

Certifikovaná metodika pro provádění dlouhodobé prognózy produkce odpadů v ČR včetně revize prognózy (Výsledek V9).

15. 10. 2021

Konečný uživatel výsledků: **Ministerstvo životního prostředí**
Vršovická 1442/65
Praha 10, 100 10

Název projektu: Prognózování produkce odpadů a stanovení složení komunálního odpadu**Číslo projektu:** TIRSMZP719

Řešitel projektu: Vysoké učení technické v Brně
Fakulta strojního inženýrství,
Ústav procesního inženýrství
Antonínská 548/1, 60190 Brno

Doba řešení: 1. 1. 2019 – 31. 12. 2021

Informace o autorském týmu:

Ústav procesního inženýrství, Fakulta strojního inženýrství

Vysoké učení technické v Brně, Technická 2896/2, Brno

www.upi.fme.vutbr.cz



Hlavní řešitel projektu: doc. Ing. Martin Pavlas, Ph.D.

Garant výsledku: Ing. Radovan Šomplák, Ph.D.

Autorský kolektiv: Ing. Veronika Smejkalová, Ing. Martin Rosecký, Ing. Lenka Szásziová, Ph.D., Ing. Kristína Šramková, Ing. Jakub Kúdela, Ph.D., Ing. Ivan Eryganov, Ing. Zbyněk Bouda, RNDr. Jana Suzová, RNDr. Pavel Popela, Ph.D.

Garant MŽP: Ing. Gabriela Bulková

Další informace o projektu:

Cílem projektu veřejné zakázky je identifikovat, popsat a nastavit proces prognózování produkce všech odpadů v ČR ve výhledu až 10 let a současně nastavit postupy pro stanovení složení směsného komunálního odpadu z obcí a komunálního odpadu.

Program veřejných zakázek v aplikovaném výzkumu a inovacích pro potřeby státní správy BETA2 byl schválen usnesením vlády České republiky č. 278 ze dne 30. 3. 2016 a je zaměřen na podporu aplikovaného výzkumu a inovací pro potřeby orgánů státní správy. Poskytovatelem finančních prostředků je Technologická agentura ČR.

Obsah

SEZNAM OBRÁZKŮ	4
SEZNAM TABULEK	4
SEZNAM ZKRATEK.....	5
1. PŘEDMLUVA	6
2. ÚČEL METODIKY	6
3. CÍL METODIKY	6
4. VÝCHODISKA.....	7
5. ÚVOD DO METODIKY, ZÁKLADNÍ POJMY A POŽADAVKY NA PROGNÓZU.....	9
6. NÁSTROJE A PŘEDPOKLADY	14
6.1. VÝPOČETNÍ TECHNIKA	14
6.2. DATABÁZE	14
7. SLEDOVANÉ UKAZATELE	14
7.1. VOLBA DETAILU ÚZEMNÍHO ČLENĚNÍ	15
7.2. VOLBA PŘEDMĚTNÉ FRAKCE ODPADŮ	15
7.3. VOLBA OBDOBÍ PROGNÓZY	15
7.4. NASTAVENÍ CÍLŮ PRO SCÉNÁŘOVÁ ŘEŠENÍ	15
8. DÍLČÍ ČÁSTI METODIKY	16
8.1. PŘÍPRAVA VSTUPŮ.....	17
8.1.1. Shromáždění a příprava vstupních dat	17
8.1.2. Analýza specifíků dat	18
8.1.3. Agregace dat.....	19
8.2. PRE-PROCESSING DAT	20
8.2.1. Analýza odlehlých hodnot v datech	20
8.2.2. Analýza skoků v datech a změn trendu.....	21
8.3. PROCESSING DAT	22
8.3.1. Odhad vývoje vlivných faktorů	22
8.3.2. Odhad vývoje produkce odpadu	22
8.3.3. Vyrovnávání dat	25
8.3.4. Intervalové odhady	26
8.4. POST-PROCESSING DAT – PROJEKCE (SCÉNÁŘOVÁ ŘEŠENÍ).....	26
9. ZPRACOVÁNÍ A INTERPRETACE VÝSLEDKŮ	27
10. REVIZE POUŽITÝCH PŘÍSTUPŮ V METODICE A KONTROLA VÝSLEDKŮ PROGNÓZY	27
11. ZÁVĚREČNÁ ZPRÁVA	28
12. SEZNAM PŘÍLOH.....	29
REFERENCE	30

Seznam obrázků

Obr. 1: Ukázka nesouladu při různém provedení agregace dat a odhadu budoucích hodnot pomocí trendu	12
Obr. 2: Schematické znázornění metodiky prognózování produkce odpadu	13
Obr. 3: Dostupné metody pro odhad vývoje produkce odpadu analyzované v projektu TIRSMZP719	23

Seznam tabulek

Tab. 1: Běžně užívané metody pro prognózování produkce odpadu	8
Tab. 2: Jednotlivé kroky postupu metodiky prognózování a reference na detailnější informace	16

Seznam zkratk

Zkratka	Význam
BAU	business-as-usual
CENIA	Česká informační agentura životního prostředí
ČR	Česká republika
ISOH	Informační systém odpadového hospodářství
kat. č.	katalogové číslo odpadu
KO	komunální odpad
MŽP ČR	Ministerstvo životního prostředí České republiky
OH	odpadové hospodářství
ORP	obec s rozšířenou působností
POH ČR	Plán odpadového hospodářství České republiky
SKO	směsný komunální odpad
TIRSMZP719	projekt Prognózování produkce odpadů a stanovení složení komunálního odpadu
ZÚJ	základní územní jednotka

1. Předmluva

Tento dokument předkládá certifikovanou metodiku pro provádění dlouhodobé prognózy produkce odpadů v ČR. Autoři metodiky děkují Mgr. et Mgr. Jiřímu Kalinovi, Ph.D. za průběžnou diskuzi nad vývojem metodiky z pozice experta TAČR a Ing. Gabriele Bulkové za konstruktivní komentáře ze strany zadavatele projektu. *Certifikovaná metodika pro provádění dlouhodobé prognózy produkce odpadů v ČR včetně revize prognózy* byla podrobena dvěma oponentským posudkům. Autoři metodiky chtějí poděkovat oponentům za důkladné prostudování dokumentů a velmi cenné komentáře a doporučení, které byly ve většině případů do metodiky promítnuty. Jmenovitě se jedná o doc. Ing. Jana Slavíka, Ph.D. (Fakulta sociálně ekonomická, Univerzita J. E. Purkyně v Ústí nad Labem) a doc. Ing. Marka Omelku, Ph.D. (Matematicko-fyzikální fakulta, Univerzita Karlova v Praze).

Struktura textu je koncipována tak, že nejprve je nastíněn účel metodiky (kap. 2), následuje cíl metodiky (kap. 3) a východiska metodiky (kap. 4). Kap. 5 shrnuje základní principy metodiky a používané základní pojmy. Kap. 6 poskytuje výčet potřebných nástrojů a předpokladů k provedení prognózy. Následně je pozornost věnována sledovaným ukazatelům v kap. 7. Kap. 8 pak představuje dílčí části samotné metodiky zahrnující přípravu vstupů, pre-processing, processing a post-processing dat. Zpracování a interpretaci získaných výsledků prognózy se věnuje kap. 9. Podmínkami použití metodiky v budoucnu a s tím spojenou revizí metodiky se zabývá kap. 10, následuje závěr (kap. 11). Součástí tohoto dokumentu jsou přílohy: (Příloha 1 až Příloha 9), které představují aplikaci metodiky v rámci projektu TIRSMZP719. Součástí příloh je tedy detailnější popis k použitým matematickým metodám a jejich výběru s ohledem na konkrétní aplikaci.

Předem je třeba říct, že z hlediska matematické statistiky je predikce budoucího vývoje vždy velmi obtížná (a často do značné míry neřešitelná) úloha. I nejsofistikovanější predikční modely v sobě vždy skrývají předpoklad, že se veličiny budou vyvíjet obdobně jako v minulosti. I krátkodobé prognózy je zapotřebí brát velmi opatrně a dlouhodobější prognózy spíše orientačně ve smyslu: jak by vývoj produkce odpadu mohl přibližně vypadat, pokud by se nic nezměnilo (např. změnou legislativy). V aplikaci na produkci odpadů je problematika navíc komplikovaná tím, že k dispozici jsou často jen velmi krátké časové řady. Z tohoto důvodu autoři metodiky doporučují vždy s novými daty prognózu aktualizovat a metodiku predikce pravidelně revidovat.

2. Účel metodiky

Certifikovaná metodika pro provádění dlouhodobé prognózy produkce odpadů v ČR byla vytvořena pro podporu hodnocení současného stavu (aktuálního trendu) a odhadu dalšího vývoje odpadového hospodářství (OH) v ČR. Představená metodika stanovuje postup pro prognózu produkce všech odpadových frakcí. Metodika byla využita při tvorbě software *TiramisO* spravovaného Ministerstvem životního prostředí (MŽP ČR). Software *TiramisO* slouží k vytvoření prognózy a scénářů (projekcí) produkce odpadů. *TiramisO* bude využíván MŽP ČR ke zpracování a hodnocení strategických dokumentů v oblasti OH ČR zejména Plánu odpadového hospodářství ČR (POH ČR). Metodika, popř. software *TiramisO*, bude dále k dispozici pro všechny subjekty, které mají zájem prognózu provádět a využívat. Je vhodné, aby s metodikou pracovaly subjekty, které mají patřičnou požadovanou erudici v relevantních oborech (viz Příloha 9).

3. Cíl metodiky

Cílem metodiky je sjednocení postupů používaných v ČR pro provádění dlouhodobých prognóz (např. na dalších 15 let) produkce odpadů v ČR. Certifikovaná metodika zajistí adekvátní přístup k prognózování u všech aktérů v OH na různých úrovních (národní, regionální apod.) při tvorbě dlouhodobých prognóz produkce

odpadů. Metodiku bude možné využít při zpracování a hodnocení strategických dokumentů v oblasti OH ČR zejm. POH ČR a dalších. Je tedy určena pro všechny subjekty zapojené do oblasti plánování v OH (MŽP ČR, kraje apod.).

Využití metodiky je považováno za záruku, že informace získané při prognózování mají relevantní a kvalitativní úroveň. Tato relevantnost a kvalita bere v potaz problematiku přesnosti prognózy, kdy může být horizont požadované prognózy větší než dostupná datová základna (např. prognóza na 15 let za použití deseti let historických údajů). V této situaci je zapotřebí obezřetnost a pečlivé dodržování relevantních matematicko-statistických postupů.

Koncepce metodiky je obecné povahy a sumarizuje požadavky na prognózu produkce odpadu. Metodika stanovuje obecný postup jednotlivých navazujících kroků, které by měla každá prognóza obsahovat, případně by mělo být diskutováno vynechání některého kroku v řešené úloze. Součástí metodiky jsou přílohy k textu (Příloha 1 až Příloha 9), které popisují konkrétní aplikaci metodiky stanovenou zadavatelem požadované prognózy, tj. MŽP ČR. Tato aplikace je také implementována do softwarového prostředí *TiramisO*, které je ve vlastnictví zadavatele projektu TIRSMZP719 – MŽP ČR. Softwarem bude disponovat MŽP ČR a předpokládá se poskytnutí určité veřejné varianty k využití odbornou veřejností. V pravidelných intervalech je nutno provést revizi navržených matematicko-statistických postupů implementovaných v softwaru, neboť datová základna se bude rozšiřovat o nová data a mohou se objevit nové vazby v systému či nové důležité odpadové frakce. Návrh této revize je součástí předložené metodiky (viz kap. 10). Dílčím cílem jsou i doporučení autorů metodiky k jednotlivým krokům prognózy.

Při vytváření prognózy produkce je nutné dodržovat postupy stanovené touto metodikou. Přílohy k tomuto textu (Příloha 1 až Příloha 9) slouží jako ukázka aplikace metodiky na konkrétní úloze, kde jsou k dispozici dílčí analýzy jednotlivých kroků.

4. Východiska

V kontextu projektu TIRSMZP719, Výsledek V9 – *Certifikovaná metodika pro provádění dlouhodobé prognózy produkce odpadů v ČR včetně revize prognózy* – navazuje na přístupy a postupy v oblasti prognózování produkce odpadů představené v předchozích Výsledcích V5 až V8 projektu TIRSMZP719. Hlavním východiskem pro tvorbu této metodiky je dokument *Impact Assessment* při Evropské komisi, tzv. *European Reference Model on Municipal Waste Management* (Economia Research & Consulting et al, 2014). V tomto dokumentu je uvedena jednotná metodika EU zahrnující modelování produkce odpadu. Důležitou informací, která byla zmíněna v *Annex 8* tohoto dokumentu je, že do celkové koncepce modelu byly pro evropské státy zahrnuty prognózy produkce odpadu, které si jednotlivé státy tvoří samy. Pokud není k dispozici prognóza pro konkrétní stát (příp. metodika pro její vytvoření), je možné provést výpočet prognózy dle doporučení v *Annex 8* podle metodologie představené v dokumentu (ETC/EEA, 2012). Tentýž model byl představen v příspěvku (Andersen et al., 2007). Projekt TIRSMZP719 svými výstupy reaguje na neexistenci metodiky prognózování produkce odpadu v podmínkách ČR.

Přístup k řešení prognózování produkce všech odpadů v mezinárodním měřítku v podobě rešerše (pro ČR i zahraničí) byl popsán ve (Výsledek V5, 2019). Zde jsou zmíněny stěžejní rešeršní práce v této oblasti a následuje výčet užívaných metod podpořených citacemi na články, kde byly aplikovány (tab. 1). Literární rešerši na téma modelování produkce komunálního odpadu (KO) nabízí práce (Beigl et al., 2008), kde bylo studováno 45 modelovacích přístupů pro odhad současné nebo budoucí produkce KO s využitím ekonomických, socio-demografických a jiných dat. V rešerši (Goel et al., 2017) bylo studováno 106 publikací z období 1972–2016. Výsledky ukazují, že je stále větší pozornost věnována metodám strojového učení pro sestavení modelu. Autoři rešerše (Goel et al., 2017) zdůrazňují, že použitelnost modelů může být významně

omezena z důvodu nedostatku potřebných vstupních dat. Pro prognózování produkce odpadu se ve většině případů v dostupné literatuře využívá některá z následujících metod, viz tab. 1. Při volbě vhodného přístupu je nutné rozlišovat prediktivní a prognostické modely. Cílem prediktivních přístupů je popsat vazby systému a na základě těchto vazeb odhadnout současnou produkci odpadu pro zvolené území. Prognostické modely se zabývají odhadem budoucí produkce odpadu (cíl této metodiky). Délka predikčního horizontu a detail vstupních dat významně ovlivní volbu modelovací metody. Volba metody pro modelování je často limitována nedostatkem historických dat, což může být způsobeno změnou metodiky pro zpracování dat (např. shromažďování dat aktuálním způsobem v databázi ISOH – Informační systém odpadového hospodářství – v ČR probíhá od roku 2009), nebo shromažďováním neveřejných dat soukromými subjekty. Zásadní informací pro uživatele výsledků prognózy je vyjádření neurčitostí, obvykle v podobě konfidenčních nebo predikčních intervalů.

Tab. 1: Běžně užívané metody pro prognózování produkce odpadu

Metoda	Doporučené články
Vícerozměrná lineární regrese	Andersen and Larsen (2012), Hřebíček et al. (2017),
Zobecněné lineární modely	Karpušenkaitė et al. (2016),
Rozhodovací stromy	Johnson et al. (2017),
Umělé neuronové sítě	Noori et al. (2010),
Klasické modely analýzy časových řad	Petridis et al. (2016), Ghinea et al. (2016),
Strojové učení pro časové řady	Abbasi and Hanadeh (2016), Noori et al. (2009),
Scénářové přístupy	Dwivedy and Mittal (2010), Cole et al. (2014).

Rozsáhlá rešerše modelovacích přístupů zpracovaná ve (Výsledek V5, 2019) ukázala, že bylo vyvinuto značné množství modelů pro prognózování produkce odpadu. Aplikace jednotlivých modelů je limitována zejména charakterem dostupných dat a délkou predikčního horizontu. S ohledem na zaměření úlohy řešené v projektu TIRSMZP719 (viz Příloha 1 až Příloha 9) se jeví jako nejvhodnější přístup založený na klasických modelech analýzy časových řad. Nalezené přístupy ve Výsledku V5 však neposkytují požadované vlastnosti uvedené v požadavcích na prognózu (Příloha 2). Proto je vhodné vyvinout nový přístup pro tvorbu prognóz produkce odpadu v podmínkách ČR, který je představen v přílohách textu metodiky (Příloha 1 až Příloha 9).

Na výběr vhodných metod datové analýzy a matematického modelování prognózy produkce odpadů a také na celkový přístup k řešení projektu se zaměřil (Výsledek V6, 2019). Doporučení pro volbu modelovací metody je shrnuto v kap. 8.3.2, obr. 3. Ve (Výsledku V7, 2020) byl navržen a popsán postup řešení a matematický model prognózování produkce odpadů, včetně požadavků na výslednou prognózu a pre-processing dat. Výběru odpadových toků pro prognózování, nastavení okrajových podmínek výpočtů pro konkrétní úlohy (např. nastavení parametrů pre-processingu), návrhu výstupů a přístupu v oblasti adaptace modelu pro případ rozšíření datové sady aj. aspekty se pak věnoval (Výsledek V8, 2021). Výše uvedené dílčí výstupy jsou zpracovány ve Výsledku V9 (tento text) do certifikované metodiky pro provádění dlouhodobé prognózy produkce odpadů v ČR. Metodika popisuje obecný pohled na prognózování produkce odpadu s konkrétními doporučeními. V přílohách je detailně popsán postup pro potřeby prognózy řešené v projektu TIRSMZP719.

Prezentovaná metodika zohledňuje dostupná data, která mají pro každou konkrétní aplikaci metodiky celou řadu specifik a ovlivňují volbu matematicko-statistických metod. Důležitým východiskem metodiky jsou data z OH, která jsou aktuálně shromažďována v systému ISOH. Uživatel metodiky má možnost využít vlastní

zdroje dat z OH, pokud jsou k dispozici. V této oblasti je nutné vzít v potaz metodiku samotného shromažďování dat, délku dostupné časové řady i výhled do budoucna (rozšiřování datové základny). Dalším východiskem jsou data vlivných faktorů, kde rovněž velkou úlohu hraje způsob získávání dat s přímou vazbou na konzistentnost a relevanci výsledků. Nedílnou součástí prezentované metodiky je výběr vhodné matematicko-statistické metody, který je proveden na základě zpracované detailní rešerše (v nadnárodním měřítku) a na základě analýzy současných odpadových dat, viz obr. 3 v kap. 8.3.2. Výsledná metodika vychází z dílčích zkušeností lokálních studií, avšak představuje první systematickou metodiku pro provádění dlouhodobé prognózy produkce odpadů v ČR. Při sestavování modelů je nutné rozlišovat prognózy KO a průmyslových odpadů. Produkce KO je významně ovlivněna vývojem demografie. Dále je vhodné rozlišovat původce odpadu (obce, firmy a bez rozlišení původce) a nebezpečné vlastnosti odpadu (nebezpečný odpad, ostatní odpad). V případě obecního odpadu je možné demografický vývoj relativně snadno zahrnout především pro agregovaná data na krajské a státní úrovni. Naproti tomu je odpad vyprodukovaný ve firmách významně ovlivněn mobilitou pracovníků a typem pracovní činnosti. V případě průmyslových odpadů je tato vazba obtížně identifikovatelná z důvodu nerovnoměrné distribuce konkrétních typů odpadu závislé na konkrétní průmyslové činnosti. U průmyslových odpadů je u produkce odpadů častější výskyt změn adekvátní dynamice daného průmyslového odvětví. To se odráží i v četnosti anomálií v datech, které je možné identifikovat v rámci pre-processingu dat (kap. 8.2). Pro uživatele výsledků je nutné si uvědomovat vypovídací hodnotu prognóz např. v podobě intervalových odhadů.

Mezi kvalifikační předpoklady pro zpracovatele prognózy patří především erudice v oborech matematické statistiky (identifikace odlehklých pozorování, korelační analýza, regresní analýza, analýza časových řad) a operačního výzkumu (matematické programování – lineární, kvadratické a nelineární programování, vyrovnávání dat aj). Zpracovatel prognózy by měl disponovat vzhledem do sociálních, demografických a ekonomických aspektů. Stěžejní je obecný vhled do problematiky OH. V případě zaměření odhadu vývoje na specifické odpadové toky je pak nutná detailní znalost problematiky.

5. Úvod do metodiky, základní pojmy a požadavky na prognózu

Na základě této metodiky mohou vznikat různé prognózy produkce odpadu dle jejího účelu. Výsledná prognóza je vždy výstupem matematicko-statistického modelu o více částech. Každý vytvořený model by měl obsahovat následující části:

- bodový odhad,
- intervalový odhad,
- diagnostika modelu (ověření splnění předpokladů využitých metod),
- analýza citlivosti.

Tyto body jsou diskutovány pro konkrétní prognózu produkce odpadu řešenou v projektu TIRSMZP719 v přílohách k tomuto dokumentu (Příloha 1 až Příloha 9).

V této kapitole jsou vysvětleny základní pojmy, se kterými se pracuje v tomto dokumentu a které jsou předpokladem pro porozumění metodice. Jde však pouze o vysvětlení či upřesnění, které je nad rámec běžných znalostí statistiky se zaměřením na specifickou problematiku OH. Dále je představen úvod do metodiky spolu s požadavky na prognózu. Všechny níže zmíněné kroky pak jsou detailně popsány v dalších kapitolách.

5.1 Základní pojmy

Frakce odpadu

Prognóza produkce odpadu se zaměřuje na vybraná katalogová čísla (kat. č.) odpadu, a to buď individuálně, nebo v agregované podobě. Primárně je možno prognózu cílit na data a jejich agregace vycházející z *Katalogu odpadů* (Vyhláška č. 8/2021 Sb.), kde je definováno 879 kat. č. odpadu (označených šestimístním kódem), 111 podskupin odpadu (označených čtyřmístním kódem) a 20 skupin odpadu (označených dvoustupňovým kódem). Dále je možno odpad prognózovat také pro agregace kat. č., které vychází ze specifických požadavků zadavatele prognózy. Při dodržení všech postupů je možné se zabývat také prognózou složení jednotlivých kat. č. Frakcí odpadu se pak myslí odpad, ať už ve formě jednotlivého kat. č. nebo jejich agregace. Obecně (mimo tento text) je používán i termín „složka odpadu“, např. (Výsledek V4, 2020), který však reflektuje především fyzikální aspekty daného odpadu.

Vlivné faktory

Na produkci a složení odpadu mají vliv i některé faktory, které je vhodné zahrnout do modelu prognózy odpadu. Mezi nejčastější patří demografické, sociálně-ekonomické či environmentální vlivy. Zejména v případě demografických a sociálních dat je vazba očekávatelná u produkce KO. I data vlivných faktorů je nutné podrobit důkladnému pre-processingu dat.

Splnění územních bilancí

S ohledem na současné územní členění je vhodné uvažovat čtyři úrovně uspořádání, a to stát ČR, kraje, obce s rozšířenou působností (ORP) a obce (viz kap. 8.1.1). Dostupná data (např. o produkci odpadu) pro tyto jednotlivé úrovně berou v potaz hierarchickou strukturu. Tzn. např. množství produkovaného odpadu v ČR je rovno součtu daného typu odpadu vyprodukovaného ve všech krajích. Součet produkce odpadu v příslušných ORP odpovídá produkci v kraji, kterému ORP náleží atd. Splnění požadavku platnosti územních bilancí je vyžadováno i pro výslednou prognózu.

Splnění odpadových bilancí

Hierarchická struktura v datech o produkci odpadu vzniká také z důvodu agregace kat. č. odpadu (např. frakce popisující separovaný papír produkovaný obcemi se skládá z kat. č. odpadu 15 01 01 a 20 01 01, přičemž je nutno podotknout, že od roku 2021 došlo ke změně legislativy a podskupina 15 01 již nebude v této agregaci uvažována). V historických datech o produkci odpadu se tedy vlivem agregace dat vyskytují vazby mezi frakcemi odpadu. Např. pokud jsou sečtena všechna kat. č. definovaná pro KO, musí to být rovno výsledné produkci KO. Požadavek na splnění územních odpadových bilancí je pak i na výslednou prognózu. Tímto způsobem je vhodné uvažovat i o interakcích mezi jednotlivými typy odpadů.

Pre-processing

Při práci s časovými řadami je nutné se vypořádat s některými jejich vlastnostmi, které by mohly výsledný model negativně ovlivnit. Přestože data o produkci odpadu procházejí při každoročním hlášení víceetapovou kontrolou, není možné odstranit veškeré chyby v evidenci. V rámci pre-processingu dat je proto nutné věnovat dostatečnou pozornost nejen odlehilým hodnotám, ale také skokům v datech (náhlá změna hodnoty s navazujícím trendem, který není vždy stejný jako trend před skokem, viz kap. 8.2), změnám v trendu a dalším možným anomáliím a případně provést odbornou korekci. Vlivem pre-processingu dat dochází k tomu, že nejsou splněny územní a odpadové bilance.

Prognóza

Prognóza, v pojetí této metodiky, představuje nejpravděpodobnější scénář budoucího vývoje. Vychází z historických dat a není do něj začleněn (až na nutné výjimky) expertní aspekt, tj. změna trendu vlivem očekávaných zásahů do OH. Prognóza nemá schopnost reagovat na legislativní a jiné zásahy do systému, ke kterým v budoucnu dojde. Prognóza je dále označována, v rámci tvorby scénářů, jako business-as-usual (BAU) scénář (kap. 8.4). Vzhledem k modelování na základě historických dat jsou do prognózy promítnuty legislativní zásahy z minulosti. Je snaha jejich vliv eliminovat pre-processingem dat.

V kontextu této metodiky je také používán pojem „odhad vývoje“ vlivných faktorů či produkce odpadu (kap. 8.3.1). Na tento dílčí krok metodiky lze nahlížet rovněž jako na „dílčí“ prognózu, nicméně kvůli zajištění jasnosti konceptu metodiky není termín *prognóza* v tomto smyslu používán. Prognóza tedy označuje celou metodiku, a ne její dílčí část.

Přístup k prognóze uvedený v této metodice pracuje s následujícími předpoklady:

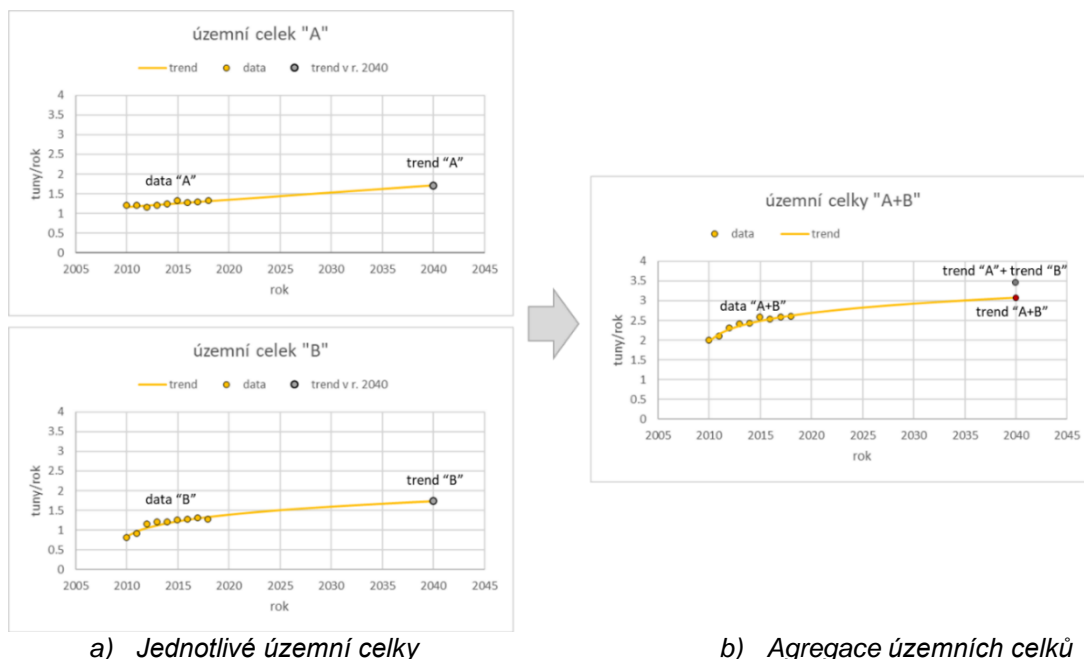
- Volba metody pro prognózu je ovlivněna charakterem dostupných dat a délkou predikčního horizontu.
- Do modelu jsou zahrnuty pouze takové vlivné faktory, které lze prognózovat s vyšší přesností, než je prognóza dat z OH bez těchto dat.
- V prognóze se předpokládá budoucí vývoj produkce odpadu za stávajících podmínek, nejsou tedy zohledněny budoucí změny v systému. U prognózy se tyto jevy projeví, až když se stanou součástí historických dat.
- Historická data odráží dopad vnitřních i vnějších vlivů realizovaných v minulosti (legislativní změny, cena druhotných surovin, technologický pokrok atd.) na produkci odpadu. Kvalitní pre-processing dat zajistí adekvátní reakci modelu na tyto vlivy, konkrétně se jedná o skoky v datech a změny trendu, viz kap. 8.2.2. Vědomé snížení počtu použitých historických dat s sebou automaticky nese dopad na kvalitu a přístup k prognóze (kap. 8.3.2).
- V prognózovaných hodnotách je zachována hierarchická struktura územních jednotek a frakcí odpadu.

Vyrovňávání dat

Splnění územních a odpadových bilancí je nutné zachovat v odhadech budoucí produkce. Obecně při odhadu vývoje není platnost těchto vazeb zajištěna. Funkce modelující trend v datech totiž nejsou obecně lineární, neplatí tedy aditivita. Platnost vazeb je dosažena s pomocí metody vyrovňávání dat, kdy dojde ke korekci odhadovaného vývoje v datech tak, aby platily zmíněné bilance.

Teoretický koncept vyrovňávání dat vychází z faktu, že odhady budoucích hodnot nejsou v logickém souladu, tzn. nesplňují základní vazby (např. územní a odpadové bilance, viz výše). U odhadnutých budoucích produkcí se předpokládá jejich zatížení náhodnou chybou (např. součet odhadovaných produkcí na nižších územních celcích se neshoduje s předpokládanou produkcí nadřazeného územního celku), která reflektuje proveditelnou přesnost odhadu vývoje, která se zhoršuje pro krátké časové řady.

Na obr. 1 je znázorněna ukázka, kdy je pro odhad vývoje použito nelineárních trendů a není tedy splněna územní bilance. V případě nelineárních trendů totiž záleží na pořadí agregace dat a modelu trendu. Odhad vývoje produkce odpadu by měl dojít k souladu, ať už je analýza trendu provedena v jakékoliv fázi agregace dat. V prvních dvou grafech na obr. 1 a) je znázorněn odhad vývoje pro dva dílčí územní celky (označeny A a B). Následný součet jejich vývoje v roce 2040 je pak znázorněn na obr. 1 b) šedou tečkou. Ve třetím grafu obr. 1 b) je znázorněna i situace, kdy se nejprve sečtou historické produkce dílčích územních celků A a B a až následně se udělá odhad vývoje. Hodnota vývoje tohoto trendu v roce 2040 je pak znázorněna na obr. 1 b) červenou tečkou. Z obr. 1 b) je patrné, že dochází k nesouladu při různém pořadí provedení agregace dat a modelu trendu.



a) Jednotlivé územní celky

b) Agregace územních celků

Obr. 1: Ukázka nesouladu při různém provedení agregace dat a odhadu budoucích hodnot pomocí trendu

Prostý odhad vývoje by tedy měl být korigován s cílem zachovat hierarchickou strukturu území a frakcí odpadu tak, aby platily územní a odpadové bilance.

Cílem vyrovnávání dat je minimální změna výsledků původních odhadů vývoje pro všechny typy odpadu a územní celky tak, aby byly zajištěny bilanční vztahy v datech. Do vyrovnání dat se implementují váhy, které zohledňují různou produkci územních celků a kvalitu odhadu vývoje dle zvoleného kritéria (Příloha 5). Odhadované hodnoty se pak získají za pomoci optimalizačních technik (kap. 8.3.3).

Intervalové odhady (konfidenční a predikční intervaly)

Bodové odhady prognózy by měly být doplněny intervalovými odhady (kap.8.3.4), které umožňují kvantifikovat nejistotu bodového odhadu. Při vyjádření nejistoty prognóz jsou rozlišovány konfidenční a predikční intervaly. Konfidenční interval vyjadřuje nejistotu při odhadu parametrů rozdělení a je vyjádřen jako interval spolehlivosti pro střední hodnotu závisle proměnné, jedná se tedy o intervalový odhad trendu. Predikční interval pak stanovuje nejistotu pro individuální pozorování závisle proměnné. Neurčitost prognózy roste s délkou predikčního horizontu, což způsobuje širší konfidenční a predikční intervaly. Vyjádření nejistoty formou intervalových odhadů by mělo být součástí každé prognózy. Avšak není vhodné intervalové odhady tvořit, pokud jsou časové řady příliš variabilní, nebo prognózy vycházejí z malého množství dat. Konstrukce intervalových odhadů je také ovlivněna přístupem k prognóze (lineární regrese, analýza časových řad, metody strojového učení atd.).

Projekce

Projekce vzniká na základě nějakého definovaného scénáře budoucího vývoje. Projekce odráží expertně nastavené okrajové podmínky, ovšem tak, aby co nejvíce odrážela historický průběh. Tedy, aby byla co nejvíce v souladu s prognózou budoucího vývoje. Takto je možné modelovat budoucí produkci odpadu např. vzhledem ke stanoveným cílům separace odpadu v požadovaném roce. Projekce tedy umožňuje modelovat změny systému ještě před jejich realizací prostřednictvím scénářů. Projekce lze chápat jako scénářová řešení budoucího vývoje podle stanoveného požadavku (kap. 8.4). Protože je v případě projekcí výsledek modelován

při splnění určitých podmínek, není možné tvořit intervalové odhady. Obvykle se tedy přistupuje se k více scénářovým variantám budoucího vývoje.

5.2 Stručné shrnutí postupu metodiky

Navržená metodika je komplexní a pracuje s několika předpoklady a dílčími kroky. Pro prognózování jsou potřebná odpadová data a data vlivných faktorů. Klíčovým bodem metodiky je pak zajištění řádného pre-processingu vstupních dat, včetně odborného ošetření anomálií, které se v datech nacházejí. Důležitým aspektem je splnění tzv. územních a odpadových bilancí.

Pro korektní použití představené metodiky hraje zásadní úlohu splnění následujících požadavků na prognózu:

- pre-processing dat,
- splnění územních bilancí,
- splnění odpadových bilancí,
- intervalové odhady, pokud to konkrétní časová řada umožňuje,
- popis vazeb mezi odpadovými toky,
- tvorba projekcí.

Všechny zmíněné kroky jsou detailně popsány v následujících kapitolách. Grafický přehled o jednotlivých krocích spolu s odkazem na sekci s detailními informacemi dává obr. 2. V hlavních kapitolách dokumentu je popsána metodika v obecné rovině. Detaily volby při konkretizaci jednotlivých kroků metodiky, provedených v rámci projektu TIRSMZP719, jsou pak uvedeny v přílohách.



Obr. 2: Schematické znázornění metodiky prognózování produkce odpadu

Kromě obecného postupu a požadavků na prognózu je obsahem této metodiky doporučení k jednotlivým krokům na základě zkušeností řešitelů projektu TIRSMZP719. Tyto části metodiky (tj. doporučení) by bylo dobré aktualizovat vzhledem k novým poznatkům a zvětšující se datové sadě v rozumném časovém intervalu (doporučuje se jednou za 5 let).

6. Nástroje a předpoklady

V následující části textu jsou popsány potřebné nástroje a předpoklady pro provedení prognózy produkce odpadu dle uvedené metodiky. Dále jsou zde představeny požadavky z oblasti výpočetní techniky a přístupů do databází. Obory, jejichž expertízu je vhodné pro vypracování prognózy zajistit, jsou specifikovány v příloze (Příloha 9).

6.1. Výpočetní technika

Software

Důležitým aspektem výpočtů je rozsah úlohy, tj. množství územních jednotek a typů odpadu. Pro malé úlohy (jednotky prognózovaných údajů) je možno prognózu řešit pomocí běžně dostupných nástrojů jako je například MS Excel. V případě komplexnějších úloh (zahrnujících tisíce prognózovaných údajů) je nutné využít pokročilejších nástrojů pro statistické zpracování dat a optimalizační výpočty. Příkladem těchto nástrojů může být prostředí R pro statistickou analýzu a Julia pro optimalizaci, které byly použity v projektu TIRSMZP719. Nicméně může být použit jakýkoli jiný vhodný software pro statistické zpracování dat a optimalizační výpočty.

Hardware

Z pohledu hardwaru není požadováno žádné speciální vybavení. V případě komplexnějších úloh (zahrnujících tisíce prognózovaných údajů) je vhodné uvažovat o více jádrovém způsobu výpočtů.

6.2. Databáze

Zásadním předpokladem pro provedení prognózy jsou kvalitní a co možná nejúplnější vstupní data. Je nutné mít na vědomí, že kvalitu či úplnost dat nelze nahradit sofistikovanými postupy. Pro prognózování jsou zapotřebí odpadová data a data vlivných faktorů. Celostátní databázi dat o produkci odpadů je ISOH, spravovaný Českou informační agenturou životního prostředí (CENIA). Příloha 1 obsahuje ukázkový požadavek na přípravu dat. V požadavku jsou specifikována potřebná kat. č. odpadu a úroveň územního uspořádání. V případě nutnosti mít data z různých úrovní územního uspořádání je zapotřebí připravit několik požadavků a datových sad. Pro regionální a mikro-regionální prognózy je možné pracovat i s daty krajských úřadů, ORP nebo přímo obcí. Je třeba mít na paměti, že v takových datových sadách není obsažena verifikace dat, jako tomu je v databázi ISOH.

Data možných vlivných faktorů zahrnují např. demografická data, socio-ekonomická data, environmentální data a další. Tato data jsou k dispozici v mnoha databázích, které jsou zajištěny různými subjekty jako např. Český statistický úřad, Ministerstvo financí ČR nebo Česká národní banka. Velké množství dat je volně dostupných. V případě specifického požadavku je možné kontaktovat přímo správce dané databáze. Je třeba zajistit konzistenci sběru dat v celé časové řadě.

7. Sledované ukazatele

V této kapitole je diskutována volba důležitých ukazatelů (kap. 7.1 až kap. 7.4), které je nutno stanovit před samotným provedením prognózy. Tuto volbu je nutno učinit nejen na základě požadavku zadavatele zpracování prognózy, ale také na základě znalostí dostupné datové základny a interakce mezi daty. V této fázi je nutno zaměřit se na volbu územního detailu, frakcí odpadů, období prognózy a nastavení cílů pro případná scénářová řešení. V této kapitole je specifikována volba zmíněných ukazatelů pro účely projektu TIRSMZP719.

7.1. Volba detailu územního členění

Odpadová data v systému ISOH jsou aktuálně dostupná na čtyřech úrovních územního uspořádání, a to obce, ORP, kraje a ČR. Pro volbu detailu územního členění je nutno zohlednit, že data nižších úrovní (např. obce) mají svá specifika (např. vysoká variabilita dat, vyšší výskyt odlehklých či nulových hodnot) a je tedy nutné věnovat značnou pozornost pre-processingu dat. V rámci processingu dat je nutno brát v potaz hierarchickou strukturu dat. Je tedy silně doporučeno provádět výpočty na vícero úrovních, díky čemuž je možno využít zprášující techniky vyrovnávání dat (viz kap. 8.3.3).

7.2. Volba předmětné frakce odpadů

Sledovaným ukazatelem je většinou hmotnost odpadu uvedená v tunách za rok (t/r), případně podíl na celku (% hm) – viz kap. 8.1.1. Různé typy odpadu jsou v databázi ISOH evidovány prostřednictvím kat. č. (viz *Katalog odpadů*). Pro prognózování lze zvolit kat. č. jednotlivě, či je možno prognózovat některá vhodně zvolená kat. č. v agregované formě, resp. částí hmotnostního toku (např. separovaný papír kat. č. 20 01 01 + papírové obaly kat. č. 15 01 01 produkovanými obcemi, přičemž je nutno podotknout, že od roku 2021 došlo ke změně legislativy a podskupina 15 01 již nebude uvažována jako KO). U agregace je nutné uvažovat i vazby pro bilanční vztahy. Je vhodné vzít do úvahy rovněž specifický vývoj některých frakcí odpadu či ojedinělé chování produkce specifického odpadu (viz kap. 8.1.2).

7.3. Volba období prognózy

Odpadová data pro ČR jsou v databázi ISOH k dispozici v ročním detailu. Pro stanovení délky horizontu prognózy je tedy nutno vzít v úvahu množství dostupných dat. Obecně nelze jasně stanovit hranici mezi krátkodobou a dlouhodobou prognózou, ale je nutné vzít na vědomí, že šířka intervalů spolehlivosti se dále do budoucnosti rozšiřuje. Speciálně u horizontu prognózy výrazně delšího, než je časová řada historických dat, je žádoucí doplnit prognózy o vyjádření nejistoty odhadů, kap. 8.3.4. V rámci diagnostiky modelu se doporučuje prognózu verifikovat tak, že jako vstupní data se bude uvažovat zkrácená časová řada historických dat. Po provedení prognózy na této zkrácené časové řadě se kvalita odhadu budoucí produkce porovná se skutečnými daty, která byla před provedením prognózy vyloučena a zachována právě pro toto porovnání.

7.4. Nastavení cílů pro scénářová řešení

Pro prognózu produkce odpadu je vhodné stanovit očekávané scénáře, tj. využít možnost zásahu uživatele do nastavení metodiky za účelem úpravy základní prognózy vzhledem k cílům projekce (např. vyšší separace materiálově využitelných složek). Základní scénář BAU (prognóza) vychází z historických dat a prognóz. Další scénáře (projekce) pak mohou výsledné hodnoty od scénáře BAU odchylovat. Mohou reflektovat např. expertní odhad o efektu budoucích opatření, která se dosud nemohla v historických datech projevit. Scénáře lze také využít pro modelování budoucích stavů zajišťujících cíle (např. cíl pro separaci odpadu). Vytvořené scénáře vychází ze základní prognózy, a tudíž jsou ve scénářích maximálně zohledněna historická data. Tyto simulované scénáře by měly odrážet realitu v rámci potenciálu pro změnu v jednotlivých lokalitách a pro vybrané typy odpadu. Je tedy nutno provést stanovení vhodných podmínek pro tvorbu scénáře (viz kap.8.4). Stejně jako pro prognózu i pro projekce je požadováno splnění územních a odpadových bilancí.

8. Dílčí části metodiky

Tato část se zaměřuje na jednotlivé kroky samotné metodiky prognózování produkce odpadu. Hlavní úkony, popsané níže, jsou následující:

- příprava vstupů,
- pre-processing dat,
- processing dat,
- post-processing dat.

Přehled jednotlivých kroků metodiky prognózy spolu s odkazem na relevantní kapitolu s detailními informacemi je uveden v tab. 2. Ta může sloužit i jako kontrolní seznam provedených kroků při práci na prognóze dle představené metodiky. V hlavních kapitolách dokumentu je popsána metodika v obecné rovině doplněna o doporučení řešitelů projektu TIRSMZP719. Detaily volby při konkretizaci jednotlivých kroků metodiky, pro aplikaci řešenou v rámci projektu TIRSMZP719, jsou pak uvedeny v přílohách, viz Příloha 1 až Příloha 9.

Tab. 2: Jednotlivé kroky postupu metodiky prognózování a reference na detailnější informace

#	Kategorie	Krok	Bližší popis
<input type="checkbox"/> 1	Sledované ukazatele	Volba detailu územního členění	Kap. 7.1
<input type="checkbox"/> 2		Volba předmětné frakce odpadu	Kap. 7.2
<input type="checkbox"/> 3		Volba období prognózy	Kap. 7.3
<input type="checkbox"/> 4		Nastavení cílů pro scénářová řešení	Kap. 7.4
<input type="checkbox"/> 5	Příprava vstupů	Shromáždění a příprava vstupních dat	Kap. 8.1.1
<input type="checkbox"/> 6		Analýza specifík dat	Kap. 8.1.2
<input type="checkbox"/> 7		Agregace dat	Kap. 8.1.3
<input type="checkbox"/> 8	Pre-processing	Analýza odlehklých hodnot v datech	Kap. 8.2.1
<input type="checkbox"/> 9		Analýza skoků v datech a změn trendu	Kap. 8.2.2
<input type="checkbox"/> 11	Processing	Odhad vývoje vlivných faktorů	Kap. 8.3.1
<input type="checkbox"/> 12		Odhad vývoje produkce odpadu	Kap. 8.3.2
<input type="checkbox"/> 13		Vyrovňávání dat	Kap. 8.3.3
<input type="checkbox"/> 14		Intervalové odhady	Kap. 8.3.4
<input type="checkbox"/> 15	Post-processing	Projekce – scénářová řešení	Kap. 8.4
<input type="checkbox"/> 16	Prezentace výsledků	Zpracování a interpretace výsledků	Kap. 9

8.1. Příprava vstupů

Tato kapitola je zaměřena na detaily přípravných činností v rámci vstupů pro prognózování. Hlavní body postupu, popsané v následujících podkapitolách, jsou následující:

- shromáždění a příprava vstupních dat,
- analýza specifík dat,
- agregace dat.

8.1.1. Shromáždění a příprava vstupních dat

Data potřebná pro prognózování lze ve všeobecnosti rozdělit do dvou oblastí, a to odpadová data a data vlivných faktorů. Obecně je nejdříve nutno provést analýzu dostupnosti požadovaných dat. V případě několika různých databází je pak nutno dbát na případnou rozdílnou metodiku zpracování dat v těchto databázích.

Odpadová data

Základním zdrojem dat o produkci odpadu v ČR je databáze ISOH. Odpadová data pro ČR jsou v databázi ISOH k dispozici v ročním intervalu (databáze viz kap. 6.2). Nejdetailnější územní členění, v rámci databáze ISOH, je tzv. základní územní jednotka (ZÚJ). Databáze obsahuje údaje od roku 2002, avšak pro analýzu je vhodné používat data od roku 2009, kdy byla poprvé použita metodika dopočtu podlimitních původců (více v metodice pro příslušný rok – *Matematické vyjádření soustavy indikátorů OH*). V případě vysokého výskytu anomálií v počátečním roce časové řady se doporučuje data v tomto roce vědomě vynechat.

Doporučení 1: Rok 2009 vykazuje v ISOH významně vyšší odlehlost dat z důvodu změny metodiky shromažďování dat. Bez detailní analýzy pro konkrétní odpady se doporučuje tento rok vynechat a pokračovat s datovou sadou od roku 2010.

Data jsou z databáze ISOH exportována ve vybraném detailu uvažovaného územního členění, typicky obce nebo ORP. Data nižších úrovní mají svá specifika (např. vysoká variabilita dat, vyšší výskyt odlehklých či nulových hodnot) a je tedy nutné věnovat značnou pozornost jejich pre-processingu. Nicméně představená metodika využívá i principů agregace a vyrovnávání dat, kdy díky této metodě dochází v některých případech k vyhlazení variability v datech. Je tedy doporučeno provádět výpočty na vícero úrovních územního uspořádání, díky čemuž je možno využít této zpřesňovací techniky. Je nutno také upozornit na to, že z různých důvodů (např. sloučení či vznik nových obcí) ne všechny obce vykazují produkci odpadu v celé časové řadě. Dokonce z různých důvodů někdy hlášení nepodávají ani obce, které už delší dobu existují. Prognózu pak z tohoto důvodu není vhodné provádět čistě na základě historických dat pro všechny aktuálně existující obce.

Doporučení 2: Doporučuje se využít průběh produkce u obcí s podobnými charakteristikami. Příkladem může být přístup řešitele projektu TIRSMZP719 pro prognózu bioodpadu popsany v Příloha 4.

Data o produkci odpadu jsou v ISOH evidována s kódy nakládání A00, AN60 (vykazuje obec nebo firma) a BN30 (vykazuje partner obce). Prognózu produkce odpadu lze zaměřit na vybraná kat. č. odpadu, a to buď individuálně, nebo v agregované podobě. Primárně je většinou prognóza cílena na data a jejich agregace vycházející z *Katalogu odpadů*. Může však být přínosné zabývat se prognózou produkce odpadu i s jiným systémem agregací, než definuje *Katalog odpadů*. V tomto případě lze prognózu provádět přímo agregací prognóz produkce kat. č. nebo je nutno ji provádět separátně v případě přítomnosti specifických vazeb (viz 8.1.2).

Všechny výše zmíněné detaily je nutno specifikovat v požadavku dat z ISOH (viz kap. 6.2 a Příloha 1). Je nutno podotknout, že pokud jsou data požadována na různých úrovních územního uspořádání, je vhodné podávat tyto požadavky zvlášť.

V případě dat z různých databází (např. CENIA a data poskytnutá krajskými úřady), je nutné ověřit platnost územních a odpadových bilancí v těchto historických datech.

Získaná odpadová data udávají celkové množství daného odpadu v jednotkách tun za daný rok a pro danou územní jednotku. Toto celkové množství je v případech některých odpadových toků (viz 8.1.2) vhodné vztáhnout na jednoho obyvatele (kg/osoba a rok) za použití demografických dat. Před tímto přepočtem je však doporučeno nejprve provést agregaci dat do vyšších územních celků (viz kap. 8.1.3). V závislosti na cílech a zaměření prognózy může být zajímavým ukazatelem podíl složek odpadu v definovaném odpadovém toku. Např. složení směsného komunálního odpadu (SKO) umožní odhadnout potenciál separace, což je zásadní informace pro cílené navyšování separace odpadu.

Data vlivných faktorů

Řada přístupů pro modelování současné nebo budoucí produkce odpadu využívá vztahů mezi produkcí odpadu a vývojem socio-ekonomických, demografických, environmentálních a jiných faktorů. Pokud je cílem prognóza produkce odpadu do budoucnosti, je nezbytné disponovat odhadem vývoje jednotlivých faktorů, které do modelů vstupují. Je vhodné zanést do prognózy vývoj takových vlivných faktorů, které mají silnou vazbu na data z OH a současně lze odhadovat jejich vývoj s vyšší kvalitou než lze odhadovat vývoj dat z OH. Pokud by tato kvalita nebyla dostatečná, pak by spíše došlo k zanesení dalších chyb do modelu prognózy odpadu než k jeho zpřesnění. Je také nutno věnovat pozornost tomu, pro jaký detail územního uspořádání jsou data vlivných faktorů k dispozici a reflektovat jejich výběr s ohledem na výběr detailu územního členění pro data produkce odpadu. Analýzu vazeb v datech (např. korelaci) je tedy nutno provést nejen ve vztahu k jednotlivým typům odpadu, ale také vzhledem k produkci daného odpadu na určité úrovni územního uspořádání. Je dobré mít na paměti, že i silné vazby mezi produkcí odpadu a vlivnými faktory se mohou v čase měnit. V případě dlouhodobých prognóz se tato situace dá očekávat.

Doporučení 3:

- a) Při hledání vazeb mezi daty o produkci odpadu a vlivnými faktory se doporučuje znormovat data logickou jednotkou (počet obyvatel, rozloha atd.).
- b) Pro zahrnutí vlivných faktorů do prognózy je nutné, aby prognóza vlivných faktorů pokrývala predikční horizont.
- c) S ohledem na prognózu produkce odpadu se nedoporučuje využívat jiná data než demografická. Data z ostatních oblastí (sociologie, ekonomie atd.) mají obvykle větší variabilitu než data o produkci odpadu a neexistují jejich dlouhodobé prognózy. Krátkodobé prognózy vykazují příliš velkou nejistotu i na rok dopředu. Navíc zde hrozí změna v nalezených vazbách v čase, které mohou mít významně negativní dopad na kvalitu prognózy.

Je vhodné podotknout, že některé vlivné faktory mají cyklický charakter. Ten se může projevit v závislosti na typu prognózy. V případě krátkodobé prognózy se uplatňují spíše cykly s kratším časovým charakterem (např. v detailu měsíců) a projevují se hlavně u sezónních odpadových toků (např. biologicky rozložitelný odpad). Při dlouhodobějších prognózách (v detailu let) se uplatňují socio-ekonomické vlivy. Pro uplatnění v prognóze odpadu však platí, že je nutné disponovat dostatečně kvalitním odhadem vývoje vlivných faktorů, které mají dostatečnou vazbu s odpadovými daty. Další možností je historická data očistit o vliv faktorů s cyklickým charakterem v rámci pre-processingu dat (kap. 8.2).

Vzhledem k potřebnému detailnímu vzhledu do problematiky a souvislostí je doporučeno spolupracovat s odborníky z dané oblasti vlivných faktorů.

8.1.2. Analýza specifík dat

V závislosti na cílech prognózy je potřeba u některých odpadových toků věnovat náležitou pozornost jejich specifickým vazbám, a tedy je nutno toto zohlednit v rámci metodiky. Modelovým příkladem může být KO, na

kterém jsou pro názornost popsány některé z těchto specifických vazeb. Lze uvažovat prognózu složek separovaného odpadu (např. papír, plast nebo bioodpad), které jsou součástí KO. Dalším důležitým odpadovým tokem, spadajícím pod KO, je SKO. Je nutno si uvědomit, že čím větší bude separace, tím méně bude produkován SKO a naopak. Avšak tato vazba mezi SKO a separovanými složkami není absolutní, tj. nově tříděný separovaný odpad pochází jen částečně z SKO a s různou intenzitou pro jednotlivé odpady (viz Příloha 2). Toto úzké propojení je nutno reflektovat i v prognózách jednotlivých odpadových toků. Touto vazbou navíc dochází i k ovlivnění složení SKO a je nutno k tomu náležitě přistupovat, pakliže je složení odpadu součástí zadání provedení prognózy.

Doporučení 4: Při identifikaci vazeb se doporučuje v prvním kroku provést korelační a regresní analýzu. Některé vazby mohou být výrazně ovlivněny lokálními podmínkami a lze očekávat odlišné vazby v různých skupinách územních celků. V globálním pohledu mohou tyto lokální závislosti zcela zaniknout. Pro jejich identifikaci je vhodné hledat územní celky obdobného charakteru. Doporučuje se použít shlukovou analýzu a hledat vazby pro více disjunktních souborů dat. Je velmi pravděpodobné, že se nedostaví požadovaný výsledek z důvodu nedostatku vysvětlujících proměnných. V takovém případě se doporučuje využít hmotnostních bilancí ve zvolených uzlech (území) za pomoci vyrovnávání dat na základě optimalizačních technik, viz Příloha 5.

Dále, produkce některých složek separovaného odpadu může být silně závislá na její aktuální výkupní ceně. Další specifickou vazbou u KO je charakter producenta odpadu. Je možno rozlišovat mezi odpadem, který produkují obce (především pak domácnosti) a který produkují firmy či instituce. Pokud je totiž brán v úvahu přepočet produkce odpadu na jednoho obyvatele (viz kap. 8.1.1), čili zahrnutí demografického faktoru, je nutno zohlednit právě vazbu na producenta odpadu. Obecně specifické vazby úzce souvisí s daným typem prognózovaného odpadu a je proto nutno spolupracovat s odborníky z dané oblasti OH.

Dalším specifikem, v rámci analýzy vstupních dat, může být změna legislativy, která proběhla v rámci časové řady. Příkladem může být rok 2014, kdy dle nově platné legislativy (Vyhláška č. 321/2014 Sb.), musí obec zajistit místo separátního sběru bioodpadu, což mělo vliv na množství této vyříděné složky (především v souvislosti se sběrem bioodpadu ze zahrad).

Dále je vhodné uvažovat o provázanosti např. sociálně-ekonomických dat a dat pro prognózu v OH. Je doporučeno zvážit možnost vyhlazení historických odpadových dat vlivem cyklu v datech (především ekonomická data), případně vlivem extrémních situací, jako je např. pandemie v roce 2020. Zde se očekává významný vliv především na průmyslové odpady. Jistý dopad lze však očekávat u všech typů odpadu a není vhodné tento údaj zahrnovat do prognózy bez patřičné úpravy.

8.1.3. Agregace dat

Jak bylo zmíněno v kap. 8.1.1, data jsou z databáze ISOH exportována ve vybraném detailu uvažovaného územního členění, typicky obce nebo ORP. Představená metodika však využívá i principů agregace dat, a proto je nutno ji provést. Agregace dostupných dat bude uvažována jak z hlediska územního členění (obec, ORP, kraj, stát – viz kap. 8.3.3), tak z hlediska kat. č. nebo celých frakcí odpadů (např. papír, plast, sklo, separované složky, KO – viz kap. 5). Agregaci dat je pak vhodné zahrnout do techniky vyrovnávání dat (viz kap. 5), která umožňuje splnění územních a odpadových bilancí. Jinými slovy součet odhadovaného vývoje pro nižší územní celky se rovná odhadovanému vývoji agregovaného územního celku a součet odhadovaného vývoje jednotlivých frakcí odpadu se rovná odhadovanému vývoji agregovaného odpadového toku.

Doporučení 5: Od roku 2021 dojde ke změně metodiky vykazování KO do databáze ISOH (viz Příloha 2). Obalový materiál produkováný obcemi, který byl doposud v některých případech evidováný v podskupině 15 01, nebude již dále zařazen pod 15 01. Tento odpad budou obce vykazovat pod příslušným kat. č. skupiny 20. Např. odpad dřívě evidováný kat. č. 15 01 01 (papírové a lepenkové obaly) bude od roku 2021 obcemi vykazováný pod kat. č. 20 01 01 (papír a lepenka). Doporučuje se

tedy tuto změnu vykazování řešit agregací příslušných kat. č. ještě před provedením pre-procesingu dat (kap. 8.2) a dalších kroků metodiky.

8.2. Pre-processing dat

Při práci s časovými řadami je nutné se vypořádat s některými jejich vlastnostmi, které by mohly výsledný model prognózy negativně ovlivnit. V rámci pre-procesingu dat je proto nutno věnovat pozornost odlehlým hodnotám a skokům v datech či změnám v trendu. Tato specifika se často vyskytují v datech nižších úrovní (např. obce a ORP). Je vhodné poznamenat, že pre-processing dat je nutné provést nejen pro data z OH, ale i pro vybrané nezávisle proměnné. Je-li to možné z pohledu dostupnosti dat, doporučuje se také vyhlazení cyklického vlivu relevantních vlivných faktorů (pokud byly identifikovány, viz kap. 8.1.1). Tento bod významně zvyšuje nároky na aktualizaci výpočtů, zejména z důvodu shromažďování dat z různých zdrojů. Je tedy vhodné zvážit přínos a náročnost tohoto bodu pro uživatele. Toto vyhlazení je nutno provést před analýzou odlehlých hodnot, skoků v datech či změnách v trendu.

Doporučení 6: Jak je řečeno výše, využití socio-ekonomických dat je velmi omezené pro prognózování. S úspěchem lze ale využít vyhlazení historických dat. S ohledem na dostupná data je vhodné vyhlazení realizovat pouze na agregovaných datech za ČR. Hlavní myšlenka je založena na nalezení významných faktorů produkce odpadu. Následné nalezení průměrného vývoje a cyklu v datech vlivných faktorů umožní zahrnout tento vliv pomocí regrese do historických dat o produkci odpadu.

Doporučení 7: V rámci pre-procesingu je rovněž vhodné pracovat na normovaných datech. U identifikace odlehlých hodnot (viz kap. 8.2.1) normování není nutné. Avšak v případě detekce skoků a změn trendu v datech (viz kap. 8.2.2) je důrazně doporučeno. Pokud by data nebyla znormovaná, bylo by zapotřebí určit kritickou mez pro každou časovou řadu. V případě znormování je pak možno určit pouze jednu kritickou mez, stejnou pro všechna území. S ohledem na aktualizaci výpočtů prognózy není zapotřebí údaj o počtu obyvatel (pro pre-procesingu) aktualizovat každoročně. Tato aktualizace je doporučena v určitém časovém intervalu (např. 5 let). Před samotnou aktualizací by však měla být použita poslední dostupná data.

8.2.1. Analýza odlehlých hodnot v datech

Při analyzování dat bývá obvykle jedním z prvních kroků identifikace těch pozorování, která se výrazně odlišují od většiny ostatních. V oblasti analýzy časových řad se jedná zejména o identifikaci anomální sekvence jako celku a o identifikaci anomální subsekvence v rámci jedné časové řady. Anomální sekvence (řada jako celek) může v případě dat o produkci odpadu zachytit územní jednotky s abnormálně nízkou/vysokou produkcí konkrétní frakce odpadu nebo např. odlišný trend celé řady. Dochází tedy k porovnávání celé vybrané časové řady s ostatními časovými řadami. Druhý formulovaný problém odlehlých hodnot se zabývá spíše chováním v rámci konkrétní sekvence a snaží se detekovat neočekávaný vývoj v této sekvenci. Zde tedy dochází spíše k porovnávání subsekvencí v rámci vybrané časové řady.

Shrnutí: Pro dostatečně dlouhou řadu existuje celá řada metod detekce bodových odlehlých hodnot (např. klasifikace, shluková analýza, statistické, informačně teoretické a spektrální metody). Dostatečná délka řady nebývá specifikována, ale v jednotlivých studiích se typicky pracuje s menším množstvím delších časových řad (alespoň stovky pozorování), výjimečně se objevují metody pro časové řady obsahující pouze vyšší desítky pozorování. Pro krátké časové řady (méně než 20 pozorování) nebyly nalezeny jednotné přístupy. Doporučuje se přístup, který nejdříve odstraní trend a potom analyzuje rezidua. Tento postup lze považovat za kombinaci statistických metod. Kombinuje metody pro nestacionární řady (odstranění trendu) a klasické testy pro odlehlé hodnoty. Pro vyhlazení trendu se doporučuje využít Holtovu metodu a následně analyzovat rezidua s využitím Grubbsova testu. Podrobný popis přístupu je v příloze (Příloha 3).

Doporučení 8: Pro některé typy dat může být vhodné analyzovat odlehlé hodnoty v podobě logaritmické transformace dat o produkci odpadu. Množství identifikovaných odlehlých hodnot se touto transformací bude snižovat. To lze považovat za žádoucí v případě, že očekáváme heteroskedasticitu v datech, tj. multiplikativní změny v produkci na rozdíl od aditivních. Úskali nastává v časových řadách s výskytem nulových hodnot, které nelze zlogaritmovat. V případě vhodného ošetření nulových hodnot se doporučuje logaritmickou transformaci pro potřeby analýzy odlehlých hodnot zvážit. Vhodné ošetření by mělo významně redukovat nebezpečí, že se nulová hodnota stane automaticky odlehlou nebo bude mít velkou páku na podobu trendu. V této chvíli autoři metodiky nemají k dispozici vhodný přístup k ošetření nulových hodnot. Tento bod by měl být předmětem revize metodiky v budoucnosti.

8.2.2. Analýza skoků v datech a změn trendu

Skoky v datech a změny trendu jsou náhlé změny hodnot, chování nebo vlastností časové řady. Díky detekci těchto druhů anomálií je možné zjistit, kdy se struktura časové řady mění a tím i lépe pochopit problém, který je zkoumán. Pomocí testů je možné zjistit, zda se v časové řadě nacházejí skoky a změny trendu a také určit jejich polohu. Změna v chování těchto systémů může být způsobena jak externím zásahem, tak i interní systematickou změnou. Nesprávná identifikace může vést k chybným závěrům a nepřesným modelům prognózy. Pokud jsou u konkrétního kat. č. známy důvody skoků a změn trendu, je vhodné zvážit přínos jejich identifikace a odstranění bodů časové řady pro další výpočty (viz Příloha 4), aby nedošlo ke ztrátě zásadní informace v datech. V případě identifikace skoku či změny trendu se doporučuje hodnoty předcházející této anomálii dále nepoužívat.

Pro analýzu těchto skoků a změn trendu v datech existuje řada vhodných matematických metod. Kromě statistických testů, je možné, pro omezené množství dat, použít i jednoduchou vizuální metodu. Tato metoda je založena na vykreslení závislé i nezávislé proměnné a následném vizuálním hledání okamžitých změn ve struktuře dat. Pro detekci změn trendu a skoků v datech se používají 2 skupiny metod: 1. učení s učitelem a 2. učení bez učitele. První ze zmíněných skupin využívá "databázi" anomálií, na níž vytvoří model strojového učení, jehož cílem je klasifikovat, zda a o jaký typ anomálie se jedná. Využití principů metod učení s učitelem se doporučuje, pokud je databáze anomálií k dispozici. Druhá skupina metod hledá skryté vzorce chování v datech. Nejčastěji se používá srovnávání rozdělení dat před a po možné změně. Pro metody učení bez učitele je nutné disponovat časovou řadou dostatečné délky. Pokud nelze využít ani jeden ze zmíněných přístupů, doporučuje se využít přístup navržený pro řešení prognózy v projektu TIRSMZP719, viz Příloha 3.

Doporučení 9:

1) V rámci pre-processingu se obecně doporučuje identifikovat skoky a změny trendu v datech. Pokud však došlo k vnějšímu zásahu (např. legislativa), který měl za následek změnu produkce, doporučuje se skoky a změny trendu v datech neidentifikovat. U těchto časových řad se očekává systematicky vývoj v podobě S-křivky a je vhodné použít pro prognózu celou historii časové řady. Pro taková kat. č. byl navržen speciální přístup (kap. 8.3.2, Příloha 4).

2) Pro detekci skoků v datech a změn trendu se doporučuje dodržet následující body:

- Využít vizualizace dat, pokud to množství dat umožňuje.
- Neidentifikovat více skoků nebo změn trendu v jedné časové řadě, pokud časová řada nemá více než 15 hodnot.
- Pokud je tvořen obecný postup, pak data znormovat.
- Zaměřit se na úhly mezi dílčími subsekvencemi dat a úhly spojnic historických dat s osou x, viz Příloha 3.
- Pro další výpočty využívat část časové řady za skokem nebo změnou trendu.

- Stanovit minimální počet dat, který má být po pre-processingu zachován po další kroky přístupu.

8.3. Processing dat

Tato kapitola je zaměřena na jednotlivé kroky zpracování dat a prognózování. Hlavní body postupu, popsané v podkapitolách, jsou následující:

- odhad vývoje vlivných faktorů,
- odhad vývoje produkce odpadu,
- vyrovňování dat,
- intervalové odhady.

8.3.1. Odhad vývoje vlivných faktorů

Odhad vývoje vlivných faktorů je nutno provádět na relevantní úrovni územního uspořádání, samozřejmě s ohledem na dostupnost dat. Vzhledem k potřebnému detailnímu vhladu do problematiky a souvislostí je doporučeno spolupracovat s odborníky z dané oblasti vlivných faktorů. Obvyklou metodou pro odhad vlivných faktorů je korelační analýza. Konkrétní vazby pak je možné popsat na základě regresní analýzy nebo metod využívajících strojové učení.

Doporučení 10:

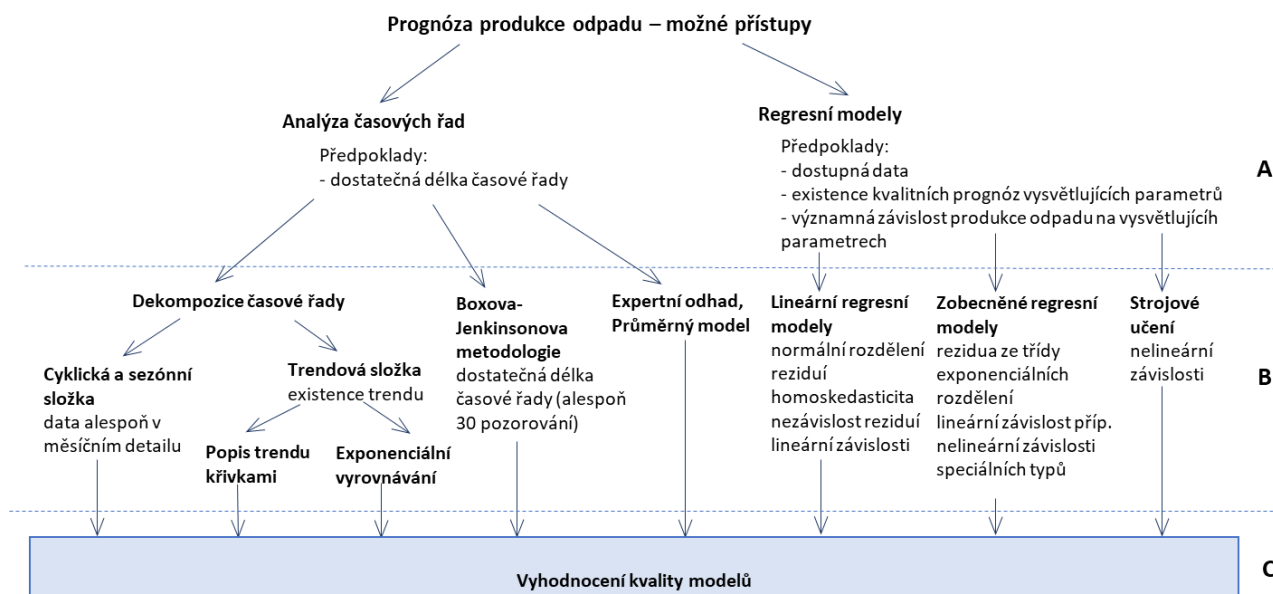
Po pečlivém výběru vlivných faktorů je nutno disponovat i jejich odhadem vývoje do budoucna. Právě kvalita těchto odhadů vývoje by měla být jedním z určujících aspektů pro výběr vlivných faktorů. Do prognózy by měly být zahrnuty jen takové vlivné faktory, které mají silnou vazbu (např. korelaci) na data z OH a současně lze odhadovat jejich vývoj s vyšší kvalitou než data z OH. Pokud by tato kvalita nebyla dostatečná, pak by spíše došlo k zanesení dalších chyb do modelu prognózy odpadu než k jeho zpřesnění, viz (Výsledek V6, 2019). Na základě výsledků získaných v rámci zpracování (Výsledek V6, 2019) se nedoporučuje zahrnovat socio-ekonomická data do prognózy. Pro dlouhodobou prognózu se doporučuje používat pro popis vazeb pouze demografická data, viz Příloha 2.

8.3.2. Odhad vývoje produkce odpadu

Postup výběru vhodné statistické metody pro odhad vývoje produkce odpadu, popsán v jednotlivých krocích níže, je zobrazen na obr. 3 ve formě stromové struktury. Tento postup je obecně platný a byl navržen v rámci projektu TIRSMZP719 za použití rozsáhlé rešerše v dané oblasti v mezinárodním měřítku.

Postup volby přístupu pro odhad vývoje produkce odpadu:

1. Stanovení významných parametrů (podmínka: dostupná data pro všechny územní celky systému).
2. Přepočet dat aktivitním členem na jednotkové množství (podmínka: aktivitní člen je významný parametr).
3. Úroveň A: volba obecného přístupu (podmínka: platnost předpokladů přístupů).
4. Úroveň B: volba vhodných metod a výpočet odhadu vývoje těmito metodami (podmínka: platnost předpokladů metod a zohlednění časové náročnosti výpočtu).
5. Úroveň C: vyhodnocení kvality výsledků modelů posouzením kritérií R^2 (koeficient determinace), $MAPE$ (střední absolutní procentuální chyba) a chyby odhadu vývoje – použití metody s nejvyšší kvalitou modelu.



Obr. 3: Dostupné metody pro odhad vývoje produkce odpadu analyzované v projektu TIRSMZP719

Tato část processingu tedy vychází z výběru vhodné matematicko-statistické metody pro odhad vývoje produkce odpadu dle charakteru dostupných historických dat (viz obr. 3). Není tedy možno obecně popsat postup, poněvadž silně závisí na vybrané metodě. Je tedy alespoň nastíněna obecná metoda zvolená v rámci projektu TIRSMZP719 (detailní informace viz Příloha 4 a lze ji brát jako doporučení autorů metodiky).

V rámci úlohy řešené v projektu TIRSMZP719 (viz Příloha 1 až Příloha 9) byla pro odhad vývoje produkce odpadu zvolena metoda založená na modelování trendu v historických datech pomocí vhodných funkcí. Důvodem jsou nedostupné prognózy vlivných faktorů pro dostatečně dlouhý časový horizont, takže není možné využít metody z větve regresních modelů podle obr. 3. Krátká časová řada v ročním detailu dat zase omezuje využití některých přístupů pro analýzu časových řad.

Obecně je možno pro modelování trendu v datech využít různých tvarů regresních funkcí. Při výběru vhodné regresní funkce je nutné zohlednit vlastnosti vybraných funkcí ve vztahu k očekávanému chování závislosti. Je vhodné hledat kompromis mezi kvalitou proložení podle zvoleného kritéria (např. R^2 , $MAPE$) a vlastnostmi vybraných funkcí. Autoři metodiky vyhodnotili následující vlastnosti jako podstatné:

- **Monotonie** – tvar funkce pro model trendu má za úkol zajistit, že se trend v průběhu horizontu prognózy nemění z rostoucího na klesající a naopak. Trend se předpokládá za monotónní, případné oscilace okolo trendu způsobené sezónní nebo cyklickou složkou není možné na krátké časové řadě ročního detailu sledovat. Požadavkem na monotonii se také sníží riziko přeučení modelu. Autoři metodiky doporučují využít pro modelování trendu mocninnou funkci. Výhodou je její široké uplatnění pro rostoucí i klesající trend.
- **Omezený růst** – některé časové řady nastolily v historických datech velmi významný růst (resp. pokles), který může mít až exponenciální charakter. Takový vývoj je často zaznamenán po změně systému, např. sběrem nové frakce odpadu. Není možné očekávat, že by tak zásadní růst daná časová řada vykazovala dále po celý horizont prognózy. Trend produkce odpadu by v takovém případě rostl nad reálné meze. Pravděpodobnější scénář je takový, že až se bude produkce dané frakce blížit její maximální možné výtěžnosti, růst bude postupně zpomalovat. V takových případech se doporučuje modelovat trend pomocí křivky ve tvaru písmene S, tzv. S-křivky.

V některých případech není vhodné hledat model trendu pomocí složitého funkčního předpisu. Jedná se zejména o časové řady s vysoce variabilními daty, nebo o případy kdy model trendu nemá přínos oproti jednoduchému modelu konstantního proložení dat (průměrem dat, poslední hodnotou atd.).

K proložení dat jednoduchým modelem s konstantní hodnotou se doporučuje přistoupit v následujících případech:

1. Vyloučením dat po pre-processingu zůstává pro modelování trendu příliš krátká časová řada. Jako minimální množství dat pro modelování trendu křivkami v projektu TIRSMZP719 jsou požadovány aspoň 4 body historických dat v posledních 6 letech a současně minimálně 5 bodů historických dat. Počet minimálního počtu dat lze přizpůsobit konkrétní délce časové řady.

2. Model trendu v datech s využitím výše popsanych funkcí je málo kvalitní. Jako kritérium se doporučuje využít R^2 , kritickou mez R^2 je možné přizpůsobit.

3. Jednoduchý model s konstantním průběhem vede ke srovnatelným výsledkům jako složitější model.

Jako speciální případ se doporučuje uvažovat takové časové řady, které v posledních letech evidovaly nulovou produkci modelovaného odpadu. V takových časových řadách se nepředpokládá opětovné zahájení produkce tohoto odpadu a doporučuje se extrapolovat jako nulovou hodnotu (podrobnosti viz Příloha 4). Model trendu je následně korigován v rámci vyrovnavání dat, viz 8.3.3.

Pokud data zahrnují dostatečný počet historických dat, je možné sledovat i cyklickou složku časové řady. Potřebný počet takových dat nelze obecně určit, odvíjí se zejména od délky cyklu, který se v datech vyskytuje. Pokud je však cyklická složka v datech identifikovaná, je vhodné ji zohlednit také při extrapolaci.

Odpadová data jsou v systému ISOH uvedena v absolutním množství odpadu (tj. tuny za danou územní jednotku a rok). Pokud jsou k dispozici demografické projekce, doporučuje se tuto informaci zahrnout do prognózy produkce odpadu od občanů. V takových případech je absolutní množství odpadu nejdříve přepočteno na jednoho obyvatele (v kg/osoba a rok) a následně se zahrne odhad demografického vývoje pro získání informace o absolutní produkci odpadu. Před tímto přepočtem je však doporučeno nejprve provést agregaci dat do vyšších územních celků.

Doporučení 11: V některých případech může být výhodné využití principů poissonovské regrese. V porovnání s modelem trendu pomocí nelineární úlohy sice dosahuje méně přesných výsledků. Avšak značnou výhodou jsou významně nižší požadavky na výpočetní čas. Poissonovskou regresi se tedy doporučuje aplikovat, pokud je zásadní snižovat výpočetní náročnost přístupu z důvodu častého opakování výpočtu. V případě modelování trendu S-křivkou je možné využít beta regresi, výpočetní náročnost beta regrese je ale srovnatelná s modelováním trendu křivkami. Problém dále nastává s nulovými hodnotami produkce, protože beta regrese nepřipouští krajní hodnoty intervalu (0, 1), na kterém je definovaná. Je tedy nutné přeškálovat vstupní data vhodným způsobem.

Doporučení 12: V některých případech odpadů s vazbou může být přínosné vytvořit fiktivní odpadový tok a prognózovat poměr dané frakce k tomuto fiktivnímu toku. Jako příklad lze uvést separovaný KO (papír, plast, sklo atd.) a zbytkový SKO. Výhodou tohoto přístupu může být zejména snížení variability v prognózovaných hodnotách. Úskalí je však ve velké závislosti na kvalitní prognóze agregované časové řady. Ze zkušeností jsou i tyto řady často velmi variabilní a jejich trend není jasně v datech patrný. Proto se nedoporučuje tuto úpravu využít ve všech prognózách, ale před její aplikací zvážit možná úskalí tohoto přístupu.

8.3.3. Vyrovnávání dat

Metoda vyrovnání dat (angl. data reconciliation) umožňuje relevantně uchopit problematiku prognózování na základě i krátkých časových řad při splnění bilančních podmínek pro výsledné prognózy. Teoretický koncept vyrovnávání dat vychází z faktu, že odhady budoucích hodnot nejsou v logickém souladu, tzn. nesplňují územní nebo odpadovou bilanci dané agregací dat. U odhadnutých budoucích hodnot se předpokládá jejich zatížení náhodnou chybou, která reflektuje proveditelnou přesnost odhadu vývoje vzhledem k délce a variabilitě časové řady. Tato část processingu dat zajistí, že prostý trend je korigován s cílem zachovat hierarchickou strukturu území a frakcí odpadu tak, aby platily požadované vazby.

Cílem vyrovnávání dat je minimálně změnit výsledky původních odhadů vývoje pro všechny typy odpadu a územní celky tak, aby byly zajištěny bilanční vztahy v datech. Odhadované hodnoty je pak možné získat za pomoci optimalizačních technik. Do vyrovnávání dat je vhodné implementovat váhy, které zohledňují různé množství produkovaného odpadu a kvalitu odhadu vývoje dle zvoleného kritéria. Jedná se o váhy pro zajištění totožné významnosti vstupních dat (použitých pro odhadu vývoje) a také váhy zohledňující kvalitu samotného odhadu vývoje produkce odpadu.

Doporučení 13: Model pro vyrovnávání dat lze formulovat v aditivní nebo multiplikativní podobě (viz Příloha 5). Volba podoby modelu závisí především na výpočetních možnostech zvoleného řešiče. Multiplikativní podoba zápisu má významnou výhodu u odpadů s velkou variabilitou mezi jednotlivými frakcemi (např. KO). V takových případech se vyskytují na různých úrovních území producenti s řádově rozdílnou produkcí což v případě aditivní podoby působí numerické a zaokrouhlovací chyby. Na formě zápisu dále závisí nastavení vah modelu (viz Příloha 5).

Váhy pro zajištění totožné významnosti vstupních dat

Cílem procesu navržení těchto vah je, aby veškeré vstupní údaje byly stejně významné při stejné kvalitě dat. V rámci systému hierarchického uspořádání přirozeně dochází k tomu, že množství odpadu v ČR je mnohonásobně větší než např. v ORP. To samé však může nastat i pro případ dvou různě velkých ORP. Systém vah tedy zajistí, že je chyba minimalizována pro všechna území stejně (čili jde o druh znormování). V případě multiplikativní podoby modelu je minimalizován procentuální odklon od vstupních hodnot a není tedy nutné váhu pro zajištění totožné významnosti vah zahrnovat.

Váhy zohledňující kvalitu odhadu vývoje produkce odpadu

Při provádění odhadu vývoje výsledný model reflektuje historická data s různou kvalitou proložení (těsností). To je dáno především variabilitou v historických datech. Obecně lze předpokládat, že čím je kvalitnější model, tím jsou odhadovaná data v budoucnu blíže skutečné hodnotě. Tento předpoklad je vhodné zapracovat do procesu vyrovnání dat pomocí vah, viz Příloha 5. Detailní popis matematického modelu, postupu řešení a nastavení vah pro vyrovnání dat udává Příloha 5. Zvýšenou pozornost je nutné věnovat takovým situacím, kdy je trend v historických datech konstruován na různých úrovních území rozdílným způsobem. Z důvodu různého přístupu ke konstrukci trendu není určení váhy na různých úrovních jednotné. Je tedy nutné zkorigovat interval, kterého mohou váhy nabývat v jednotlivých úrovních území (viz Příloha 5).

Doporučení 14: Pro rozsáhlé úlohy může dojít k situaci, kdy nebude možné najít optimální řešení. Obvykle je v takových situacích k dispozici alespoň relaxované řešení, tzn. bilance nejsou splněny přesně. Pro některé tvorby prognóz to nemusí být problém. V opačných případech se doporučuje zmenšit optimalizační úlohu tak, že budou bilancovány menší množiny odpadových toků, např. pouze pro jednotlivá kat. č.

8.3.4. Intervalové odhady

Výstupem dosavadních kroků metodiky jsou bodové odhady množství produkovaného odpadu. Avšak jedním ze základů teorie pravděpodobnosti je skutečnost, že pro spojitou náhodnou veličinu pravděpodobnost nabytí konkrétně stanovené hodnoty je nulová. Kvůli tomu jsou ve velké míře používány intervalové odhady, které jsou schopné se zadanou spolehlivostí stanovit meze pro výskyt bodové hodnoty prognózy. Způsob konstrukce intervalových odhadů je závislý na zvoleném přístupu pro odhad vývoje produkce. V některých případech (např. lineární regrese, některé metody analýzy časových řad) budou využity přístupy pro konstrukci intervalových odhadů, které se běžně s danou metodou využívají. V ostatních případech se doporučuje využít principů metody parametrického bootstrapu. Kvůli manipulaci s daty, která zajistí splnění jejich hierarchické struktury (tj. vyrovnávání dat), nejsou klasické metody pro stanovení intervalových odhadů přímo použitelné. Je nutno vzít toto v potaz a přistoupit ke konstrukci jisté obdoby intervalových odhadů pro prognózu možných realizací vývoje množství odpadů, které specifika vyrovnávání dat reflektují. Tento postup, který kombinuje vyrovnávání dat a tvorbu intervalových odhadů je jeden z možných. Důležité je, aby tvůrci prognóz dodrželi přestavené požadavky (viz kap. 5).

Doporučení 15: Na základě projektu TIRSMZP719 byla doporučena konstrukce intervalových odhadů na základě metody známé jako parametrický bootstrap. Detailní popis výpočtu dává Příloha 6.

8.4. Post-processing dat – Projekce (scénářová řešení)

Doposud se vycházelo z historických dat a výsledkem je prognóza (základní scénář). Nastávají situace, kdy je do systému zasahováno z vnějšku a tím je měněn historický trend. Tyto vlivy (legislativní, technologický pokrok atd.) nelze předpovídat. Proto se často přistupuje ke konstrukci projekcí, tj. scénářů budoucího vývoje ve vazbě na konkrétní okrajové podmínky zvolené autorem scénáře. Scénáře je možné vytvářet s ohledem na cíle OH. Vzhledem k územní hierarchii, je také vhodné využít možnosti rozpadu celostátních cílů, které lze rozdělit do jednotlivých regionů podle jejich potenciálu pro změnu.

U územních celků, kde je velký potenciál pro změnu, lze očekávat větší změnu (např. zvyšování třídění odpadu). Obecně platí, že je vhodné se co nejméně odchylovat od základního scénáře a zároveň splnit požadavky na scénář.

Doporučení 16: Autoři nedoporučují pro projekce konstruovat intervalové odhady. Přínos scénářových řešení je zejména v případech, kde se vyskytují vazby mezi daty. Pro tvorbu scénářových řešení je nutné stanovit potenciál pro změnu. Pro scénářová řešení platí následující (Příloha 7):

- Scénář nepřekračuje potenciál pro změnu stanovený pro konkrétní územní jednotku.
- Všechny územní jednotky vykazují posun ke splnění scénáře, pokud to potenciál dovoluje.
- Změna pro splnění scénáře na nižších celcích než ČR probíhá tak, že se jednotlivá území v míře plnění potenciálu nepředbíhají a jsou monotónní.

9. Zpracování a interpretace výsledků

Vhodné zpracování výsledků zahrnuje grafickou prezentaci a tabulkové výstupy, viz Příloha 8. Grafické výstupy prognózy umožňují přehledně sledovat historický vývoj, a navíc očekávaný vývoj v čase. Speciálním případem grafických výstupů jsou mapy, které umožní rychlé vizuální porovnání situace v jednotlivých územních celcích pro daný rok. Forma tabulky s konkrétními číselnými hodnotami je pak užitečná zejména pro následné využití výsledků.

V obou typech výstupů je pak nutno stanovit požadované detaily zobrazení výsledků prognózy, a to zejména úroveň území, volba konkrétního území, agregace území, skupina odpadu podle *Katalogu odpadů*, podskupina odpadu, kat. č. odpadu, agregace kat. č. nebo rok prognózy.

Shrnutí: Pro grafický a tabulkový výstup se doporučuje znázornit následující:

- historická data,
- data vyřazená při pre-processingu,
- trend, resp. bilanci namísto skutečných historických dat v případě neveřejných dat,
- prognóza v podobě konfidenčních a predikčních intervalů,
- projekce, viz Příloha 8.

Výstupy je doporučeno tvořit v absolutním množství i vztahené na zvolenou jednotku (např. na obyvatele).

10. Revize použitých přístupů v metodice a kontrola výsledků prognózy

Metodika (tj. obecný postup) je navržena tak, aby byla i v budoucnu platná. Je však silně doporučeno v pravidelných intervalech provést revizi a konkretizaci dílčích kroků metodiky (např. revizi navržených matematicko-statistických postupů) a také ověřit, zda je postup jako celek ještě aktuální. Dá se očekávat rozšiřování datové základny a mohou se objevit nové vazby, nové vlivné faktory nebo nové důležité odpadové toky. Doporučuje se provádět revizi dílčích kroků metodiky v pravidelných intervalech po 5 letech.

Konkretizace dílčích kroků metodiky byla provedena v rámci projektu TIRSMZP719 a je uvedena v přílohách tohoto dokumentu (viz seznam příloh v kap. 12). Tato konkretizace metodiky odráží současný stav OH včetně dostupné datové základny. Praktická aplikace metodiky, v rámci projektu TIRSMZP719, se zaměřila na důležité odpadové toky současnosti.

Po provedení prognózy je zapotřebí disponovat plánem na kontrolu výsledků, a to v krátkodobém i dlouhodobém horizontu (pokud je to možné). Doporučuje se vynechat část nejnovějších historických dat a provést prognózu na zkrácené časové řadě. Data, která nevstupovala do prognózy, se využijí pro posouzení kvality odhadu očekávaného vývoje. Je vhodné analyzovat výsledky i pro jednotlivé části prognózy. Z pohledu pre-processingu nebo chyb ve vstupních datech se doporučuje vizuální analýza více nezávislých pozorovatelů pro ověření kritických mezí v rámci jednotlivých analýz.

V budoucnu je potřeba počítat s aktualizací vytvořené prognózy v případě doplnění dat (např. o další rok). V tom případě je možné provést nový výpočet a prognózu tím aktualizovat.

11. Závěrečná zpráva

Struktura závěrečné zprávy o výsledcích studie by měla obsahovat následující údaje:

- 1) Zadavatel studie.
- 2) Autor – dodavatel studie, včetně sub-kontraktorů z mezioborové spolupráce na studii.
- 3) Popis vstupních a okrajových podmínek studie:
 - zájmové území a jeho členění,
 - prognózované odpadové toky,
 - vstupní data (odpadová data, data vlivných faktorů) – zdroj, detail a množství,
 - predikční horizont,
 - specifické vazby, pokud jsou přítomny,
 - nastavení cílů pro scénářová řešení.
- 4) Popis konkretizace dílčích kroků metodiky (pre-processing, processing, post-processing):
 - zvolené metody pro jednotlivé kroky,
 - splnění předpokladů použitých metod,
 - splnění předpokladů samotné prognózy,
 - popis dílčích výsledků jednotlivých kroků metodiky a diskuse.
- 5) Shrnutí výsledků prognózy:
 - grafické výstupy,
 - tabulkové výstupy,
 - diskuse.
- 6) Porovnání výsledků studie s výsledky předchozí studie nebo srovnatelného studie (pokud existuje).
- 7) Diskuse o potenciálních faktorech, které mohly mít vliv na výsledky prognózy.
- 8) Závěr – shrnutí realizované prognózy, případná doporučení pro další práci.
- 9) Přílohy (pracovní protokoly a detailní výsledky).

12. Seznam příloh

Níže je uveden seznam příloh, zpracovaných v samostatném dokumentu „*Certifikovaná metodika pro provádění dlouhodobé prognózy produkce odpadů v ČR včetně revize prognózy (Výsledek V9) – PŘÍLOHY*“. Přílohy (Příloha 1 až Příloha 9) představují konkrétní využití metodiky pro prognózu, která byla řešena v rámci projektu TIRSMZP719. V budoucnu je možné doplňovat přílohy k metodice v podobě dalších aplikací.

Příloha 1: Požadavek na odpadová data z databáze ISOH.

Příloha 2: Vstupní a okrajové podmínky metodiky v rámci projektu TIRSMZP719.

Příloha 3: Pre-processing – analýza anomálií v datech.

Příloha 4: Processing – odhad vývoje produkce odpadu.

Příloha 5: Processing – vyrovnávání dat.

Příloha 6: Processing – intervalové odhady.

Příloha 7: Post-processing – projekce (scénářová řešení).

Příloha 8: Zpracování a interpretace výsledků.

Příloha 9: Požadovaná expertíza.

Reference

- Abbasi, M., Hanandeh, A.E., 2016. Forecasting municipal solid waste generation using artificial intelligence modelling approaches. *Waste Management*, 56, 13-22. doi: 10.1016/j.wasman.2016.05.018.
- Andersen, F., M., Larsen, H.V., 2012. FRIDA: A model for the generation and handling of solid waste in Denmark. *Resources, Conservation and Recycling*, 65, 47-56, doi: 10.1016/j.resconrec.2012.04.004.
- Andersen, F.M., Larsen, H., Skovgaard, M., Moll, S., Isoard, S., 2007. A European model for waste and material flows. *Resources, Conservation and Recycling*, 49, 421-435. doi: 10.1016/j.resconrec.2006.05.011.
- Beigl, P., Lebersorger, S., Salhofer, S., 2008. Modelling municipal solid waste generation: A review. *Waste Management*, 28(1), 200-214. doi: 10.1016/j.wasman.2006.12.011.
- Cole, Ch., Quddus, M., Wheatley, A., Osmani, M., Kay, K., 2014. The impact of Local Authorities' interventions on household waste collection: A case study approach using time series modelling. *Waste Management*, 34(2), 266-272. doi: 10.1016/j.wasman.2013.10.018.
- Dwivedy, M., Mittal, R.K., 2010. Future trends in computer waste generation in India. *Waste Management*, 30(11), 2265-2277. doi: 10.1016/j.wasman.2010.06.025.
- ETC/EEA, 2012. Revision of the MSW Generation Projection Equations Based on Additional Data Points for 2009 and 2010, prepared by Andersen et al.
- Eunomia Research & Consulting, Copenhagen Resource Institute, 2014. Development of a Modelling Tool on Waste Generation and Management – Headline Project Report. Final Report for the European Commission DG, prepared by Gibbs et al.
- Ghinea, C., Drăgoi, E.N., Comăniță, E.-D., Gavrilăscu, M., Câmpeanu, T., Curteanu, S., Gavrilăscu, M., 2016. Forecasting municipal solid waste generation using prognostic tools and regression analysis. *Journal of Environmental Management*, 182, 80-93. doi: 10.1016/j.jenvman.2016.07.026.
- Goel, S., Ranjan, V.P., Bardhan, B., Hazra, T., 2017. Forecasting solid waste generation rates. *Modelling Trends in Solid and Hazardous Waste Management*, 35-64. doi: 10.1007/978-981-10-2410-8_3.
- Hřebíček, J., Kalina, J., Soukopová, J., Horáková, E., Prášek, J., Valtá, J., 2017. Modelling and Forecasting Waste Generation – DECWASTE Information System. 12th International Symposium on Environmental Software Systems (ISESS), 433–445. doi: 10.1007/978-3-319-89935-0_36.
- ISOH – Informační systém odpadového hospodářství. Ministerstvo životního prostředí [online]. 2021. Dostupné z: <https://isoh.mzp.cz/>.
- Johnson, N.E., Ianiuk, O., Cazap, D., Liu, L., Starobin, D., Dobler, G., Ghandehari, M., 2017. Patterns of waste generation: A gradient boosting model for short-term waste prediction in New York City. *Waste Management*, 62, 3-11. doi: 10.1016/j.wasman.2017.01.037.
- Karpušenkaitė, A., Ruzgas, T., Denafas, G., 2016. Forecasting medical waste generation using short and extra short datasets: Case study of Lithuania. *Waste Management & Research*, 34(4), 378–387. doi: 10.1177/0734242X16628977.
- Lebersorger, S., Beigl, P., 2011. Municipal solid waste generation in municipalities: Quantifying impacts of household structure, commercial waste and domestic fuel. *Waste Management*, 31(9–10), 1907-1915. doi: 10.1016/j.wasman.2011.05.016.
- Noori, R., Abdoli, M.A., Ameri Ghasrodashti, A., Jalili Ghazizade, M., 2009. Prediction of municipal solid waste generation with combination of support vector machine and principal component analysis: A case study of Mashhad. *Environmental Progress and Sustainable Energy*, 28 (2), 249-258. doi: 10.1002/ep.10317.

Noori, R., Karbassi, A., Salman Sabahi, M., 2010. Evaluation of PCA and Gamma test techniques on ANN operation for weekly solid waste prediction. Journal of Environmental Management, 91(3), 767-771. doi: 10.1016/j.jenvman.2009.10.007.

Petridis, N.E., Stiakakis, E., Petridis, K., Dey, P., 2016. Estimation of computer waste quantities using forecasting techniques. Journal of Cleaner Production, 112(4), 3072-3085. doi: 10.1016/j.jclepro.2015.09.119.

Šomplák, R., Smejkalová, V., Rosecký, M., Bouda, Z., Suzová, J., Roupec, J., Szásziová, L., Popela, P., Kúdela, J., Pavlas, M., 2019. Výzkumná zpráva popisující výběr vhodných metod datové analýzy a matematického modelování prognózy produkce odpadů (Výsledek V6). TIRSMZP719.

Šomplák, R., Smejkalová, V., Rosecký, M., Kúdela, J., Popela, P., Pavlas, M., 2019. Výzkumná zpráva popisující přístup k řešení prognózování produkce odpadů v mezinárodním měřítku (Výsledek V5). TIRSMZP719.

Šomplák, R., Smejkalová, V., Rosecký, M., Popela, P., Kúdela, J., Eryganov, I., Šramková, K., Pavlas, M., 2020. Výzkumná zpráva popisující princip a matematický model prognózování produkce odpadů (Výsledek V7). TIRSMZP719.

Šomplák, R., Smejkalová, V., Rosecký, M., Popela, P., Kúdela, J., Eryganov, I., Šramková, K., Pluskal, J., Pavlas, M., 2021. Souhrnná komplexní výzkumná zpráva popisující nastavení a řešení zpracování dlouhodobé prognózy produkce odpadů v ČR (Výsledek V8). TIRSMZP719.

Veverka, Z., Bouda, Z., Suzová, J., Gregor, J., Kropáč, J., Pavlas, M., 2020. Metodika pro stanovení složení směsného komunálního odpadu z obcí a komunálního odpadu (Výsledek V4). TIRSMZP719.

Vyhláška č. 321/2014 Sb. Vyhláška o rozsahu a způsobu zajištění odděleného soustředování složek komunálních odpadů.

Vyhláška č. 8/2021 Sb. Vyhláška o Katalogu odpadů a posuzování vlastností odpadů (Katalog odpadů).