

# Matematické vyhodnocení kvality zdrojů pitné vody pro hlavní město Prahu

Ing. Radka Hušková, Ing. Jaroslava Palasová, Ing. Rudolf Pecinovský, CSc.  
Pražské vodovody a kanalizace a. s.

---

Kvalita pitné vody dodávané jednotlivými úpravkami pro hlavní město Prahu je průběžně kontrolována podle ČSN 75 7111 *Pitná voda* (dále jen *Norma*). Ta stanovuje více než 100 ukazatelů, které je třeba průběžně měřit, a zároveň definuje hodnoty, které by neměly být překročeny, resp. podkročeny. Z tohoto množství singulárních dat si však lze těžko udělat celkový obraz o kvalitě vody z úpraven a na jejich základě je téměř nemožné prohlašovat cokoli o dlouhodobých trendech vývoje kvality vody dodávané vyhodnocovaným vodním zdrojem nebo porovnávat mezi sebou kvality několika zdrojů. Pokusili jsme se proto najít charakteristiku, která by vycházela z výše jmenované *Normy*, avšak počet hodnot charakterizujících kvalitu testované vody by byl natolik malý, aby umožňoval jednoduše a rychle vzájemně porovnat kvalitu vody z různých zdrojů nebo kvalitu vody v různých obdobích.

Předkládané řešení bylo konzultováno s hygieniky a na základě jejich doporučení jsme přidali některá další kritéria hodnocení výsledných údajů, aby získaná data bylo možno využít i s přihlédnutím na hygienická hlediska.

V tomto případě je cílem matematického vyhodnocení kvalita vody na výstupu z úpraven Želivka, Káraný a Podolí. Tyto úpravny jsou zdrojem pitné vody pro hlavní město Prahu. Jejich podíl na celkové výrobě pitné vody je následující: Želivka 62 %, Káraný 25,5 %, Podolí 12,5 %.

## Charakteristika problému

Při vypracování příslušného postupu bylo nutno vzít v úvahu následující omezení:

- Ve výsledku by se měly projevit pokud možno všechny *Normou* stanovené ukazatele a případně i některé další měřené ukazatele.
- Hodnoty jednotlivých ukazatelů se řádově liší.
- Ukazatelé jsou různých typů: u většiny se hlídá povolené maximum, u jiných povolené minimum, další se musejí pohybovat ve vymezeném rozsahu.
- Různé ukazatele jsou pro posouzení kvality vody různě důležité. Tato důležitost je dána v *Normě* charakterem sledovaného limitu (Nejvyšší mezná hodnota, Mezná hodnota, Indikační hodnota, Doporučená hodnota).

## Výběr vyhodnocovaných ukazatelů

Při zpracování dat bylo rozhodnuto zahrnout do výsledku naměřené hodnoty všech ukazatelů s následujícími výjimkami:

- Ukazatele, jejichž hodnoty nemají exaktní podklad a jejichž stanovení je do značné míry subjektivní (chuť, pach).
- Ukazatele, jejichž požadované hodnoty se řídí okamžitou potřebou na výstupu z úpraveny (celkový aktivní chlor)
- Ukazatele, jejichž hodnota se může cestou ke spotřebiteli výrazně změnit (teplota, rozpuštěný kyslík).

Sada ukazatelů specifikovaných normou je velmi široká a obsahuje řadu ukazatelů, jejichž hodnoty jsou trvale velice nízké. Tím vyhlazují výsledné průměry, takže případné rozdíly jsou méně patrné. Po konzultaci s hygieniky byla proto definována vybraná podmnožina ukazatelů, které jsou z hygienického hlediska považovány pro vyhodnocování

kvality vody za nejdůležitější. Tato podmnožina by měla zároveň lépe odrážet případné fluktuační kvality vody.

### Eliminace řádových rozdílů hodnot – normalizace

Jedním z problémů, s nimiž jsme se museli vypořádat, byl velký rozsah naměřených hodnot. Tyto hodnoty se totiž pohybují u některých ukazatelů v řádu stotisícin a u jiných v řádu stovek. Při posuzování kvality upravené vody se nezajímáme o vlastní naměřené hodnoty, ale daleko více nás zajímá to, zda naměřené hodnoty splňují *Normou* předepsané limity a s jakou rezervou tyto limity splňují či nakolik je překračují. Rozhodli jsme se proto nepočítat v další analýze s naměřenými hodnotami, ale převést je nejprve na tzv. normalizované hodnoty, které nám lépe umožní sledovat vztah naměřených hodnot k hodnotám, jejichž dodržení je předepsáno *Normou*.

**Normalizovaná hodnota** je bezrozměrné číslo vyjadřující dodržení nebo nedodržení *Normou* předepsaných limitů a případnou rezervu jejich splnění resp. míru jejich překročení. Platí pro ní následující:

- Je vždy **nezáporná**.
- Dosáhne-li naměřená hodnota právě mezní hodnoty dané *Normou*, bude normalizovaná hodnota měření právě rovna 1 (jedné).
- Bude-li naměřena **lepší** hodnota, než stanovuje norma (tj. u ukazatelů s hlídaným maximem hodnota nižší než toto maximum a u ukazatelů s hlídaným minimem hodnota vyšší než toto minimum), bude normalizovaná hodnota **menší než 1**.
- Bude-li naměřena **horší** hodnota, než stanovuje *Norma*, bude normalizovaná hodnota **větší než 1**.
- Čím dále bude změřená hodnota dále od kritické hodnoty uvedené v *Normě*, tím dále bude normalizovaná hodnota od jedničky.

Principy naznačené v předchozích odstavcích můžeme pro většinu ukazatelů splnit takovou definicí normalizace, při níž bude platit:

- mezi normalizovanou hodnotou a hodnotou ukazatele s hlídaným maximem platí přímá úměra,
- mezi normalizovanou hodnotou a hodnotou ukazatele s hlídaným minimem platí úměra nepřímá.

Zvláštní povahu mají čtyři mikrobiologické a biologické ukazatele, u nichž je nejvyšší přípustnou a zároveň i nejmenší možnou hodnotou nula. Vzhledem k jejich zvláštní povaze a výjimečné důležitosti bylo rozhodnuto vyhodnocovat tyto ukazatele zvlášť.

Vzhledem k nulovosti nejvyšší povolené hodnoty není u těchto ukazatelů možno splnit podmínku jednotkovosti normy nejvyšší povolené hodnoty pomocí přímé úměry. Protože jsme zároveň požadovali, aby byla výrazně detekována i malá překročení a na druhou stranu jsme nechtěli, aby nejvyšší hodnoty normy příliš rostly, byla pro tyto ukazatele nahrazena přímá úměra speciální logaritmickou závislostí.

Mezi sledovanými ukazateli jsou dva (Ca+Mg a pH), u kterých je požadováno udržování hodnot v předem daném rozsahu. Jednoduchý normalizační vztah vede k normalizovaným hodnotám, které se v nejlepším případě pohybují těsně pod 1. Bylo by sice možno definovat nějaké rafinovanější normalizační pravidlo, avšak výsledek by byl neúměrně citlivý na zvolenou normalizační funkci a nastavení jejích parametrů. Bylo proto rozhodnuto žádný složitější normalizační vzorec nevytvářet a tyto dva ukazatele nenormalizovat.

### Zakomponování rozdílné důležitosti ukazatelů – supernorma

Různé ukazatele jsou různě důležité. Abychom nesměšovali údaje o životně důležitých ukazatelích s údaji o ukazatelích, které jsou spíše orientační, rozdělili jsme je do několika skupin, které korespondují s ČSN 75 7111:

- Mikrobiologické a biologické ukazatele s nulovou nejvyšší povolenou hodnotou (4 ukazatele).
- Ukazatele s hlídaným (nenulovým) maximem – ty byly rozděleny ještě do čtyř podtříd:
  - Látky s přímým toxickým účinkem (35 ukazatelů).
  - Toxické látky a mikrobiologické a biologické ukazatele s nulovou nejvyšší povolenou hodnotou (13 ukazatelů).
  - Smyslově postižitelné látky (13 ukazatelů).
  - Ostatní (10 ukazatelů).
- Ukazatele s hlídaným minimem (5 ukazatelů).
- Ukazatele s hlídaným rozsahem (2 ukazatele, jejichž hodnoty se nenormalizují).

Abychom mohli v dalším kroku počet výstupních údajů dále zmenšit a aby se přitom do výsledné hodnoty více promítly hodnoty důležitějších ukazatelů, byly jednotlivým ukazatelům přiřazeny váhy, kterými se normalizovaná hodnota umocní. Tím dosáhneme toho, že se u důležitých ukazatelů vliv jejich odchylky od *Normou* definované sledované hodnoty zvýrazní. Takto obdrženu úpravu normalizované hodnoty označujeme jako **supernormu**. Váhy byly jednotlivým ukazatelům přiřazeny podle následujícího klíče:

- Ukazatele, pro něž je v *Normě* definována *Mezná hodnota přijatelného rizika* nebo *Nejvyšší mezná hodnota* mají přiřazenu váhu 2.
- Ukazatele, pro něž je v *Normě* definována *Maximální mezná hodnota* mají přiřazenu váhu 1,5.
- Ukazatele, pro něž je v *Normě* definována *Doporučená hodnota* (a to jak minimální, tak maximální) nebo *Indikační hodnota* mají přiřazenu váhu 1.
- Vodivosti byla přiřazena váha 1,5, přestože *Norma* označuje mezní hodnotu jako indikační. Učinili jsme tak proto, že překročení této hodnoty většinou signalizuje vážnější výkyv v kvalitě vody.
- Mikrobiologickým a biologickým ukazatelům s nulovou nejvyšší povolenou hodnotou (4 ukazatele) byla přiřazena váha 1, protože tyto ukazatele jsou vzhledem ke své důležitosti vyhodnocovány samostatně, takže bychom větší vahou stejně nic nezískali.

### Histogramové kontingenční tabulky

U každého ukazatele vždy spočteme minimální, maximální a průměrnou naměřenou hodnotu, kterou následně normalizujeme, resp. supernormalizujeme. Pro výpočet průměru, minima a maxima jsme zvolili období 1 rok. Veškeré zpracovávané údaje byly zkontrolovány použitím různých testů pro extrémně odlehlá čísla případně dohledáním surových dat. Z takto získaných údajů vytvoříme tzv. *histogramové kontingenční tabulky*, z nichž je možno snadno vyčíst rozložení normalizovaných hodnot pro danou úpravu a danou skupinu ukazatelů.

Řádky této tabulky jsou vyhrazeny jednotlivým zdrojům vody, resp. jednotlivým rokům, sloupce jsou pak vyhrazeny intervalům normalizovaných, resp. supernormalizovaných hodnot. Do buněk tabulky se pak zapisuje kolik normalizovaných, resp. supernormalizovaných hodnot stanovení ukazatelů příslušné skupiny padlo do daného intervalu. Jedno z možných uspořádání kontingenční tabulky vidíte na následujícím obrázku.

		Průměry supernormy všech ukazatelů s hlídaným minimem či maximem																									
TDůl	(Vše)																										
TNX	(Vše)																										
počet z SNPrům																											
NázevVodámy	Rok	0,0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1,0	1,1	1,2	1,3	1,4	1,5	1,6	1,7	1,8	1,9	2,0	2,1	2,2	2,3	2,4	2,5
Podolí	1995	45	10	4	3	1	1	2	3		2	1		1													
	1996	50	10	4		4	1		2	3	1	2	1														
	1997	51	9	5	3	1			1	3		2	1														
	1998	52	9	3	3	1	2	1		1	2	1	1														
	1999	52	10	4		1	2	1		2		1	1	2													
Podolí počet		250	48	20	9	8	6	4	6	7	5	7	3	4													
Želivka	1995	48	9	2	2	2	1	1	1	3	3														1		
	1996	53	9	3			2	1		3	3	1														1	
	1997	54	10	1		1	1	1	1	2	2	2										1					
	1998	54	10	1		2	1	1	2		2	2							1								
	1999	54	10	1		1	1	3	1		2	2										1					
Želivka počet		263	48	8	2	6	6	7	5	8	12	7							1			2		1	1		
Káraný I+II	1995	50	8	3	4	2		1	1	1	1					1		1									
	1996	54	8	5	2	1	2			1		1		1			1										
	1997	55	11	1	3	1			1		1			2	1												
	1998	56	8	3	2	1	1				1	1		1	1		1										
	1999	56	8	4	1		2	1			1				1		2										
Káraný I+II počet		271	43	16	12	4	6	2	2	3	4	1	1	5	1	2	3	1									
Káraný III	1995	51	6	4	3	3	1		2		1				1					1							
	1996	53	8	5	1	3	2			1		1		1					1		1						
	1997	55	9	3	3		1	1		1			1	1	1												
	1998	55	9	2	3	1	1		1		1		1	1	1												
	1999	57	7	4	1	1		2			1			1		2											
Káraný III počet		271	39	18	11	8	5	3	3	1	4	2	1	4	1	4			1	1							
	Počet z 1995	194	33	13	12	8	3	4	7	4	7	1		1	1	1		1	1					1			
	Počet z 1996	210	35	17	3	8	7	1	2	6	4	3	2	1	2		1	1						1			
	Počet z 1997	215	39	10	9	2	3	2	3	5	4	4	1	1	3	1	1					1					
	Počet z 1998	217	36	9	8	5	5	2	3	2	6	3	2	1	2		2	1									
	Počet z 1999	219	35	13	2	3	5	7	1	2	4	3	1	2	2		2	2				1					

### Statistické charakteristiky supernormalizovaných hodnot

Histogramové kontingenční tabulky jsou sice velice informativní charakteristikou jednotlivých zdrojů pitné vody pro Prahu, avšak dají se těžko jednoduše popsat a ještě obtížněji se dají využít při popisu různých trendů – např. chceme-li ukázat, jak se mění celková sledovaná kvalita vody v posledních několika letech nebo jak se mění v průběhu jednoho roku. Pro stručnou charakteristiku kvality pitné vody několika málo údaji byl zvolen průměr supernorem. Kvalita pitné vody na výstupu z úpravní je pak charakterizována dvojicí či trojicí čísel:

- průměrem supernorem průměrů mikrobiologických a biologických ukazatelů s hlídaným nulovým maximem (hodnota supernormy je pro tyto ukazatele shodná s hodnotou normy) a
- průměrem supernorem průměrů ukazatelů s hlídaným minimem či maximem nezařazených do předchozí skupiny,
- případně také průměrem supernorem průměrů ukazatelů vybraných hygieniky.

### Biologické a mikrobiologické ukazatele s nulovou nejvyšší povolenou hodnotou

	Rok				
	1995	1996	1997	1998	1999
Vodárna					
Podolí	1,4	1,4	1,425	1,275	1,
Želivka	1,	1,	1,	1,	1,
Káraný I+II	1,25	1,35	2,1	1,925	1,65
Káraný III	1,15	1,55	1,925	1,7	2,

### Chemické ukazatele s jednou hlídanou mezí

#### Všechna stanovení

	Rok				
	1995	1996	1997	1998	1999
Vodárna					
Podolí	0,149	0,134	0,124	0,12	0,13
Želivka	0,162	0,151	0,142	0,132	0,137
Káraný I+II	0,132	0,113	0,109	0,113	0,117
Káraný III	0,13	0,121	0,111	0,113	0,113

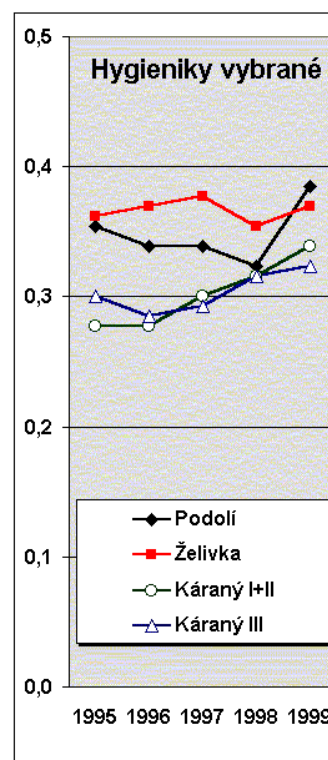
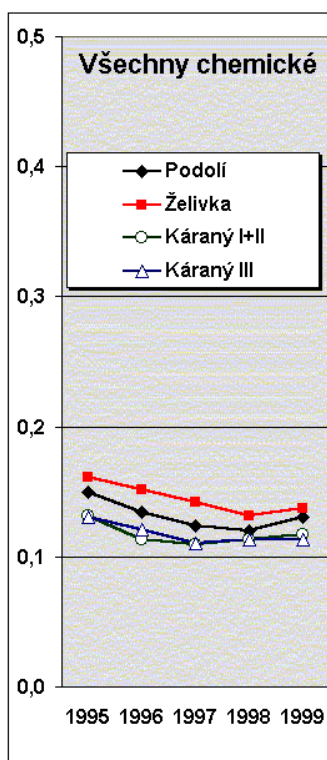
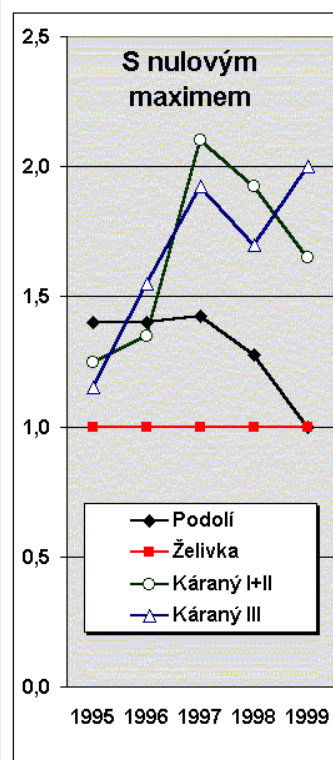
#### Stanovení vybraná hygieniky

	Rok				
	1995	1996	1997	1998	1999
Vodárna					
Podolí	0,354	0,338	0,338	0,323	0,385
Želivka	0,362	0,369	0,377	0,354	0,369
Káraný I+II	0,277	0,277	0,3	0,315	0,338
Káraný III	0,3	0,285	0,292	0,315	0,323

Tuto trojici lze samozřejmě doplnit trojicemi, v nichž se místo průměrů supernorem průměrů počítá průměr supernorem minim či průměr supernorem maxim.

Uvedené tabulky je samozřejmě možné velice snadno vyjádřit graficky.

### Grafy meziročního průběhu průměru supernorem



### Hrubý nástin řešení

Celý proces u nás probíhá následovně:

1. Každá laboratoř, tj. provozní laboratoře vodáren a laboratoř speciálních analýz a kontroly, dodá všechna svoje naměřená data v souboru zformátovaném definovaným způsobem.

2. Data se centrálně zpracují, doplní se sadou pomocných údajů (např. charakteristiky jednotlivých ukazatelů) a vytvoří se z nich jedna databáze.
3. Data v této databázi se otestují a podezřelé hodnoty se vrátí zdrojovým laboratořím k přezkoumání, zda se jedná o překlep nebo o skutečně naměřenou hodnotu.
4. Po eliminaci podezřelých hodnot se z dat vypočtou minima, průměry a maxima hodnot jednotlivých ukazatelů. Takto získané hodnoty se normalizují a supernormalizují.
5. Vytvoří se histogramové kontingenční tabulky normalizovaných a supernormalizovaných hodnot.
6. Vypočtou se statistické charakteristiky supernormalizovaných hodnot a na jejich podkladě se zkonstruuji příslušné grafy.

## Závěr

Byl vypracován postup matematického zpracování velkého objemu naměřených dat týkajících se kvality pitné vody z jednotlivých úpraven PVK a.s. Data jsou zpracována tak, aby ve výsledném zjednodušení byly zahrnuty veškeré vybrané naměřené údaje související s kvalitou vyráběné vody. Každá úprava je tak v každém roce porovnávaného období 1995 – 1999 charakterizována třemi trojicemi čísel vycházejících z průměrů supernorem minim, průměrů a maxim:

- mikrobiologické a biologické ukazatele s povoleným nulovým maximem,
- chemické ukazatele spolu s mikrobiologickými a biologickými ukazateli s vyšším než nulovým povoleným maximem,
- podmnožina ukazatelů vybranou podle doporučení hygieniků.

Výsledky jsou zpracovány graficky. Pro detailnější charakteristiku úpravy pak slouží kontingenční histogramové tabulky, které názorně ukazují rozložení jednotlivých normalizovaných, resp. supernormalizovaných hodnot.

Závěr zpracování údajů je jednoznačný a srozumitelný. Takto lze jednoduchým způsobem vyjádřit míru kvality vyráběné vody, kde jsou v několika málo číslech poměrně spravedlivě zkoncentrovány desetitisíce údajů týkajících se sledování kvality vyráběné vody (po ustálení specifikace vstupních dat je v současné době zpracováváno výše uvedeným způsobem téměř 50 000 údajů ročně).

Tento náhled na kvalitu vody, který je výsledkem výše popsaného matematického zpracování, bychom chtěli využít i pro sledování kvality vody v průběhu roku, aby bylo možno postihnout i sezónní nebo jiné výkyvy v kvalitě vyráběné vody.