického kongresu (Nešetřil 2014).

5 Shrnutí a závěr

Podzemí je obtížně přístupné přímému pozorování a bývá navíc velice heterogenní. Proto při studiu podzemní vody čelíme značné nejistotě. Práce reaguje na nejistotu ve výsledcích modelů podzemní vody. V práci je představeno shrnutí existujících klasifikací nejistot (zejm. koncepčního modelu). Je zde definována epistemologická nejistota a její projevy: nejistota koncepčního modelu a nejistota struktury modelu. Koncepční model není možné verifikovat či validovat, nýbrž je možné jej pouze potvrdit (*confirmation*) – podle Refsgaarda et al. (2006) je možné posoudit jeho obhajitelnost (*tenability*). Pro vypořádání se s takovou nejistotou je v práci obhajována a využita strategie více jednoduchých modelů. Ta nespočívá ve vytváření alternativních sad parametrů (varianty, scénáře) ani alternativních geologických modelů, ale ve využití jednoduchých alternativních modelů, které simulují jiné procesy, používají jinou metodu či představují odlišné koncepční přístupy (jsou principiálně odlišné). Tyto jednoduché modely mohou být formulovány jako testování hypotéz a jednoznačně snižovat nejistotu porozumění fungování studovaného systému. Zvolená strategie v podstatě odpovídá podle klasifikace Uusitalo et al. (2015) kombinaci přístupů „emulace modelu“ a „více modelů“ a také odpovídá tzv. „*multiple model ensemble*“. Podle Refsgaarda et al. (2007) odpovídá kategorii „simulace více modely – modely s odlišnou strukturou“. Využití modelů pro testování hypotéz je spolu s dalšími uveden v navržené klasifikaci využití modelů. V práci je diskutováno použití principu ekvifinality a tohoto pojmu v oblasti modelování podzemní vody. Strategie je demonstrována na unikátních případových studiích, které vycházejí z malého množství vstupních dat.

Bredhoeft (2005) uvádí, že mnozí autoři navrhují využití alternativních koncepčních modelů, ale že nikdy nezaznamenal uskutečnění tohoto přístupu v praxi, protože modeláři pracují s jedním koncepčním modelem, který mění až v případě, kdy jej není možné nakalibrovat. Předkládaná práce naopak demonstruje, jak je možno uplatnit alternativní koncepční modely v praxi – totiž používat (alespoň v první fázi) více jednoduchých modelů.

V kap. 2 je prezentováno využití alternativních jednoduchých modelů v modelování různých hydrogeologických úloh. Jedna úloha je řešena více modely, které jsou založeny na různých přístupech a předpokladech. Vysoká nejistota modelů je dána nedostatkem adekvátních vstupních dat. Prezentované případové studie srovnávají jednoduché a komplexní modely, testují alternativní hypotézy sadou jednoduchých modelů či představují alternativní prediktivní modely. Obhajitelnost modelů v případových studiích byla semikvantitativně vyhodnocena podle metodiky Refsgaarda et al. (2006).

Předkládaná práce nevytváří jen alternativní sady parametrů (varianty, scénáře) či alternativní geologické modely, ale jednoduché alternativní modely, které simulují jiné procesy, používají jinou metodu či představují odlišné koncepční přístupy (jsou principiálně odlišné). V některých případových studiích se neuplatňuje princip ekvifinality, protože modely se doplňují (např. kap. 2.3).

Multidimenzionální environmentální data je třeba vhodným způsobem spravovat, zobrazit a analyzovat, abychom mohli snadno sestavovat adekvátní alternativní koncepční modely. K tomu je možno využít existující software pro správu dat o podzemní vodě a životním prostředí a existující nástroje pro vizualizaci a analýzu dat. Práce obsahuje důkladnou rešerši takových nástrojů. Velmi se

však osvědčilo využití nástrojů, přístupů a technologií z naprosto jiné oblasti – manažerské informatiky. Úlohy řešené při zpracování dat o podzemní vodě a životním prostředí jsou do značné míry podobné úlohám business intelligence. Jedná se o:

- Načítání archivních dat z různorodých zdrojů (databáze, soubory atd.). Není ani tak důležité interaktivně editovat data, ale především je zobrazovat a analyzovat.
- Tvorba tiskových sestav, které jsou automaticky generovány, ale přitom je možno je ručně upravovat (formáty pro MS Word a MS Excel).
- Spouštění modelů a analýz nad komplexními daty.

Proto byl vyvinut informační systém HgIS (kap. 3), který využívá nástroje, přístupy a technologie BI. Důležitým požadavkem na HgIS bylo, aby umožňoval snadné načítání rozmanitých dat a export dat pro využití v nástrojích pro speciální analýzu dat a modelování. Tuto funkcionalitu zabezpečuje ETL (konkrétně PDI), která umožňuje automatizovaně importovat data z různých zdrojů a to i málo strukturovaných. Data jsou uložena v databázi s navrženým datovým modelem, který vychází z existujících standardů a produktů. V jeho návrhu se z BI uplatnily přístupy pro návrh datových skladů a OLAP. HgIS má některé komponenty a funkce GIS. Nebyla nalezena práce, která by využívala nástroje BI v hydrogeologii.

Předkládané práce identifikovala potřeby hydrogeologů a navrhla koncepční přístup pro práci s daty. Hydrogeologové po desetiletí běžně používají nástroje GIS, ale nikoliv BI. Tato práce představuje BI hydrogeologické komunitě. HgIS prokazuje funkčnost zvoleného přístupu a otevírá příležitosti pro využití dílčích komponent od jiných dodavatelů.

Použitá literatura

- BEVEN, Keith a Peter YOUNG, 2013. A guide to good practice in modeling semantics for authors and referees. *Water Resources Research*. Roč. 49, č. 8, s. 5092–5098. ISSN 1944-7973.
DOI: 10.1002/wrcr.20393
- BOULIL, Kamal, Florence LE BER, Sandro BIMONTE, Corinne GRAC a Flavie CERNESSON, 2014. Multidimensional modeling and analysis of large and complex watercourse data: an OLAP-based solution. *Ecological Informatics*. Roč. 24, s. 90–106. ISSN 1574-9541.
DOI: 10.1016/j.ecoinf.2014.07.001
- BREDEHOEFT, John, 2005. The conceptualization model problem—surprise. *Hydrogeology Journal*. Roč. 13, č. 1, s. 37–46. DOI: 10.1007/s10040-004-0430-5
- GUPTA, Hoshin V., Martyn P. CLARK, Jasper A. VRUGT, Gab ABRAMOWITZ a Ming YE, 2012. Towards a comprehensive assessment of model structural adequacy. *Water Resources Research*. Roč. 48, č. 8, W08301. DOI: 10.1029/2011WR011044
- HADAČ, Jakub, 2014. *Transformace zeměpisných souřadnic*. Bakalářský projekt. Technická univerzita v Liberci.
- HILL, Mary C. a Claire R. TIEDEMAN, 2007. *Effective groundwater model calibration : with analysis of data, sensitivities, predictions, and uncertainty*. Hoboken N.J.: Wiley. ISBN 978-0-471-77636-9.
- JODAS, Tomáš, 2014. *Webová prezentace prostorových dat*. Bakalářská práce. Technická univerzita v Liberci.
- KREJBICH, David, 2013. *Automatizace konverze datových formátů pro databázový systém*. Bakalářská práce. Technická univerzita v Liberci.

- KREJBICH, David, 2015. *Využití nástrojů business intelligence pro hodnocení kvality přírodních vod*. Diplomová práce. Technická univerzita v Liberci.
- KREYE, Melanie E., Yee Mey GOH a Linda B. NEWNES, 2011. Manifestation of uncertainty – a classification. *Proceedings of the 18th International Conference on Engineering Design (ICED 11): Impacting Society Through Engineering Design, Vol 6: Design Information and Knowledge*. Roč. 6, s. 96–107. ISSN 2220-4334.
- KUBRICHT, Jiří, 2008. *Distribuční sklad PHM společnosti Benzina a.s. Lokalita Šumperk-Vikýřovice : Analýza rizik starých ekologických zátěží*. Závěrečná zpráva. České Budějovice: ENVIRO C.B. s.r.o.
- KUHN, Thomas S., 1970. *The structure of scientific revolutions*. Second Edition, enlarged. University of Chicago Press. International encyclopedia of unified science. Foundations of the unity of science. ISBN 0-226-45803-2.
- LÁSKA, Jakub, 2011. *Import laboratorních rozborů do databáze EnviroInsite*. Semestrální projekt. Technická univerzita v Liberci.
- MATULA, S., G. B. MEKONNEN, K. BÁŤKOVÁ a K. NEŠETŘIL, 2014. Simulations of groundwater-surface water interaction and particle movement due to the effect of weir construction in the sub-watershed of the river Labe in the town of Děčín. *Environmental Monitoring and Assessment*. Roč. 186, č. 11, s. 7755–7770. ISSN 0167-6369, 1573-2959. DOI: 10.1007/s10661-014-3964-6
- MATULA, Svatopluk, Getu BEKERE, Kamil NEŠETŘIL a Kamila ŠPONGROVÁ, 2009. Hydrogeological modelling of different scenarios of groundwater movement and discharge in the watershed of the River Labe in North-west Bohemia. In: Pavel KOVAR, Petr MACA a Jana REDINOVA, ed. *Water Policy 2009, Water as a Vulnerable and Exhaustible Resource. 23 – 26 June 2009: Proceedings of the Joint Conference of APLU and ICA*. Prague: CULS Prague, s. 48–51. ISBN 978-80-213-1944-8. Dostupné z: <http://www.fzp.czu.cz/waterpolicy2009>
- NEŠETŘIL, Kamil, 2008a. *Aktualizovaná analýza rizika na lokalitě Luštěnice : Matematický model*. Závěrečná zpráva. Liberec: ET CONSULTING s.r.o.
- NEŠETŘIL, Kamil, 2008b. *DS PHM Benzina a.s. Šumperk-Vikýřovice Matematický model*. Závěrečná zpráva. Praha: AQUATEST a.s.
- NEŠETŘIL, Kamil, 2008c. *Jez u Děčína : Šíření kontaminace z náplavů : Kvantitativní hodnocení*. Závěrečná zpráva 253070498000. Praha: AQUATEST a.s.
- NEŠETŘIL, Kamil, 2009a. *Jez u Děčína : Šíření kontaminace z náplavů : Matematický model*. Závěrečná zpráva 253090036000. Praha: AQUATEST a.s.
- NEŠETŘIL, Kamil, 2009b. Modelování šíření a přirozené atenuace ropných uhlovodíků – praktická aplikace. In: Nad'a RAPANTOVÁ a Arnošt GRMELA, ed. *Voda – strategická surovina pro 21. století : Sborník 10. Česko-Slovenského mezinárodního hydrogeologického kongresu: 31. 8. – 3. 9. 2009*. VŠB – Technická univerzita Ostrava, Esmedia DTP s.r.o., s. 163–166. ISBN 978-80-248-2026-2.
- NEŠETŘIL, Kamil, 2012a. Einfache und umfassende Modelle des Grundwassers : Jednoduché a komplexní modely podzemní vody. In: *Fachseminar „Wissenschaft & Praxis“: Odborný seminář „Věda a praxe“ 2009–2012*. Zittau: Hochschule Zittau/Görlitz, s. 49. ISBN 978-3-9812655-7-6.
- NEŠETŘIL, Kamil, 2012b. *Matematický model podzemní vody: areál firmy KAR-BOX s.r.o. Hořice*. Závěrečná zpráva. Technická univerzita v Liberci.
- NEŠETŘIL, Kamil, 2014. Informační systém pro správu hydrogeologických dat. In: Josef V. DATEL, Jan NOVOTNÝ a Jindřiška HAUEROVÁ, ed. *Sborník příspěvků XIV. hydrogeologického kongresu Průzkum, využívání a ochrana podzemní vody: nové úkoly a výzvy. Sborník příspěvků II. inženýrsko-*

- geologického kongresu Role inženýrského geologa v současnosti*. Liberec: TUL, ČAH, ČAIG, USB: 3 strany, abstrakt na s. 83. ISBN 978-80-903635-4-0.
- NEŠETŘIL, Kamil a Jan ŠEMBERA, 2014. Groundwater data management system. In: Jorge Marx GÓMEZ, Michael SONNENSCHNEIN, Ute VOGEL, Andreas WINTER, Barbara RAPP a Nils GIESEN, ed. *EnviroInfo 2014 – ICT for Energy Efficiency : Proceedings of the 28th International conference on informatics for environmental protection. September 10–12, 2014, Oldenburg, Germany*. Oldenburg: BIS-Verlag, Carl von Ossietzky University Oldenburg, s. 301–306. ISBN 978-3-8142-2317-9. Dostupné z: enviroinfo2014.org, <http://oops.uni-oldenburg.de/id/eprint/1919>, <http://www.iai.kit.edu/ictensure/site?mod=litdb&subject=art&pid=X13287035&action=detail>
- POPPER, Karl, 1997. *Logika vědeckého bádání*. 1. vyd. Přel. Jiří FIALA. Praha: OIKOYMENH. Oikúmené. ISBN 978-80-86005-45-4.
- REFSGAARD, Jens Christian, Jeroen P. VAN DER SLUIJS, James BROWN a Peter VAN DER KEUR, 2006. A framework for dealing with uncertainty due to model structure error. *Advances in Water Resources*. Roč. 29, č. 11, s. 1586–1597. ISSN 0309-1708. DOI: 10.1016/j.advwatres.2005.11.013
- REFSGAARD, Jens Christian, Jeroen P. VAN DER SLUIJS, Anker Lajer HØJBERG a Peter A. VANROLLEGHEM, 2007. Uncertainty in the environmental modelling process – A framework and guidance. *Environmental Modelling & Software*. Roč. 22, č. 11, s. 1543–1556. ISSN 1364-8152. DOI: 10.1016/j.envsoft.2007.02.004
- SKOŘEPA, Jaroslav, Petr CHARVÁT, Markéta HRKALOVÁ, Zdeněk JEZERSKÝ, Lucia LENCSESOVÁ, Kamil NEŠETŘIL, Ondřej NOL, Aleš PACL, Věra PĚKNÁ, Ivan PERGLER a Tomáš VRÁNEK, 2009. *Hydrogeologický monitoring a posouzení pohybu podzemních vod na hranicích Polské, Německé a České republiky v povodí toků Horní Ploučnice, Nisy a Smědé, závěrečná zpráva 2008/2009*. Číslo úkolu: J241080220000. Praha: AQUATEST a.s.
- ŠTÍREK, Pavel, 2014. *Automatizace transformací dat*. Ročníkový projekt – PRJ2. Technická univerzita v Liberci.
- TOPINKOVA, Barbora, Kamil NESETRIL, Josef DATEL, Ondrej NOL a Petr HOSL, 2007. Geochemical heterogeneity and isotope geochemistry of natural attenuation processes in a gasoline-contaminated aquifer at the Hnevice site, Czech Republic. *Hydrogeology Journal*. Roč. 15, č. 5, s. 961–976. DOI: 10.1007/s10040-007-0179-8
- UUSITALO, Laura, Annukka LEHIKONEN, Inari HELLE a Kai MYRBERG, 2015. An overview of methods to evaluate uncertainty of deterministic models in decision support. *Environmental Modelling & Software*. Roč. 63, s. 24–31. ISSN 1364-8152. DOI: 10.1016/j.envsoft.2014.09.017
- WALKER, W.E., P. HARREMOËS, J. ROTMANS, J.P. VAN DER SLUIJS, M.B.A. VAN ASSELT, P. JANSSEN a M.P. KRAYER VON KRAUSS, 2003. Defining uncertainty: a conceptual basis for uncertainty management in model-based decision support. *Integrated Assessment*. Roč. 4, č. 1, s. 5–17. ISSN 1389-5176. DOI: 10.1076/iaij.4.1.5.16466
- WARMINK, J. J., J. A. E. B. JANSSEN, M. J. BOOIJ a M. S. KROL, 2010. Identification and classification of uncertainties in the application of environmental models. *Environmental Modelling & Software*. Roč. 25, č. 12, s. 1518–1527. ISSN 1364–8152. DOI: 10.1016/j.envsoft.2010.04.011