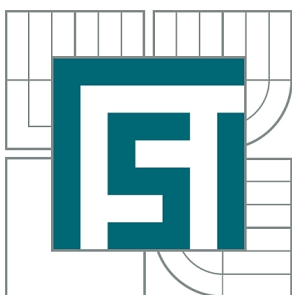




VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA STROJNÍHO INŽENÝRSTVÍ
ÚSTAV MATEMATIKY

FACULTY OF MECHANICAL ENGINEERING
INSTITUTE OF MATHEMATICS

MODELOVÁNÍ ENERGETICKÉHO ZDROJE A PLÁNOVÁNÍ JEHO PROVOZU S VYUŽITÍM POKROČILÝCH MATEMATICKÝCH METOD

MODELLING OF ENERGY PRODUCING SYSTEM AND ITS OPERATION PLANNING APPLYING
ADVANCED MATHEMATICAL METHODS

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

Bc. JANA BENÁČKOVÁ

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. MARTIN PAVLAS, Ph.D.

BRNO 2011

Vysoké učení technické v Brně, Fakulta strojního inženýrství

Ústav matematiky

Akademický rok: 2010/2011

ZADÁNÍ DIPLOMOVÉ PRÁCE

student(ka): Bc. Jana Benáčková

který/která studuje v **magisterském navazujícím studijním programu**

obor: **Matematické inženýrství (3901T021)**

Ředitel ústavu Vám v souladu se zákonem č.111/1998 o vysokých školách a se Studijním a zkušebním řádem VUT v Brně určuje následující téma diplomové práce:

Modelování energetického zdroje a plánování jeho provozu s využitím pokročilých matematických metod

v anglickém jazyce:

Modelling of energy producing system and its operation planning applying advanced mathematical methods

Stručná charakteristika problematiky úkolu:

Tato práce se zabývá jednak využitím statistických metod, jednak využitím metod matematického programování pro řešení úloh souvisejících s modelováním a plánováním provozu existujícího energetického zdroje. Uvažovaným zdrojem je zařízení, které, kromě uhlí, využívá také alternativní paliva, zejména biomasu. Vzhledem k tomu, že jsou k dispozici provozní data, je možné vytvořit model na jejich základě. Ovšem některé modely technologických uzlů vyžadují použití pokročilých statistických metod. Celkový model zdroje je pak možné použít pro návrh optimálního plánování provozu s ohledem na ekonomiku. Zejména jde o plánování dávkování paliv, výroby energií, atd.

Cíle diplomové práce:

Cílem práce je aplikovat pokročilejší statistické metody při vytváření modelu energetického zdroje založeného na provozních datech. Poté s využitím modelu navrhnout plán provozu zařízení. Poté bude úkolem vyhodnotit výsledky výpočtů a okomentovat je.

Seznam odborné literatury:

- BIRGE, John R.; LOUVEAUX, Francois. Introduction to stochastic programming. Berlin:Springer, 1997. 447s. ISBN 0-387-98217-5
- KALL, Peter and WALLACE, Stein W.. Stochastic Programming, John Wiley & Sons, Chichester, 1994
- BAXTER, Larry. Biomass-coal co-combustion: opportunity for affordable renewable energy. Fuel. 2005, 84, 10, s. 1295-1302. ISSN 0016-2361
- DE, S; ASSADI, M. Impact of cofiring biomass with coal in power plants - A techno-economic assessment. Biomass and Bioenergy. 2009, 33, 2, s. 283-293. ISSN 0961-9534

Vedoucí diplomové práce: Ing. Martin Pavlas, Ph.D.

Termín odevzdání diplomové práce je stanoven časovým plánem akademického roku 2010/2011.

V Brně, dne 22.10.2010

L.S.

prof. RNDr. Josef Šlapal, CSc.
Ředitel ústavu

prof. RNDr. Miroslav Doupovec, CSc.
Děkan fakulty

Abstrakt

Tato práce se zabývá návrhem ekonomicky efektivního využití biomasy při spoluspalování s uhlím pro reálnou teplárnu, jejímž požadavkem bylo optimální využití stávající palivové základny a naplánování ročního provozu vzhledem k měsíční poptávce po teple. Při vytváření matematického modelu teplárny, který je založen na poskytnutých provozních datech, aplikujeme regresní analýzu a stochastické programování. Celkový model energetického zdroje pak používáme pro návrh optimálního plánování provozu s ohledem na ekonomiku. Zejména jde o plánování dávkování paliv a výroby energií.

Summary

This master's thesis deals with the proposal of the cost-effective biomass and coal combustion concept for a real generation plant. The optimization of the current fuel basis exploitation and annual operation scheduling was the main goal. Applying the regression analysis and stochastic programming the mathematical model was constructed based on the operating data. The overall energy source model was implemented to the concept of the optimal operation scheduling considering the economic aspects. The fuel utilization and energy production planning are the main applications of this design.

Klíčová slova

Stochastické programování, regresní analýza, optimalizace, biomasa, teplárna

Keywords

Stochastic programming, regression analysis, optimization, biomass, heating plant

BENÁČKOVÁ, J. Modelování energetického zdroje a plánování jeho provozu s využitím pokročilých matematických metod. Brno: Vysoké učení technické v Brně, Fakulta strojního inženýrství, 2011. 62 s. Vedoucí diplomové práce Ing. Martin Pavlas, Ph.D.

Prohlašuji, že jsem svoji diplomovou práci *Modelování energetického zdroje a plánování jeho provozu s využitím pokročilých matematických metod* vypracovala samostatně pod vedením Ing. Martina Pavlase, Ph.D. s použitím materiálů uvedených v seznamu literatury.

Jana Benáčková

Ráda bych poděkovala svému školiteli Ing. Martinu Pavlasovi, Ph.D. za vedení mé diplomové práce. Dále bych chtěla poděkovat Ing. Michalu Toušovi za pomoc, přínosné rady a čas, který mi věnoval. V neposlední řadě bych ráda poděkovala RNDr. Pavlu Popelovi, Ph.D. za pomoc, podporu a užitečné připomínky.

Jana Benáčková

Obsah

1	Úvod	10
2	Motivace	11
2.1	Biomasa a její využití	11
2.2	Popis teplárenského provozu	12
3	Deterministický model teplárny	15
3.1	Indexové množiny	15
3.2	Omezení	15
3.3	Účelová funkce	17
3.4	Model	18
4	Matematické programování	19
4.1	Deterministické programování	19
4.2	Stochastické programování	20
4.2.1	Wait-and-See přístup	21
4.2.2	Here-and-Now přístup	21
4.2.3	Scénářové úlohy	21
5	Jednorozměrná regrese	23
5.1	Výběr regresní funkce	23
5.1.1	Typy regresních funkcí	23
5.2	Odhad parametrů β	25
5.3	Analýza reziduí	26
5.3.1	Normalita reziduí	27
5.3.2	Homoskedasticita a heteroskedasticita	29
5.3.3	Autokorelace reziduí	31
5.4	Testy hypotéz o regresních koeficientech	31
5.5	Index determinace	33
6	Vícerozměrná regrese	34
6.1	Výběr proměnných	34
6.2	Tvar regresní funkce	35
6.3	Vlastnosti dat	35
6.3.1	Nezávislost proměnných	36
6.3.2	Multikolinearita	36
6.4	Analýza reziduí, testy koeficientů, index determinace	37
7	Neurčitosti v modelu	41
7.1	Účinnost kotlů	41
7.2	Velikost poptávky po teple	43
7.3	Vývoj cen dotací na výrobu elektřiny z biomasy	45
8	Stochastický model teplárny	47
8.1	WS model teplárny	47
8.2	Analýza výsledků	47
9	Závěr	52

1 Úvod

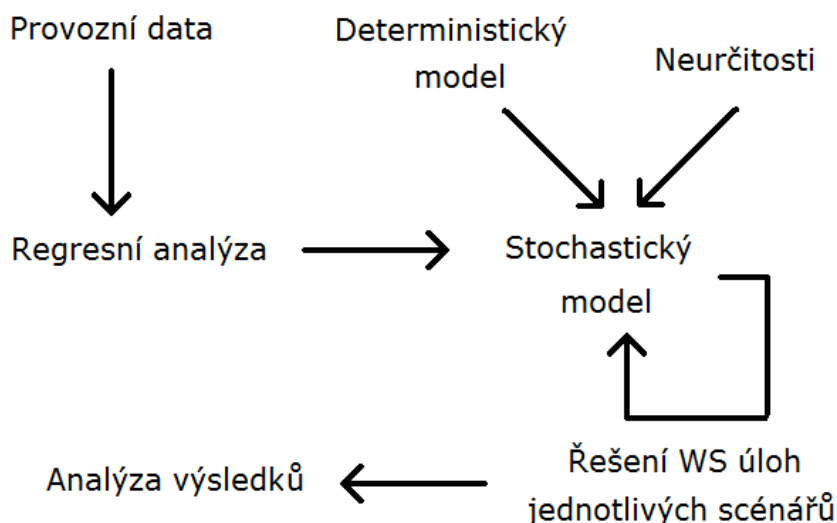
Motivací k tématu mé diplomové práce je projekt optimalizace teplárenského provozu, na kterém spolupracuje Ústav matematiky (ÚM) s Ústavem procesního a ekologického inženýrství (ÚPEI) v souvislosti s aktivitami projektu Výzkumné centrum MŠMT CQR č.1M06047. Jedná se o vhodné začlenění biomasových paliv do stávajícího energetického zdroje s cílem maximálního finančního zisku. Teplárna pracuje v kogeneračním režimu, což znamená, že současně vyrábí tepelnou a elektrickou energii.

Zisk teplárny se skládá z prodeje energií a navíc ze zelených bonusů, naopak náklady teplárny tvoří především nákup paliv a dále náklady na vlastní provoz. Zelenými bonusy se stát snaží podpořit využívání obnovitelných zdrojů energie, protože náklady na takto vyrobenou energii jsou v porovnání s energií z neobnovitelných zdrojů (fosilních paliv) mnohem vyšší, podrobnosti viz kapitola 2 a 3.

Cílem práce je sestavení vhodného stochastického modelu dané teplárny, který bude věrně popisovat reálný provoz. Jelikož máme k dispozici provozní data teplárny, je možné na jejich základě tento model vytvořit. Pro co nejpřesnější popsání tohoto provozu využíváme matematických metod, které popisujeme v kapitolách 4-6. Velmi důležité je stochastické programování, které nám umožňuje popsat a následně implementovat neurčitosti a náhodnosti reálného provozu do našeho modelu, navíc tento model následně efektivně řešit. Pro určení závislostí z poskytnutých provozních dat využíváme regresní analýzu.

Vytvořený stochastický model uvádíme v kapitole 8, poté ho implementujeme do optimalizačního softwaru GAMS (General Algebraic Modeling System). Výpočty potřebné především k regresní analýze budou prováděny ve výpočetním programu Matlab.

V poslední části diplomové práce uvádíme a popisujeme výsledky stochastického modelu teplárny, do kterého zařadíme scénáře popsanych neurčitostí (výše poptávky a hodnota zelených bonusů). Cílem je, aby tento model nám pro danou poptávku určil optimální dávkování všech dostupných druhů paliv, které nám zaručí maximální zisk. Následující schéma ukazuje vývoj práce.

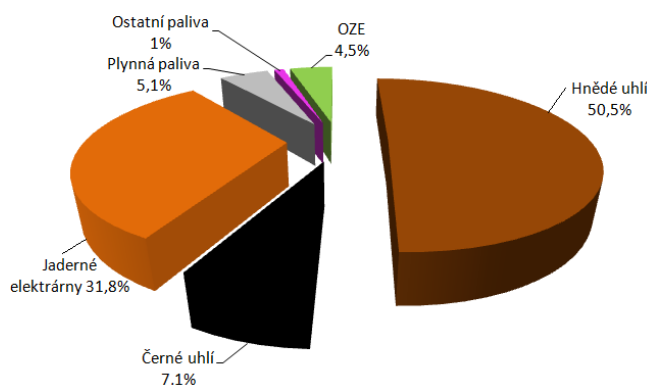


2 Motivace

V posledních letech se státy na celém světě potýkají s ekologickými problémy. Jedním z nejdiskutovanějších je globální oteplování způsobené nadměrnou produkcí skleníkových plynů při využívání fosilních paliv a z energetického hlediska je to vyčerpání světových zásob fosilních paliv (především uhlí) v horizontu desítek let. Řada států se kvůli globálnímu oteplování v rámci Kjótského protokolu zavázala ke snížení produkce skleníkových plynů (přestože oxid uhličitý CO_2 není jediným skleníkovým plynem, zaměřujeme se právě na něj). Jedním způsobem jak omezit tuto produkci a zároveň snížit závislost na fosilních palivech je využívání obnovitelných zdrojů energie (OZE). Fosilní paliva jako je uhlí totiž stále dominují v současných zařízeních na výrobu energie. Využívání fosilních paliv má negativní dopad na životní prostředí kvůli těžbě, zpracování a škodlivinám ve spalínách.

EU se zavázala snížit objem vyprodukovaných emisí skleníkových plynů o pětinu oproti produkci v roce 1990. Česká republika jako člen EU musí plnit závazky dané směrnicemi EU (2009/28/ES [1]). Závazek ČR ke splnění indikativního cíle, což je procentuální podíl OZE na hrubé domácí spotřebě elektřiny, byl ve výši 8% v roce 2010 a cíl na následující desetiletí (do roku 2020) je pro ČR 13% [1]. S nynějším nárůstem fotovoltaických elektráren se k cíli více přiblížíme, ale zatím se tyto cíle bohužel nedaří plnit, proto je zde snaha o využití dalších OZE, především biomasy.

V ČR využíváme k výrobě elektřiny energetický mix několika zdrojů. Na Obrázku 1 vidíme jejich poměr na výrobě hrubé domácí elektřiny v ČR. Polovinu palivového mixu tvoří hnědé uhlí, které se využívá ve velkých elektrárenských a teplárenských provozech. Více jak 30% tvoří jaderné palivo, které se používá ve dvou jaderných elektrárnách. Jak vidíme, OZE tvoří pouze 4,5% ze všech paliv, což není příliš uspokojivé.



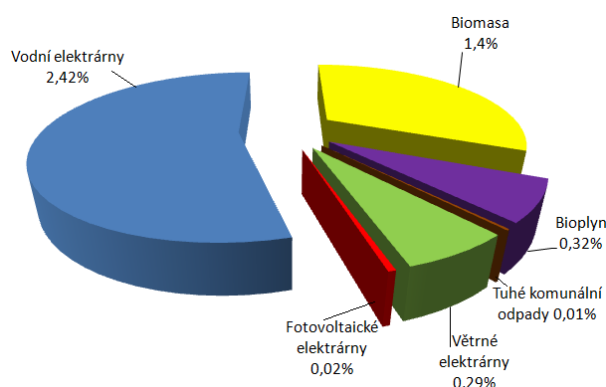
Obrázek 1: Výroba elektřiny podle zdrojů v ČR v roce 2008 [2]

Vzhledem k tomu, že biomasa patří mezi obnovitelné zdroje s vysokým potenciálem, budeme se zabývat právě tímto typem OZE.

2.1 Biomasa a její využití

Biomasa se rozumí hmota organického původu, kterou můžeme dělit na rostlinnou a živočišnou. V energetice se nejvíce používá z lesnického provozu dřevní štěpka, piliny a kůra, zatímco zemědělské provozy poskytují slámu, seno, kukuřici, řepku olejkou a v poslední době se rozšiřuje pěstování rychle rostoucích dřevin viz [4].

Biomasa má velký potenciál jako obnovitelný zdroj. Po vodních elektrárnách má největší podíl (Obrázek 2) a zřejmě je brzy překoná i díky dotacím státu.



Obrázek 2: Výroba elektřiny z OZE v ČR v roce 2008 [2]

Běžným způsobem energetického využití biomasy je přímé spalování. Dalším způsobem je zplyňování, kdy se energeticky využívá až vzniklý plyn. V současnosti je nejrozšířenější spalování biomasy s uhlím ve velkých teplárenských nebo elektrárenských kotlích. Budoucnost spalování v ČR závisí převážně na:

- ekonomických podmínkách (dotace za zelenou energii, ceny biomasy a uhlí, ...),
- dostupnosti biomasy.

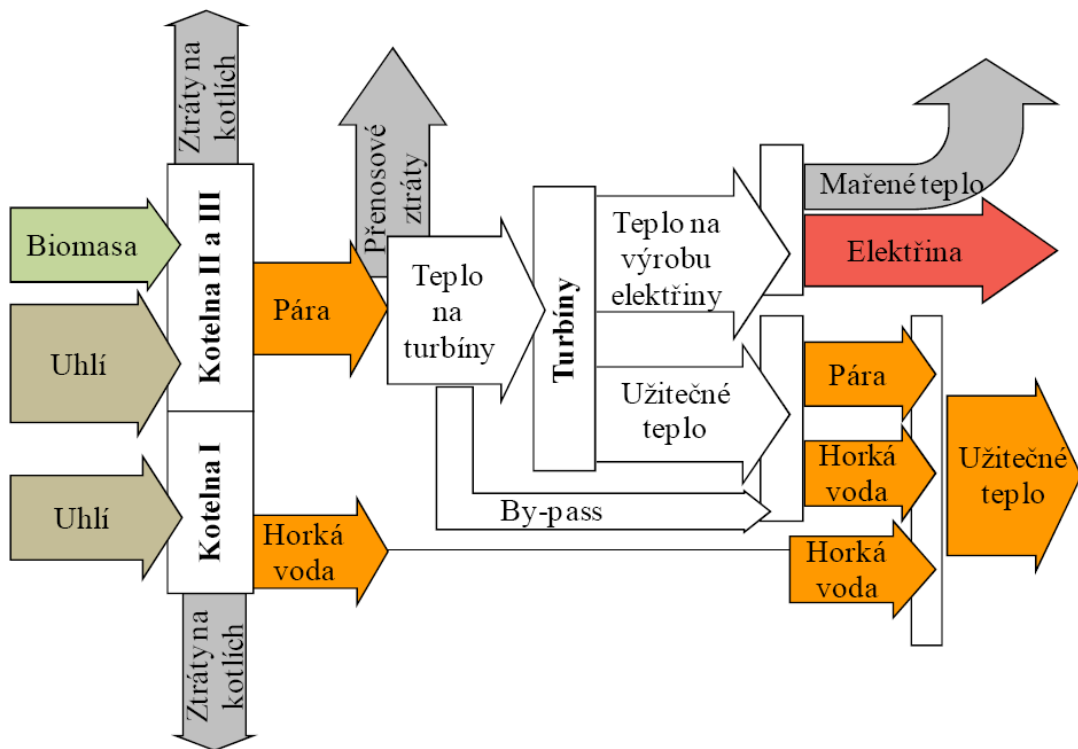
V současnosti by bez podpory státu spalování biomasy nebylo ekonomicky výhodné. Důvodem je vyšší cena pořízení biomasy, její nekonstantní vlastnosti (např. vlhkost), větší náklady na údržbu zařízení a možné snížení účinnosti kotlů při spalování.

Cílem práce je navrhnout ekonomicky efektivní využití biomasy při spalování s uhlím pro reálnou teplárnu, jejímž požadavkem je optimální využití stávající palivové základny a naplánování ročního provozu vzhledem k měsíční poptávce po teple.

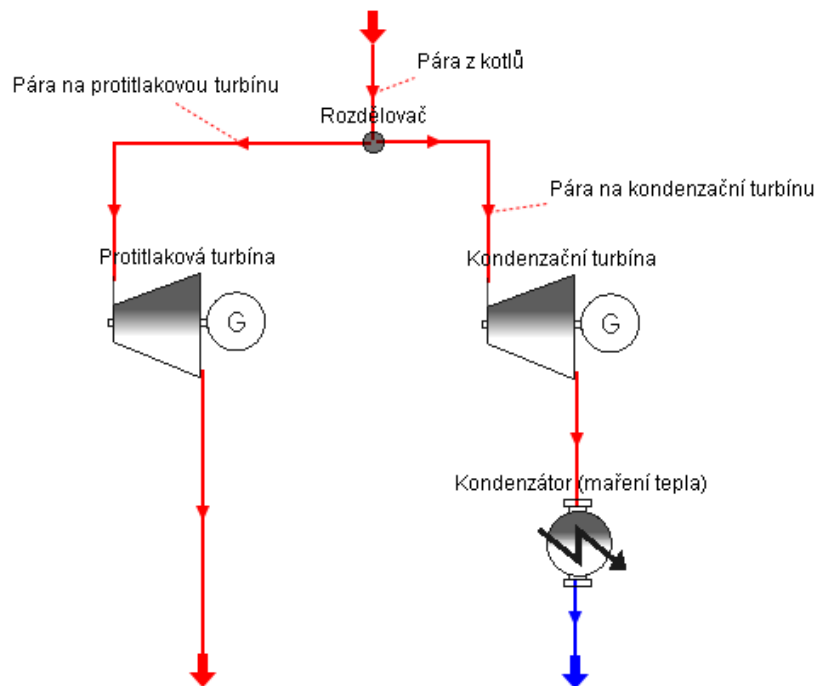
2.2 Popis teplárenského provozu

Teplárna zajišťuje dodávky tepla pro technologické účely, otop či ohřev topné a užitkové vody. Současně vyrábí také elektřinu, jde tedy o tzv. kogenerační výrobu elektřiny a tepla. Energie se získává spalováním paliv, v naší teplárně jde o uhlí a biomasu. Vyrobená pára z kotlů je přivedena na do strojovny, kde se na turbínách vyrábí teplo a elektřina. Ve strojovně jsou instalovány protitlaková parní turbína a odběrová kondenzační turbína (Obrázek 4). Po expanzi na protitlakové turbíně se pára využívá pro technologické účely a pro ohřev vody pro síť CZT (centrální zásobování teplem). Po expanzi na kondenzační turbíně je již energie v páře nevyužitelná a je vedena na vzduchový chladič (kondenzátor). Vzhledem k nedostatku informací nelze modelovat každou turbínu zvlášť a jsou tedy modelovány jako jedna turbína. V případě, kdy je to žádoucí, lze využít obtok (by-pass viz Obrázek 3) a použít tak pro výrobu tepla páru přímo z kotle. (Odbornou terminologii přebíráme z ÚPEI.)

Abychom mohli provést ekonomickou analýzu spalování biomasy a uhlí, musíme popsat vztahy, které se na tomto procesu podílí a vytvořit jejich matematický model. Ten musí odpovídat reálnému provozu, který je znázorněn na Obrázku 3.



Obrázek 3: Blokové schéma teplárny [4]



Obrázek 4: Zapojení turbín ve strojovně

Nejvíce se budeme zabývat kotli v kotelnách 2 a 3, protože zde je možnost spoluspalování biomasy s uhlím (kotel v kotelně 1 na to není uzpůsoben). Množství dávkované biomasy je ale omezeno použitou technologií, v případě kotleny 2 může biomasa zaujímat maximálně 10% celkové hmotnosti paliv a v případě kotleny 3 je to 30% z celkového příkonu v palivu (příkon v palivu je $\text{hmotnost} \times \text{výhřevnost}$). Řešení optimalizační úlohy nám určí množství dávkovaných paliv do jednotlivých kotlů, které nám zajistí splnění poptávky s maximálním ziskem.

3 Deterministický model teplárny

Model teplárny byl vytvořen s použitím reálných provozních dat. Předpokládá se, že množství dostupné biomasy je dáno pro celý rok, čímž dochází k propojení jednotlivých měsíců a úlohu nelze řešit pro každý měsíc v roce zvlášť. Protože některé z transformačních funkcí (např. popis přeměny paliva v energii při spalování) v modelu jsou nelineární, využíváme nelineární programování (NLP).

Model je realizován v systému GAMS, což je software využívaný pro modelování optimalizační úlohy. Nelineární optimalizační úlohy jsou řešeny pomocí zahrnutých řešičů problémů NLP (CONOPT, BARON). Pro uživatelsky příjemnou a komfortní práci se systémem GAMS a pro lepší prezentaci výsledků bylo vyvinuto excelovské rozhraní viz [4].

Ve své práci navazuji na již vytvořený deterministický model teplárny viz [3]. Abychom lépe pronikli do tohoto modelu, podrobně si ho popíšeme. Nejprve si zavedeme potřebné indexové množiny, poté všechna omezení a nakonec účelovou funkci.

3.1 Indexové množiny

- Písmenem J označíme množinu všech paliv, $j \in J$ označuje konkrétní druh paliva a $|J| = n_J$ je mohutnost množiny typů paliv. Množina paliv J zahrnuje všechny druhy paliv, které jsou použity v provozu. Použitá paliva jsou fosilní (uhlí) a biomasová paliva, jako např. dřevní štěpka, mláto (zbytky z výroby piva), řepné řízky (zbytky z výroby cukru), řepkový šrot (zbytky z výroby oleje), slunečnicový šrot (zbytky z výroby oleje) a podsítná frakce mulčovací kůry.
- Nyní zavedeme množinu obsahující kotelny K . Jednotlivé kotelny označíme $k \in K$ a $|K| = n_K$. Tento provoz má tři kotelny, v první se dává pouze uhlí a slouží spíše jako pomocná v době, kdy poptávku nepokryje výroba v druhé a třetí. V druhé a třetí se může dávkovat uhlí i biomasa, jejich poměr je však omezen.
- Množina T obsahuje časová období braná po měsících. t značí jednotlivé měsíce $t \in T$ a počet období $|T| = n_T$. V našem případě uvažujeme roční provoz, $n_T = 12$.
- Písmenem I označíme množinu turbín a navíc i redukční stanici. Prvky množiny $i \in I$ označují konkrétní turbíny a znovu mohutnost této množiny označíme $|I| = n_I$.
- Poslední množinu označíme M , která zahrnuje typy vyprodukovaných energií (elektřina a teplo). Její mohutnost označíme $|M| = n_M$.

3.2 Omezení

Vektor $\mathbf{x} = (x_{jk})_{j \in J, k \in K}$ (složka se dvěma indexy) označuje vstupní množství paliv dávkovaných do kotlů. Poměr vstupních paliv je omezen stávajícími technologiemi a paliva jsou omezena velikostmi zdrojů a navíc každý typ paliva má různou výhřevnost.

Indexy s velkými písmeny používáme pro vyjádření součtů, např. $\mathbf{x}_J = (x_{J,k})_{k \in K}$, kde $x_{J,k} = \sum_{j \in J} x_{jk}$.

Matice $\mathbf{E}_x = (E_{x,ijk})_{i \in J, j \in J, k \in K}$, kde $E_{x,ijk} \in \{0, 1\}$, nám vyjadřuje, které typy paliv jsou dávkovány do konkrétních kotlů v daném časovém období (sloupce mají dva indexy). Vektor $\mathbf{b}_x = (b_{x,j})_{j \in J} \in \mathfrak{R}^{n_J}$ označuje dostupné množství všech druhů paliv. Nulový vektor označíme $\mathbf{0}$ a vektor jedniček je označen $\mathbf{1}$, dimenze je určena kontextem (podle použití). Dále $\mathbf{b}_L = (b_{L,j,k})_{j \in J, k \in K} \in \mathfrak{R}^{n_J \cdot n_K}$ je vektor, který vyjadřuje technologicky nejnižší možné množství vstupů (dávkovaného paliva do kotlů) a také určuje poměr mezi vstupními palivy (sloupce mají dva indexy). Naproti tomu vektor $\mathbf{b}_U = (b_{U,j,k})_{j \in J, k \in K} \in \mathfrak{R}^{n_J \cdot n_K}$ definuje technologicky maximální možný vstup a znovu určuje poměr mezi vstupními palivy. Zde jsou omezení pro vstup:

$$\begin{aligned} \mathbf{E}_x \mathbf{x} &\leq \mathbf{b}_x, \\ \mathbf{0} &\leq \mathbf{x} \in \mathfrak{R}^{n_J \cdot n_K}, \\ \mathbf{b}_L \mathbf{1}^T \mathbf{x} &\leq \mathbf{x} \leq \mathbf{b}_U \mathbf{1}^T \mathbf{x}. \end{aligned}$$

Vektor označující energii, která vstupuje do kotlů, (složka se dvěma indexy) označíme jako $\mathbf{u} = (u_{jk})_{j \in J, k \in K}$. V modelu označuje $\alpha = (\alpha_i)_{i \in \Delta_\alpha} \in \mathfrak{R}^{n_\alpha}$ vektor koeficientů výhřevností jednotlivých typů paliv (LHV), $|\Delta_\alpha| = n_\alpha$. Funkce $\mathbf{f} : \mathfrak{R}^{n_J \cdot n_K} \times \mathfrak{R}^{n_\alpha} \rightarrow \mathfrak{R}^{n_J \cdot n_K}$ popisuje přeměnu paliva na energii (spalování). Obdobně jako pro paliva zavedeme matici $\mathbf{E}_u = (E_{u,\kappa jk})_{\kappa \in K, j \in J, k \in K}$, kde $E_{u,\kappa jk} \in \{0, 1\}$. Tato matice nám udává, které typy paliv jsou přeměněny na energii (sloupce mají dva indexy). Zavedeme si mezní vektor $\mathbf{b}_C = (b_{C,k})_{k \in K} \in \mathfrak{R}^{n_K}$, který nám vyjadřuje maximální množství vstupující energie, toto množství je určeno technologií kotlů. Omezení plynoucí z technologie kotlů a vlastností spalovacího procesu jsou tvaru:

$$\begin{aligned} \mathbf{u} &= \mathbf{f}(\mathbf{x}, \alpha), \\ \mathbf{0} &\leq \mathbf{u} \in \mathfrak{R}^{n_J \cdot n_K}, \\ \mathbf{E}_u \mathbf{u} &\leq \mathbf{b}_C. \end{aligned}$$

Vektor $\mathbf{v} = (v_{jk})_{j \in J, k \in K}$ je vektor množství páry vystupující z kotlů (složky mají dva indexy) a $\eta = (\eta_i)_{i \in \Delta_\eta} \in \mathfrak{R}^{n_\eta}$ je vektor vyjadřující účinnosti kotlů, $|\Delta_\eta| = n_\eta$. Potom funkce $\mathbf{g} : \mathfrak{R}^{n_J \cdot n_K} \times \mathfrak{R}^{n_\eta} \rightarrow \mathfrak{R}^{n_J \cdot n_K}$, popisuje přeměnu energie v páru (vyprodukované teplo ohřívá vodu, která se přemění v páru). Vektor $\mathbf{e}_v = (e_{v,jk})_{j \in J, k \in K}$, kde $e_{v,jk} \in \{1, 0\}$ udává, které výstupy z kotlů míří do strojovny na výrobu elektrické energie, a které přímo na užitkové teplo, navíc také zahrnuje deterministické ztráty (v našem případě přibližně 4%) v síti mezi kotly a strojovnou (turbíny). Nyní označíme vektor $\mathbf{e}_w = (e_{w,jki})_{j \in J, k \in K, i \in I}$, kde $e_{w,ijk} \in \{1, 0\}$, jako vektor určující, které vstupy jdou na kterou turbínu. Nyní označíme $\mathbf{w} = (w_{jki})_{j \in J, k \in K, i \in I} \in \mathbf{W}$ jako vektor energetických toků ve strojovně (složky se třemi indexy), definovaný pro detailní popis a maximální flexibilitu. Následující rovnice jsou rovnice transformace energie v páru v jednotlivých kotlích a bilanční rovnice pro výstupy z kotlů a vstupy do turbín:

$$\begin{aligned} \mathbf{v} &= \mathbf{g}(\mathbf{u}, \eta), \\ \mathbf{0} &\leq \mathbf{v} \in \mathfrak{R}^{n_J \cdot n_K}, \\ \mathbf{e}_v^T \mathbf{v} &= \mathbf{e}_w^T \mathbf{w}, \\ \mathbf{0} &\leq \mathbf{w} \in \mathbf{W} \in \mathfrak{R}^{n_J \cdot n_K \cdot n_I}. \end{aligned}$$

Nyní zavedeme vektor $\mathbf{y} \in \mathfrak{R}^{n_J \cdot n_K \cdot n_I \cdot n_M}$, který popisuje vyrobené energie (složky se čtyřmi indexy). Tento vektor je definován pro detailní popis procesu, ale velmi často indexy označující turbíny a kotle nejsou příliš důležité, proto používáme sčítací indexy (jak jsme již

uvedli na začátku této části) $\mathbf{y}_{KI} = (y_{KI,jm})_{j \in J, m \in M}$, kde $y_{KI,jm} = \sum_{k \in K} \sum_{i \in I} y_{kijm}$. Matice $\mathbf{E}_w = (E_{w,i\kappa\lambda\theta jkim})_{i,j \in J, \kappa, k \in K, \lambda, i \in I, \theta, m \in M}$, kde $E_{w,i\kappa\lambda\theta jkim} \in \{0, 1\}$ je matice udávající, které výstupní energie se vyrábí ve strojovně, a které ne (prvky matice mají čtyři indexy). Zavedeme si další vektor γ , který interpretuje účinnost turbín, $\gamma = (\gamma_i)_{i \in \Delta_\gamma} \in \mathfrak{R}^{n_\gamma}$, kde $|\Delta_\gamma| = n_\gamma$. Máme zobrazení vektorové množiny \mathbf{H} umožňující popis strojovny $\mathbf{H} : \mathbf{W} \times \mathfrak{R}^{n_\gamma} \rightarrow \mathfrak{S} \subset 2^{\mathfrak{R}^{n_J \cdot n_K \cdot n_I \cdot n_M}}$, kde $2^{\mathfrak{R}^{n_J \cdot n_K \cdot n_I \cdot n_M}}$ je množina všech podmnožin množiny $\mathfrak{R}^{n_J \cdot n_K \cdot n_I \cdot n_M}$. Nyní si zavedeme matici $\mathbf{E}_v = (E_{v,i\kappa\theta jkim})_{i,j \in J, \kappa, k \in K, \lambda, i \in I, \theta, m \in M}$, jejíž prvky $E_{v,i\kappa\theta jkim} \in \{0, 1\}$ (prvky matice mají čtyři indexy). Tato matice nám udává, které energetické výstupy vychází přímo z kotlů (neprochází turbínami). Matice obvykle obsahuje jedničky takovým způsobem, že počítá složky vektoru \mathbf{y} , je-li \mathbf{y}_{JKI} použito, potom se několik jedniček objeví pouze na hlavní diagonále a ostatní hodnoty a i většina prvků hlavní diagonály jsou rovny nule. Zavedeme matici $\mathbf{E}_K = (E_{K,i\kappa\theta jkim})_{i,j \in J, \kappa, k \in K, \lambda, i \in I, \theta, m \in M}$, jejíž prvky $E_{K,i\kappa\theta jkim} \in \{0, 1\}$, (sloupce matice mají čtyři indexy). Tato matice nám udává, které energetické výstupy vychází přímo z kotlů (neprochází turbínami). Matice obvykle obsahuje jedničky takovým způsobem, že počítá složky vektoru \mathbf{y} , je-li \mathbf{y}_{JKI} použito, potom se několik jedniček objeví pouze na hlavní diagonále a ostatní hodnoty a i většina prvků hlavní diagonály jsou rovny nule. Jednotkovou matici standardně označíme \mathbf{I} , rozměr matice je opět určen kontextem. Zápis sgn nám udává funkci signum, která přiřazuje nulovou hodnotu na nulu, kladným hodnotám přiřadí $+1$ a záporným hodnotám -1 . Pokud funkci sgn aplikujeme na vektor, aplikujeme tuto funkci na každou složku vektoru zvlášť. To je užitečné pro popsání ztrát v síti, které se uvažují pouze mezi kotli a strojovnou. Označíme diag jako vektorovou funkci, která vytváří čtvercové matice a dává vektor v argumentu na hlavní diagonálu. Pak související omezení jsou uvedena takto:

$$\begin{aligned} \mathbf{E}_w \mathbf{y} &\in \{\mathbf{H}(\mathbf{w}, \gamma), \mathbf{w} \in W\}, \\ \mathbf{0} &\leq \mathbf{y} \in \mathfrak{R}^{n_J \cdot n_K \cdot n_I \cdot n_M}, \\ \mathbf{E}_v \mathbf{y} &= \mathbf{E}_K (\mathbf{I} - \text{diag}(\text{sgn} \mathbf{e}_v))^T \mathbf{v}. \end{aligned}$$

Posledním omezením je omezení velikosti poptávky. Nyní si zavedeme matici $\mathbf{E}_D = (E_{D,\theta jkim})_{j \in J, k \in K, i \in I, \theta, m \in M}$, jejíž prvky $E_{D,\theta jkim} \in \{0, 1\}$. Matice \mathbf{E}_D je matice vyjadřující kombinaci vyrobených energií z různých zdrojů (pomocí součtu) pro splnění požadované poptávky (řádky mají jeden index a sloupce mají čtyři indexy). Potom $\mathbf{b}_{DL} = (b_{DL,m})_{m \in M} \in \mathfrak{R}^{n_M}$ je vektor dolní hranice možné poptávky, naopak vektor $\mathbf{b}_{DU} = (b_{DU,m})_{m \in M} \in \mathfrak{R}^{n_M}$ je vektor horní hranice možné poptávky, omezení je tedy tvaru

$$\mathbf{b}_{DL} \leq \mathbf{E}_D \mathbf{y} \leq \mathbf{b}_{DU}.$$

Toto jsou všechna omezení optimalizačního modelu naší teplárny. Nyní máme nadefinované použité množiny indexů, omezení a zbývá už nám pouze účelová funkce.

3.3 Účelová funkce

Funkci popisující zisk z prodeje tapla a elektřiny označíme $d_D : \mathfrak{R}^{n_J \cdot n_K \cdot n_I \cdot n_M} \rightarrow \mathfrak{R}$. Kromě toho, vektor $\pi = (\pi_i)_{i \in \Delta_\pi} \in \mathfrak{R}^{n_\pi}$ je vektor časového parametru účelové funkce, $|\Delta_\pi| = n_\pi$. Funkce $d_G : \mathfrak{R}^{n_J \cdot n_K \cdot n_I \cdot n_M} \rightarrow \mathfrak{R}$ popisuje zisk z dotací od státu. Potom $\mu = (\mu_i)_{i \in \Delta_\mu} \in \mathfrak{R}^{n_\mu}$ je vektor časového parametru účelové funkce, $|\Delta_\mu| = n_\mu$ a $c : \mathfrak{R}^{n_J \cdot n_K} \rightarrow \mathfrak{R}$ je funkce popisující výdaje za nákup paliv a provoz teplárny, také $\rho = (\rho_i)_{i \in \Delta_\rho} \in \mathfrak{R}^{n_\rho}$ je vektor časového parametru účelové funkce, $|\Delta_\rho| = n_\rho$. Pak účelová funkce je tvaru

$$\max_{\mathbf{x}, \mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{y}} \{d_D(\mathbf{y}, \pi) + d_G(\mathbf{y}, \mu) - c(\mathbf{x}, \rho)\}.$$

3.4 Model

Předchozí model, tj. účelová funkce a omezení, zapíšeme ve tvaru optimalizační úlohy:

$$\begin{aligned}
& \max_{\mathbf{x}, \mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{y}} \{d_D(\mathbf{y}, \pi) + d_G(\mathbf{y}, \mu) - c(\mathbf{x}, \rho)\} \\
& \mathbf{E}_x \mathbf{x} \leq \mathbf{b}_x, \\
& \mathbf{0} \leq \mathbf{x} \in \mathfrak{R}^{n_J \cdot n_K}, \\
& \mathbf{b}_L \mathbf{1}^T \mathbf{x} \leq \mathbf{x} \leq \mathbf{b}_U \mathbf{1}^T \mathbf{x}, \\
& \mathbf{u} = \mathbf{f}(\mathbf{x}, \alpha), \\
& \mathbf{0} \leq \mathbf{u} \in \mathfrak{R}^{n_J \cdot n_K}, \\
& \mathbf{E}_u \mathbf{u} \leq \mathbf{b}_C, \\
& \mathbf{v} = \mathbf{g}(\mathbf{u}, \eta), \\
& \mathbf{0} \leq \mathbf{v} \in \mathfrak{R}^{n_J \cdot n_K}, \\
& \mathbf{e}_v^T \mathbf{v} = \mathbf{e}_w^T \mathbf{w}, \\
& \mathbf{0} \leq \mathbf{w} \in W \in \mathfrak{R}^{n_J \cdot n_K \cdot n_I}, \\
& \mathbf{E}_w \mathbf{y} \in \{\mathbf{H}(\mathbf{w}, \gamma), \mathbf{w} \in W\}, \\
& \mathbf{0} \leq \mathbf{y} \in \mathfrak{R}^{n_J \cdot n_K \cdot n_I \cdot n_M}, \\
& \mathbf{E}_v \mathbf{y} = \mathbf{E}_K (\mathbf{I} - \text{diag}(\text{sgn } \mathbf{e}_v))^T \mathbf{v}, \\
& \mathbf{b}_{DL} \leq \mathbf{E}_D \mathbf{y} \leq \mathbf{b}_{DU}.
\end{aligned}$$

V tomto modelu nejsou zahrnuty neurčitosti a jistá míra nejistoty. Za nejdůležitější neurčitosti modelu považujeme výši poptávky po teple, výši dotace od státu a určení závislosti účinnosti kotlů na daných veličinách. Naším cílem je co nejpřesněji popsat a do modelu doplnit tyto neurčitosti a tím lépe vystihnout reálný proces dané teplárny.

Dále si v práci zavedeme teorii, kterou využíváme k popsání neurčitostí, nejprve základy stochastického programování a poté regresní analýzu.

4 Matematické programování

Optimalizační úloha se snaží o výběr toho nejlepšího prvku z nějakého souboru dostupných alternativ, tedy jde o snahu nalézt takové hodnoty proměnných, pro které daná účelová funkce nabývá optimální hodnoty, při splnění požadovaných omezení. Rozlišujeme dva typy optimalizační úlohy, a to minimalizační nebo maximalizační. Úloha optimalizace je někdy nazývána též úlohou *matematického programování*.

Funkci $f : \mathbb{R}^n \rightarrow \mathbb{R}$, označíme jako *účelovou funkci*, pak v minimalizační (resp. maximalizační) úloze hledáme takové řešení \mathbf{x}_{opt} , které odpovídá podmínce $\mathbf{x}_{opt} \in C$ a pro všechna $\mathbf{x} \in C$, kde $C \subset \mathbb{R}^n$ platí, že $f(\mathbf{x}_{opt}) \leq f(\mathbf{x})$ (resp. $f(\mathbf{x}_{opt}) \geq f(\mathbf{x})$). Množinu C označíme jako *množinu přípustných řešení*. Tuto množinu získáme díky zadaným *omezením*, které nám v prostoru \mathbb{R}^n tuto množinu vymezí viz [11].

Obecně můžeme minimalizační úlohu zapsat ve tvaru

$$\begin{aligned} &\text{minimalizuj} && f(\mathbf{x}), \\ &\text{za podmíněk} && g_i(\mathbf{x}) \circ 0, \quad i = 1, 2, \dots, n, \\ &&& \mathbf{x} \in C, \end{aligned}$$

kde $f : \mathbb{R}^n \rightarrow \mathbb{R}$ je účelová funkce, kterou minimalizujeme a funkce $g_1 : \mathbb{R}^n \rightarrow \mathbb{R}, \dots, g_n : \mathbb{R}^n \rightarrow \mathbb{R}$ se nazývají omezení nebo podmínky. Množinu C nazveme množinou přípustných řešení

$$C = \{\mathbf{x} \mid g_1(\mathbf{x}) \circ 0, \dots, g_n(\mathbf{x}) \circ 0\},$$

kde $\circ \in \{\leq, =, \geq\}$. Z této množiny vybíráme naše hledané \mathbf{x}_{opt} . Dále budeme používat zápis $\mathbf{g}(\mathbf{x}) = (g_1, \dots, g_n)^T$.

Maximalizační úlohu můžeme převést na minimalizační transformací účelové funkce následovně

$$\max_x \{f(\mathbf{x}) \mid \mathbf{x} \in C\} = - \min_x \{-f(\mathbf{x}) \mid \mathbf{x} \in C\}.$$

Matematické programování rozdělím z pohledu práce na dva hlavní přístupy, a to podle znalosti dat na

- deterministické programování (lineární, nelineární),
- stochastické programování.

4.1 Deterministické programování

V této sekci popíšeme co je deterministické programování. *Deterministický program* je matematický program, kde jsou všechny údaje plně známy. Pod pojmem "údaje" máme na mysli všechny parametry a koeficienty problému, který řešíme. Jinými slovy, v deterministickém modelu není žádná nejistota nebo náhodnost. Toto je nejvýznamnější rozdíl v porovnání se stochastickým programováním. Deterministický program je tvaru

$$\begin{aligned} &\text{minimalizuj} && f(\mathbf{x}, \mathbf{a}), \\ &\text{za podmíněk} && \mathbf{g}(\mathbf{x}, \mathbf{a}) \leq \mathbf{0}, \\ &&& \mathbf{h}(\mathbf{x}, \mathbf{a}) = \mathbf{0}, \\ &&& \mathbf{x} \in C, \end{aligned} \tag{1}$$

kde, $\mathbf{a} \in \mathbb{R}^k$ je k -dimenzionální vektor konstant. Použitím vektoru \mathbf{a} jsme zdůraznili, že všechny parametry jsou konstantní a zcela známé.

Zvláštním případem tohoto programu je úloha, která má všechny komponenty (účelová funkce a omezení) lineární a C je konvexní mnohostěn. Takový program se nazývá lineární deterministický program. Jeho forma je

$$\begin{array}{ll} \text{minimalizuj} & \mathbf{c}^T \mathbf{x} \\ \text{za podmínek} & \mathbf{A}\mathbf{x} = \mathbf{b}, \\ & \mathbf{x} \in C, \end{array}$$

kde \mathbf{c} je N -dimenzionální vektor, nazývaný vektor cen, \mathbf{A} je matice typu $n \times N$ a \mathbf{b} je n -dimenzionální vektor pravých stran. V tzv. standardní formě lineárního programu je požadavek na nezápornost vektoru \mathbf{x} ($\mathbf{x} \geq 0$) přidána k předchozímu modelu viz [12].

Nejvíce pracujeme s deterministickými "reformulacemi" (ekvivalenty), které nám pomáhají správně interpretovat náhodné parametry. Všechny tyto programy, které zahrnují náhodné parametry, se nazývají stochastické programy.

4.2 Stochastické programování

V předchozím oddílu jsme diskutovali koncepci deterministického programování a jeho model. V praktických problémech a reálných aplikacích je použití deterministických modelů velmi omezující, protože reálné modely obsahují parametry, které nejsou zcela známy. Jinými slovy, reálné problémy obsahují určitou míru nejistoty a použití deterministických modelů může vést ke zkreslení nebo dokonce ke zcela nesprávným výsledkům.

Obecně platí, že k modelování problémů obsahujících nejistoty mohou být použity odlišné přístupy. Jeden z nich je přístup *stochastického programování*, ve kterém jsou nejisté parametry modelovány generováním náhodných veličin.

Definice: Necht' (Ω, \mathcal{A}) je jevové pole a zobrazení $\xi : \Omega \rightarrow \mathbb{R}$ je takové, že platí $\{w : \xi(w) \leq x\} \in \mathcal{A}$ pro $\forall x \in \mathbb{R}$. Pak ξ nazveme náhodnou veličinou vzhledem k jevovému poli (Ω, \mathcal{A}) .

Obecně můžeme úlohu stochastického programování zapsat ve tvaru

$$\begin{array}{ll} \text{minimalizuj} & f(\mathbf{x}, \xi), \\ \text{za podmínek} & \mathbf{g}(\mathbf{x}, \xi) \leq \mathbf{0}, \\ & \mathbf{h}(\mathbf{x}, \xi) = \mathbf{0}, \\ & \mathbf{x} \in C(\xi), \end{array} \quad (2)$$

kde $\xi = (\xi_1, \dots, \xi_k)^T$, $\xi : \Omega \rightarrow \mathbb{R}$, je konečně rozměrný náhodný vektor tvořený náhodnými veličinami vzhledem k jevovému poli (Ω, \mathcal{A}) .

Všimněme si, že model (2) je v podstatě deterministický model (1), ve kterém některé konstantní parametry byly nahrazeny náhodnými veličinami.

Naším cílem je najít řešení modelu (2). Hlavním problémem je, ve kterém časovém okamžiku budeme moci provést rozhodnutí. Povaha reálného problému obvykle určí, jak a kdy se rozhodujeme. Máme dva možné přístupy, buď rozhodneme před tím (rozhodovatel obvykle nemůže čekat), než je náhodný parametr ξ pozorován. Tento přístup se nazývá Here-and-Now. Nebo můžeme rozhodnout až po pozorování ξ , kdy už je jeho hodnota známá, pak se tento přístup nazývá Wait-and-See.

4.2.1 Wait-and-See přístup

Tento přístup k stochastickému programování předpokládá znalost hodnot náhodných parametrů, které jsou v optimalizačním modelu zahrnuty. Místo Wait-and-See přístup budeme psát pouze WS přístup.

WS přístup používáme, když známe realizace ξ , než učiníme naše rozhodnutí, což předpokládá dokonalé informace o budoucnosti. V tomto případě můžeme změnit naše rozhodnutí díky pozorování, a tedy, rozhodnutí \mathbf{x} chápeme jako funkci $\mathbf{x}(\xi)$ náhodného vektoru ξ . Také účelovou funkci $f(\mathbf{x}(\xi), \xi)$ budeme chápat jako náhodnou veličinu. Tento přístup má svůj význam zvláště pro dlouhodobé plánování. Jeho tvar je

$$\begin{aligned} & \text{minimalizuj} && f(\mathbf{x}(\xi), \xi) \\ & \text{za podmíněk} && \mathbf{g}(\mathbf{x}(\xi), \xi) \leq \mathbf{0}, \\ & && \mathbf{h}(\mathbf{x}(\xi), \xi) = \mathbf{0}, \\ & && \mathbf{x}(\xi) \in C(\xi). \end{aligned}$$

WS deterministický ekvivalent programu (2), můžeme zapsat ve tvaru

$$z^{WS} = E_{\xi}(\min_{\mathbf{x}(\xi)} \{f(\mathbf{x}(\xi), \xi) \mid \mathbf{x}(\xi) \in C(\xi) \forall \xi\})$$

a jeho řešení označíme \mathbf{x}_{opt}^{WS} viz [12].

4.2.2 Here-and-Now přístup

Někdy bohužel výše uvedený WS přístup nelze použít, protože nemůžeme čekat na konkrétní realizace ξ a musíme tedy rozhodnout bez jeho znalosti. Rozhodnutí \mathbf{x} je tedy stejné pro všechny možné budoucí realizace ξ .

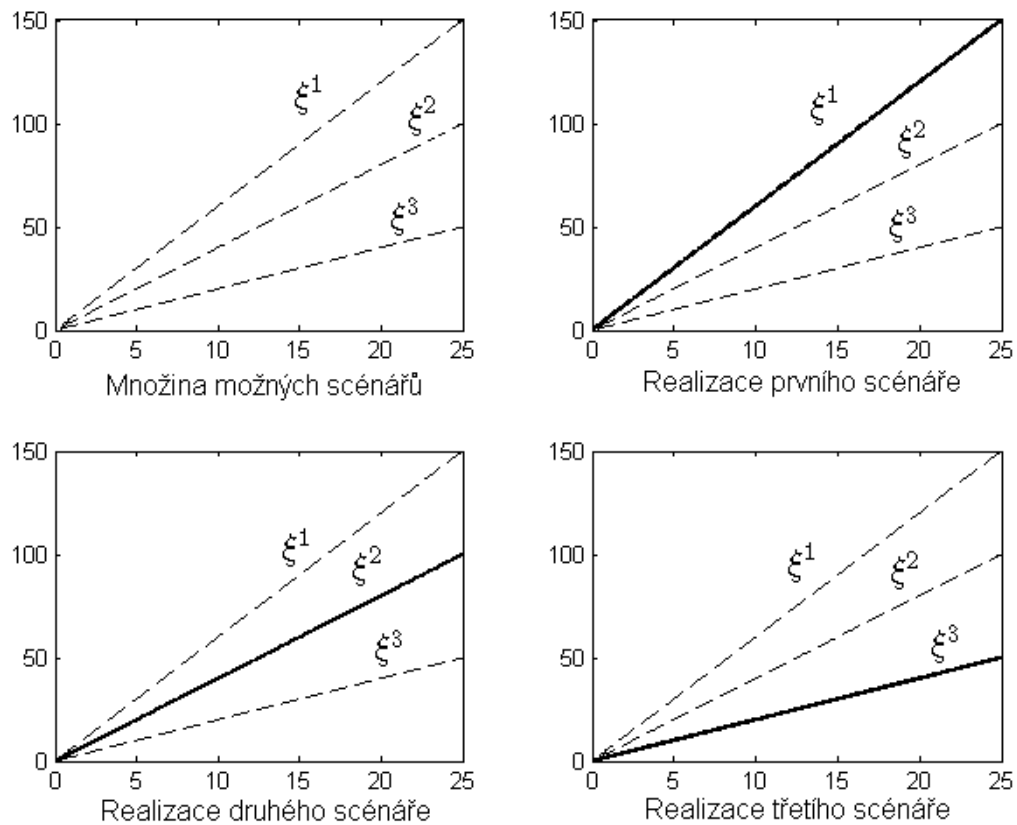
4.2.3 Scénářové úlohy

Tento přístup je založen na myšlence, že náhodné parametry ξ v základním programu (2) jsou nahrazeny jejich typickou realizací ξ^s . Pojem náhodné veličiny na jevovém poli zahrnutý v modelu nám umožňuje se zabývat jejím pravděpodobnostním rozdělením místo konstantních parametrů, jako je tomu v případě deterministického programování. Budeme předpokládat, že pravděpodobnostní rozdělení ξ je zcela známé.

Pokud konkrétní realizace náhodných parametrů ξ^s jsou pozorovány a stanou se známé nebo je možné odhadnout právě pomocí znalosti jejich rozdělení, můžeme nahradit ξ v (2) známým ξ^s . Takové konkrétní realizace ξ^s se nazývají *scénáře*. Proto program s jednotlivými scénáři (individual scenario programme - IS program) má tvar

$$\begin{aligned} & \text{minimalizuj} && f(\mathbf{x}, \xi^s) \\ & \text{za podmíněk} && \mathbf{g}(\mathbf{x}, \xi^s) \leq \mathbf{0}, \\ & && \mathbf{h}(\mathbf{x}, \xi^s) = \mathbf{0}, \\ & && \mathbf{x} \in C(\xi^s), \end{aligned}$$

Optimální řešení je označováno \mathbf{x}_{opt}^{IS} . Tento IS-program je užitečný v případě, že odborník je schopen určit typické realizace ξ viz [12]. Princip práce se scénáři je znázorněn na Obrázku 5.



Obrázek 5: Princip scénářů

Stochastické programování využíváme k modelování plánování teplárenského provozu, kde se lze setkat s několika náhodnými parametry. Jsou to neznámé velikosti poptávky po teple a vývoj cen dotací za vyrobenou elektřinu z OZE. Tyto parametry jsou podrobněji popsány v kapitole 7. Pro určení a popsání scénářů a také pro zpracování dodaných provozních dat si zavedeme v následující kapitole regresní analýzu.

5 Jednorozměrná regrese

Další potřebnou teorií je regresní analýza, jelikož máme k dispozici data z provozu, ze kterých se snažíme odvodit jejich závislosti, které použijeme ke zpřesnění modelu (kapitola 6). Využíváme jednorozměrnou i vícerozměrnou regresi. Tento text byl psán s přihlédnutím na další možné využití regresní analýzy na ÚPEI. Nejprve si stručně uvedme, čím se zabývá regresní analýza. Vycházíme z literatury [5] - [10], [14] - [20].

Tato část statistické analýzy se zabývá hledáním, zkoumáním a ověřováním závislostí mezi danými proměnnými. Hodnoty proměnných získáváme z měření reálných procesů nebo různých experimentů. Proměnné dělíme na dvě části: *nezávisle proměnné* (vysvětlující), které jsou reprezentovány nenáhodným vektorem nebo maticí \mathbf{X} a *závisle proměnné* vyjádřené nejčastěji náhodným vektorem Y . Závislost Y na \mathbf{X} vyjadřujeme pomocí tzv. *regresní funkce* η

$$y = \eta(\mathbf{x}, \beta),$$

kde $\mathbf{x} = (x_1, \dots, x_K)$ je vektor nezávisle proměnných (hodnota nenáhodného vektoru \mathbf{X}), y je závisle proměnná (hodnota náhodné veličiny Y) a β je vektor parametrů, které nazýváme *regresní koeficienty* viz [9].

Vektor Y nemusí záviset pouze na daných proměnných X_1, \dots, X_K , mohou ho ovlivňovat i další proměnné X_{K+1}, X_{K+2}, \dots , které ale mohou mít jen zanedbatelný vliv. Do modelu tedy zahrnujeme pouze proměnné, které mají zásadní vliv na hodnoty Y . Nezařazené vlivy způsobí, že každý předpis regresní funkce je zatížen chybou. Nazýváme ji *náhodnou chybou* (nebo také rušivou složkou) a označujeme ji ε .

První otázkou je, kolik máme vysvětlujících proměnných. Proto se zabýváme regresní analýzou jednorozměrnou a vícerozměrnou.

Pojem jednorozměrná regrese se používá, pokud máme pouze jednu vysvětlující proměnnou X . A hledáme závislost náhodného vektoru Y na X . Máme n daných hodnot pro obě proměnné.

5.1 Výběr regresní funkce

Pro regresní analýzu je nezbytné znát tvar regresní funkce, což bývá jedním z největších problémů regresní analýzy. Snažíme se jej volit tak, aby co nejvíce odpovídal hledané závislosti proměnných a zároveň měl co nejmenší počet regresních koeficientů β .

V regresním modelu musíme také rozhodnout, zda náhodná chyba ε je v aditivním $Y = \eta + \varepsilon$ nebo multiplikativním tvaru $Y = \eta\varepsilon$.

Při hledání η vycházíme z podrobné znalosti zkoumaného problému (známe nebo aspoň předpokládáme tvar závislosti), ze zkušeností nebo v současné době se při hledání regresní funkce dají použít vhodné databáze regresních funkcí na PC viz [20]. V případě jednorozměrné regrese nám hodně napoví vykreslený bodový graf proměnných $[x_i, y_i]$, $i = 1, \dots, n$.

5.1.1 Typy regresních funkcí

Regresní funkce se dělí na dva hlavní typy:

• Lineární regresní funkce

Linearitou se myslí lineární vzhledem k regresním koeficientům, ne k proměnným. Ukážeme nejnámější a nejpoužívanější druhy lineárních modelů:

Zcela lineární model

Ve zcela lineárním modelu (terminologie viz [6]) se používá aditivní vliv chyby i všech vysvětlujících proměnných, regresní funkce je funkcí nadroviny

$$Y = \beta_0 + \beta_1 X + \varepsilon,$$

kde β_0 je tzv. absolutní člen.

Racionální celistvé a lomené funkce

V této skupině jsou nejpoužívanější modely regresní paraboly nebo hyperboly s-tého stupně, popřípadě jejich kombinace

$$Y = \beta_0 + \beta_1 X^{\pm 1} + \beta_2 X^{\pm 2} + \dots + X^{\pm s} + \varepsilon.$$

Model lineární v parametrech

Je zobecněním předchozích druhů. Regresní funkce je tvaru

$$Y = \beta_0 + \beta_1 f_1(X) + \dots + \beta_R f_R(X) + \varepsilon,$$

kde f_i , $i = 1, \dots, R$ jsou libovolné neznámá funkce vysvětlující proměnné, které neobsahují žádné další neznámé parametry.

Př: $Y = \beta_0 + \beta_1 X^2 + \beta_2 \ln(X) + \beta_3 X + \varepsilon$.

• Nelineární regresní funkce

Lineární modely jsou pro svou jednoduchost více oblíbené, ale reálné vztahy mezi veličinami jsou dost často daleko složitější, proto je někdy nutné použít nelineární regresní funkci, kdy nelinearitou myslíme nelinearitu v regresních koeficientech.

Pro představu uvedeme příklady nelineárních regresních funkcí.

Př: $Y = \sqrt{\beta_0 + \beta_1 X} + \varepsilon$

$$Y = \beta_0 X^{\beta_1} + \varepsilon$$

$$Y = \frac{1}{\beta_0 + \beta_1 X} + \varepsilon$$

$$Y = \beta_0 e^{\beta_1 X} + \varepsilon$$

Některé nelineární funkce můžeme linearizovat a použít analýzu pro lineární modely.

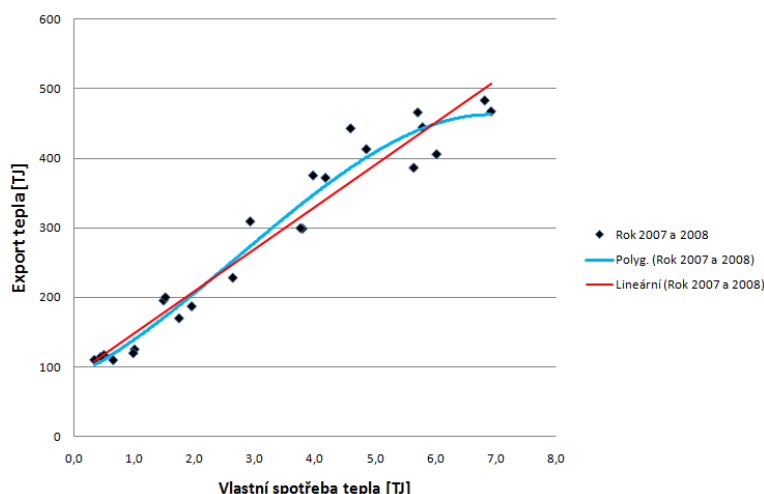
PŘÍKLAD:

Pro ukázkou si ukážeme jednu z hledaných závislostí a to závislost exportu tepla [TJ] na vlastní spotřebě tepla [TJ]. Z provozu jsme získali měsíční data z let 2007 a 2008, které jsou v Tabulce 1.

Tabulka 1: Data vlastní spotřeby tepla [TJ] a exportu tepla [TJ]

2007												
i	1	2	3	4	5	6	7	8	9	10	11	12
x	5,785	5,637	3,969	2,639	1,748	1,012	0,989	0,656	1,491	2,927	4,592	6,814
y	444,3	386,1	375	227,9	169,6	125	119,4	109,5	195,1	308,9	442,7	482,9
2008												
i	13	14	15	16	17	18	19	20	21	22	23	24
x	5,704	6,016	4,853	3,795	1,958	0,504	0,448	0,345	1,523	3,76	4,175	6,92
y	465,7	405,5	412,8	298,5	186,8	116,9	113,9	109,8	199,8	299,2	371,7	467,1

kde x je vlastní spotřeba tepla [TJ] a y je export tepla [TJ]. Nejprve si vyneseme naměřené hodnoty do dvourozměrného bodové grafu (Obrázek 6).



Obrázek 6: Graf závislosti exportu tepla na vlastní spotřebě tepla

Podle grafu můžeme soudit, že vhodnou regresní funkcí je polynom třetího stupně nebo přímka. Pro názornost zkusíme obě varianty a porovnáme je.

- regresní funkce η_1 má tvar: $y = \beta_0 + \beta_1 x$
- regresní funkce η_2 má tvar: $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3$

5.2 Odhad parametrů β

Dalším krokem je odhad neznámých koeficientů β zvolené regresní funkce pomocí vhodného vyrovnávacího kritéria. Ukážeme si nejpoužívanější kritéria.

Máme n naměřených hodnot pro obě proměnné. Vyrovnanou hodnotou $y_i, i = 1, \dots, n$ rozumíme hodnotu \hat{y}_i , kterou získáme z odhadnuté regresní rovnice η dosazením hodnot x_i . Vyrovnané hodnoty \hat{y}_i jsou aproximací y_i . Rozdíl mezi naměřenou a vyrovnanou (predikovanou) hodnotou říkáme *reziduum* a značíme je e_i . Nemá-li regresní model větší nedostatky, je možné rezidua e_i brát za odhady náhodné chyby ε , kde $e_i = y_i - \hat{y}_i$. Při výběru vhodného vyrovnávacího kritéria požadujeme viz [6], aby součet reziduí byl nulový

$$\sum_{i=1}^n e_i = 0.$$

To ale nestačí k nalezení vhodného tvaru regresní funkce, proto navíc požadujeme vhodné minimalizační kritérium.

• Metoda nejmenších čtverců

V této nejpoužívanější metodě minimalizujeme součet čtverců reziduí

$$\min Q(e) = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2.$$

Odhad parametrů β je potom získán pomocí tzv. *normálních rovnic*

$$\mathbf{b} = (\mathbf{F}^T \mathbf{F})^{-1} \mathbf{F}^T \mathbf{y},$$

kde matice $\mathbf{F} = \begin{pmatrix} 1 & f_1(x_1) & \cdots & f_R(x_1) \\ \vdots & \vdots & \cdots & \vdots \\ 1 & f_1(x_n) & \cdots & f_R(x_n) \end{pmatrix}$ a vektor $\mathbf{y} = (y_1, \dots, y_n)$. (V některých literaturách se používá místo matice \mathbf{F} matice \mathbf{X} , kde $\mathbf{X} = \mathbf{F}^T$.)

Tato metoda nám dává nejlepší nestranný odhad vektoru neznámých parametrů β .

• Metoda maximální věrohodnosti

Pokud známe rozdělení náhodné chyby ε , můžeme použít metodu maximální věrohodnosti a tím získat vhodné kritérium. Tato metoda používá k nalezení maximálně věrohodného odhadu tzv. věrohodnostní funkci $L(e_1, e_2, \dots, e_n, \theta)$, kde e_1, e_2, \dots, e_n jsou pozorování náhodné veličiny ε a θ je parametr rozdělení. Nejlepší odhad pak získáme pro takovou hodnotu θ' , pro kterou dosahuje věrohodnostní funkce svého maxima.

Př.: Má-li ε dvojitě exponenciální rozdělení, metoda maximální věrohodnosti vede k minimalizaci součtu absolutních odchylek

$$\min L_1 = \sum_{i=1}^n |e_i|.$$

• Méně používaná jsou kritéria **minimalizace největší hodnoty rezidua** $\min \max(e_i)$ a **minimalizace součtu relativních chyb** $\min \sum_{i=1}^n \frac{e_i}{y_i}$.

PŘÍKLAD:

Pokračujeme v příkladu z kapitoly 5.1. Vybrali jsme si nejběžnější kritérium pro vyrovnání hodnot y metodu nejmenších čtverců, odhad parametrů β tedy získáme vypočítáním normálních rovnic $\mathbf{b} = (\mathbf{F}^T \mathbf{F})^{-1} \mathbf{F}^T \mathbf{y}$.

Výpočty provádíme v programu MATLAB.

a) Za předpokladu lineárního tvaru regresní funkce $y = \beta_0 + \beta_1 x$ získáme z normálních rovnic odhady parametrů β :

$$b_0 = 86,8261 \quad \text{a} \quad b_1 = 60,6986.$$

Rovnice η_1 je tvaru $y = 86,8261 + 60,6986x$.

b) Za předpokladu kubického tvaru regresní funkce $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3$ získáme z normálních rovnic odhady parametrů β :

$$\begin{aligned} b_0 &= 86,2888, & b_1 &= 42,5799, \\ b_2 &= 11,3547, & b_3 &= -1,3939. \end{aligned}$$

Rovnice η_2 je tvaru $y = 86,2888 + 42,5799x + 11,3547x^2 - 1,3939x^3$.

5.3 Analýza reziduí

Ověření vhodnosti modelu začíná analýzou reziduí. Protože chceme, abychom rezidua mohli považovat za odhady rušivé složky ε , musí tedy splňovat předpoklady kladené na ε , tj., že náhodná chyba ε systematicky nezkrsluje vysvětlovanou proměnnou Y .

To znamená, že pro dané hodnoty proměnné X jsou hodnoty nepozorovatelné rušivé složky ε normálně rozdělené, nezávislé náhodné veličiny s nulovou střední hodnotou a rozptylem σ^2 . Jestliže máme více pozorování y_i pro stejná x_i , požadujeme po náhodných

chybách ε , aby měly n -rozměrné normální rozdělení s nulovými středními hodnotami a konstantním rozptylem. Tedy $E(\varepsilon) = \mathbf{0}_n$ a kovarianční matice $C(\varepsilon) = \sigma^2 \mathbf{I}_n$.

Po zjištění odhadu β máme už konkrétně danou regresní funkci, tedy i vyrovnané hodnoty \hat{y}_i a můžeme si spočítat hodnoty reziduí e_i , $i = 1, \dots, n$. Bodový odhad rozptylu σ^2 je výraz

$$s^2 = \frac{Q(e)}{n - (R + 1)},$$

kde $Q(e) = \sum_{i=1}^n (y_i - \hat{y}_i)^2$ je reziduální součet čtverců.

Postupně si ověříme všechny předpoklady kladené na rezidua: normalitu, konstantní rozptyl (homoskedasticita) a neautokorelovanost.

5.3.1 Normalita reziduí

Normalitu reziduí můžeme testovat pomocí testů dobré shody nebo pouze rozborem jejich histogramu a bodového grafu. Jakákoliv nenáhodnost vektoru \mathbf{e} ukazuje na nedostatek regresního modelu (např. špatný typ regresní funkce, chybná pozorování, ...). Konkrétně si ukážeme jeden z testů normality rozdělení viz [14].

• Modifikovaný Kolmogorovův-Smirnovův test normality

V případech malých výběrů, kde parametry normálního rozdělení nejsou známy (neúplně specifikovaný model), používáme k porovnání shody empirického a teoretického rozdělení tento test. Střední hodnotu a rozptyl odhadneme z výběrových hodnot a testujeme hypotézu H_0 , že výběr pochází ze souboru s normálním rozdělením. Předpokládaný konstantní rozptyl σ^2 odhadneme výrazem $s^2 = \frac{Q(e)}{n - (R + 1)}$ a střední hodnotu známe $\mu = 0$.

Máme náhodný výběr rozsahu n , který uspořádáme podle velikosti. Máme tedy nový seřazený soubor $e_{(1)}, e_{(2)}, \dots, e_{(n)}$, kde $e_{(1)} \leq e_{(2)} \leq \dots \leq e_{(n)}$.

Potom testovací statistika je tvaru

$$D = \max_{e_{(i)}} \left\{ \left| F(e_{(i)}) - \frac{i-1}{n} \right|, \left| F(e_{(i)}) - \frac{i}{n} \right| \right\},$$

kde $F(e)$ je distribuční funkce normálního rozdělení

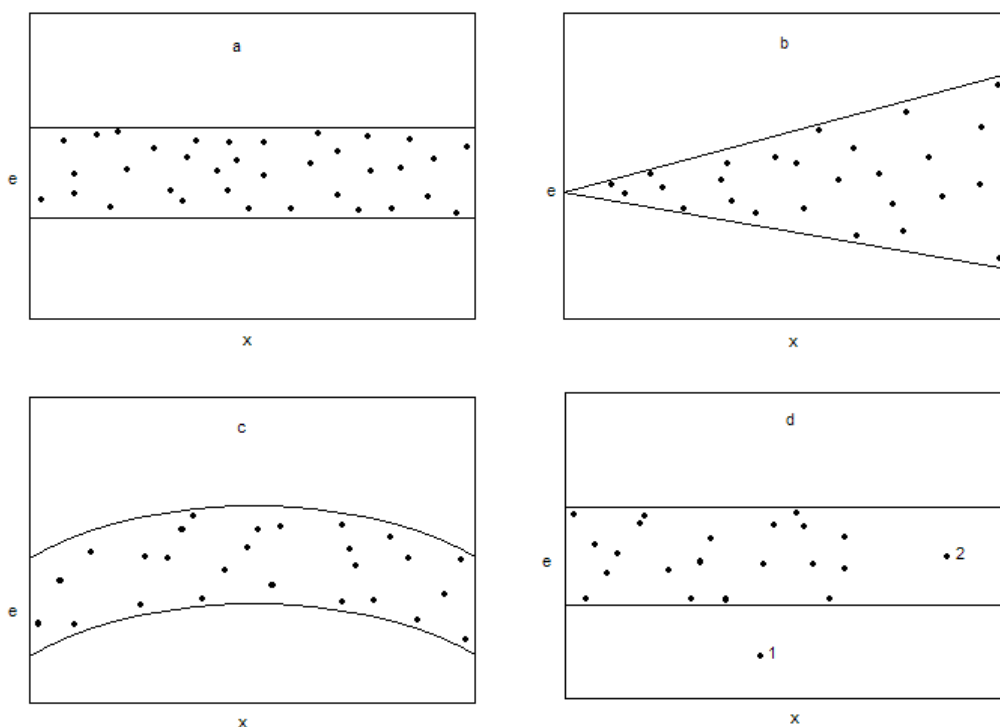
$$F(e) = \int_{-\infty}^e \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt.$$

Hypotézu H_0 zamítáme, jestliže $D > D_\alpha$, kde D_α je kritická hodnota maximální odchylky empirické distribuční funkce od teoretické. Tyto hodnoty jsou tabelovány díky Lilieforsovi (viz příloha tabulka T1). Tomuto modifikovanému K-S testu se také někdy říká Lilieforsův test normality.

• Rozbor grafů

Sestrojíme bodový graf $[x_i, e_i]$ a také histogram reziduí. Na Obrázku 7 máme příklady grafů různých rozložení bodů $[x_i, e_i]$.

- v tomto případě můžeme říct, že zvolený model dobře vystihuje závislosti proměnných, tedy rušivá složka systematicky nezkrsluje vysvětlovanou proměnnou
- rozptyl bodů se zvyšuje s rostoucím X , model nevyhovuje
- graf ukazuje nevhodnost stávajícího modelu, ale napovídá nám, jak by mohla vypadat



Obrázek 7: Rozbor grafů reziduí [7]

vhodná regresní funkce (např. pokud je stávající model lineární, vhodnější by byl kvadratický)

- d) tento graf obsahuje dva odlehlé body 1 a 2, tyto body mohou zkreslit výsledky regresní analýzy, proto se snažíme rozhodnout, zda jde o chybná měření nebo o důležité body, chybná měření zanedbáme a analýzu provedeme znovu.

PŘÍKLAD:

Pro naše dvě regresní funkce η_1 a η_2 vypočítáme hodnoty reziduí a provedeme jak modifikovaný K-S test, tak i rozbor grafů. Začneme modifikovaným K-S testem, takže si napočítáme hodnoty distribučních funkcí reziduí.

Pro test regresní funkce η_1 máme data:

Střední hodnota: $\mu = 0$

Směrodatná odchylka: $s_1 = 31,8443$

Hladina významnosti: $\alpha = 0,05$

Kritická hodnota max odchylky empirické distribuční funkce od teoretické: $D_\alpha = 0,1764$

Hodnota testovací statistiky: $D_1 = \max_{e_{(i)}} \left\{ \left| F(e_{(i)}) - \frac{i-1}{n} \right|, \left| F(e_{(i)}) - \frac{i}{n} \right| \right\} = 0,1907$

Pro test regresní funkce η_2 máme data:

Střední hodnota: $\mu = 0$

Směrodatná odchylka: $s_2 = 27,1894$

Hladina významnosti: $\alpha = 0,05$

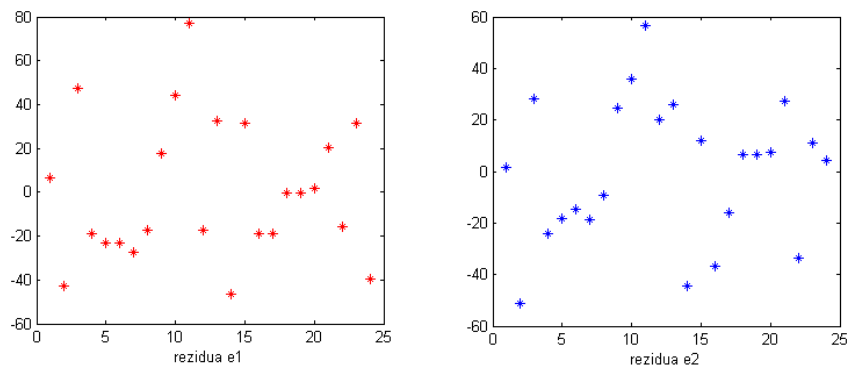
Kritická hodnota max odchylky empirické distribuční funkce od teoretické: $D_\alpha = 0,1764$

Hodnota testovací statistiky: $D_2 = \max_{e_{(i)}} \left\{ \left| F(e_{(i)}) - \frac{i-1}{n} \right|, \left| F(e_{(i)}) - \frac{i}{n} \right| \right\} = 0,1059$

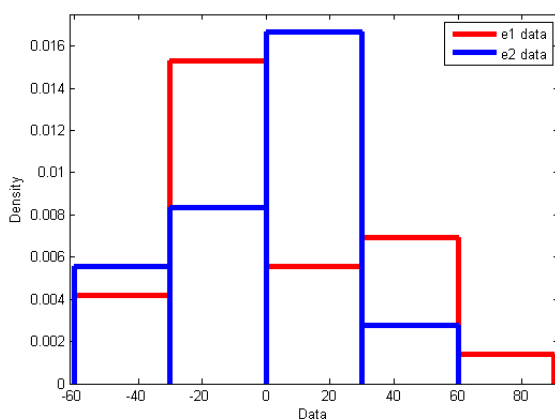
Jak vidíme z hodnot testovacích statistik $D_1 > D_\alpha > D_2$, hypotézu nezamítáme, že

rezidua získaná z regresní funkce η_2 mají normální rozdělení, kdežto rezidua získaná z η_1 normalitu nevykazují. Proto se nadále budeme zabývat pouze druhou regresní funkcí $y = 86,2888 + 42,5799x + 11,3547x^2 - 1,3939x^3$.

Při rozboru grafů reziduí si vykreslíme histogramy obou reziduí a jejich velikosti.



Obrázek 8: e_1 : velikost reziduí lineární fce, e_2 : velikost reziduí kubické funkce



Obrázek 9: e_1 - histogram reziduí lineární fce, e_2 - histogram reziduí kubické funkce

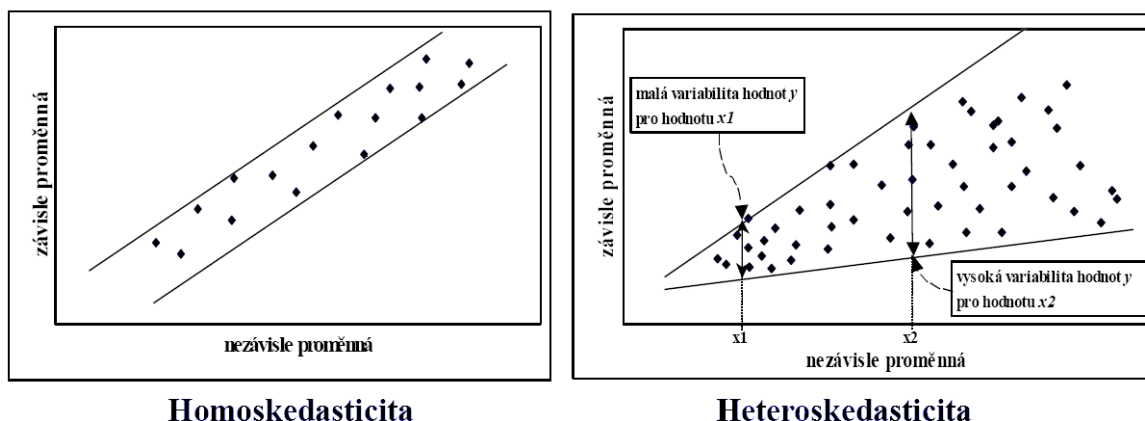
Z rozložení reziduí nemůžeme nijak rozhodnout, jejich histogram nám ale napovídá, že rezidua z η_2 mají normálnější charakter, což nám potvrdil i modifikovaný K-S test.

5.3.2 Homoskedasticita a heteroskedasticita

V klasickém lineárním modelu požadujeme tzv. homoskedasticitu veličin ε_i . *Homoskedasticita* znamená, že hodnoty závisle proměnné Y mají pro všechny hodnoty nezávisle proměnných \mathbf{X} konstantní rozptyl (variabilitu). Naopak *heteroskedasticita* vykazuje nekonstantní rozptyly. Rozdíl mezi homoskedasticitou a heteroskedasticitou je znázorněn na Obrázku 10.

• Ověření

K ověření můžeme použít např. *Bartlettův test*. Data rozdělíme do p skupin (podle hodnot proměnných X) a testujeme rovnost rozptylů ve skupinách. Jestliže počet prvků v jednotlivých skupinách je větší než 6, pak výsledná statistika má přibližně Pearsonovo rozdělení



Obrázek 10: Homoskedasticita \times heteroskedasticita rozptylů [16]

s $p - 1$ stupni volnosti.

Testujeme hypotézu homoskedasticity rozptylů $H : \sigma_1 = \dots = \sigma_p$. Testovacím kritériem Bartlettova testu je veličina B , která je definována

$$B = \frac{1}{C} \left[(n - p) \ln s^2 - \sum_{j=1}^p (n_j - 1) \ln s_j^2 \right],$$

kde p je počet podvýběrů a n_j , $j = 1, \dots, p$ je počet prvků v jednotlivých podvýběrech. Konstanta C je definovaná jako $C = 1 + (\sum_{j=1}^p \frac{1}{n_j - 1} - \frac{1}{n - p}) / 3(p - 1)$. Testovanou hypotézu zamítáme pokud platí $B \geq \chi_{1-\alpha}^2(p - 1)$, hodnoty kvantilů (viz příloha tabulka T3).

• Řešení

Heteroskedasticitu nejčastěji řešíme použitím metody vážených nejmenších čtverců, kdy se součet čtverců reziduí násobí vhodně zvolenými váhami w_i .

$$Q(e) = \sum_{i=1}^n [w_i(y_i - \hat{y}_i)]^2$$

Často používáme jako váhy hodnoty $1/y_i$ nebo $1/y_i^2$.

PŘÍKLAD:

Testováním normality jsme vyřadili lineární regresní funkci, budeme se tedy již zabývat jen kubickou funkcí, jejíž homoskedasticitu reziduí nyní otestujeme pomocí Bartlettova testu. Rezidua jsme rozdělili na čtyři podvýběry. Podvýběry jsme určili rozdělením všech reziduí podle pořadí, kde každý podvýběr má 6 členů.

Data pro výpočet výše uvedené statistiky B (výpočty provedené pomocí MATLABu):

Počet podvýběrů: $p = 4$

Hladina významnosti: $\alpha = 0,05$

α -kvantil Pearsonova rozdělení s $(p - 1)$ stupni volnosti: $\chi_{1-\alpha}^2(3) = 7,815$

Velikosti a rozptyly podvýběrů:

velikosti: $n_1 = n_2 = n_3 = n_4 = 6$

rozptyl prvního podvýběru: $s_1^2 = 911,6984$

rozptyl druhého podvýběru: $s_2^2 = 1182,2818$

rozptyl třetího podvýběru: $s_3^2 = 883,9444$

rozptyl čtvrtého podvýběru: $s_4^2 = 422,6882$

Potom statistika $B = [(n-p) \ln s^2 - \sum_{j=1}^p (n_j - 1) \ln s_j^2] / C = 0,0753$ a tedy $B < \chi_{1-\alpha}^2(3)$, což znamená, že potvrzujeme hypotézu o rovnosti rozptylů (homoskedasticitu). Tento výsledek potvrzuje i Obrázek 8.

5.3.3 Autokorelace reziduí

Dalším z požadavků na chybovou složku ε je, aby to byla nekorelovaná náhodná veličina. *Autokorelací* rozumíme závislost hodnot náhodné složky ε , která zapříčiňuje, při nenáhodném vektoru X , i závislost hodnot vysvětlované proměnné Y .

Nyní tedy rezidua e_i bereme jako konkrétní hodnoty náhodné složky. K testování autokorelace se používá Durbin-Watsonova statistika ve tvaru

$$W = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2}.$$

Hodnoty této statistiky náleží intervalu $\langle 0, 4 \rangle$. Pokud je W rovno číslu 2, rezidua nevykazují žádnou autokorelaci, hodnoty $W < 2$ značí pozitivní autokorelaci a hodnoty $W > 2$ značí negativní autokorelaci. Zjednodušeně můžeme postupovat takto: leží-li hodnota testové statistiky v intervalu $\langle 1, 4; 2, 6 \rangle$, rezidua nevykazují autokorelaci, $W < 1,4$ značí kladnou autokorelaci, $W > 2,6$ značí zápornou autokorelaci. Uvedené pravidlo je třeba chápat jen jako orientační. Tento test je přesně popsán v [5] str. 208 - 211.

PŘÍKLAD:

Pro zjištění neautokorelovanosti reziduí využijeme výše uvedenou Durbin-Watsonovu statistiku. Z výpočtů provedených v programu MATLAB je hodnota testovací statistiky $W = \sum_{i=2}^n (e_i - e_{i-1})^2 / \sum_{i=1}^n e_i^2 = 1,9132$. Tedy $W \in \langle 1, 4; 2, 6 \rangle$ a my zamítáme hypotézu o autokorelovanosti reziduí.

Tímto jsme dokončili celkovou analýzu reziduí, která nám potvrdila předpoklady na rezidua, což je jejich normální rozdělení s konstantním rozptylem a jejich neautokorelovanost.

5.4 Testy hypotéz o regresních koeficientech

Při splnění předpokladů o normalitě, homoskedasticitě a neautokorelovanosti reziduí, předpokládáme vhodnost zvoleného modelu. Pak můžeme testovat:

Celkový F-test všech regresních koeficientů

Zde se testuje hypotéza o nulovosti všech regresních koeficientů proti hypotéze, že alespoň jeden koeficient je nenulový. Při potvrzení hypotézy $H : \beta_1 = \beta_2 = \dots = \beta_K = 0$ tedy dostáváme, že regresní model můžeme zredukovat na model konstantní $Y = \beta_0 + \varepsilon$. Testovací kritérium je tvaru

$$F = \frac{Q(\hat{y})/K}{Q(e)/(n - (K + 1))},$$

kde $Q(\hat{y}) = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2$ a \bar{y} myslíme průměr hodnot y . Absolutní hodnota $|F|$, která je větší než kvantil $F_{1-\alpha/2}$ Fisher-Snedecorova rozdělení s K a $(n - (K + 1))$ stupni volnosti dovoluje zamítnout hypotézu H (viz příloha tabulka T4).

Zamítnutí této hypotézy neznamená, že všechny proměnné jsou „užitečné“, ale pouze ukazuje, že aspoň jedna proměnná má zásadní vliv. Jestliže hypotézu H potvrdíme, zkusíme zanalyzovat konstantní regresní funkci $Y = \beta_0 + \varepsilon$ nebo hledáme novou regresní funkci.

Testy hypotéz o jednotlivých regresních koeficientech

Jak už bylo řečeno, dobrá regresní funkce dostatečně vysvětluje závislost Y na X , ale zároveň požadujeme, aby měla co nejméně regresních koeficientů. Proto testujeme hypotézu o nulovosti koeficientů β_i , $i = 0, \dots, K$, které by mohly být z modelu vyloučeny pro svůj malý vliv na Y .

Pro testování těchto hypotéz se používají individuální t -testy, kdy testujeme hypotézu $H_i : \beta_i = 0$ proti hypotéze $\bar{H}_i : \beta_i \neq 0$. Testové kritérium je tvaru:

$$t_i = \frac{b_i}{s(b_i)},$$

kde $s(b_i) = s\sqrt{a_{jj}}$ a a_{jj} je j -tý diagonální prvek matice $\mathbf{A} = (\mathbf{X}^T \mathbf{X})^{-1}$. Absolutní hodnota $|t_i|$, která je větší než kvantil Studentova rozdělení $t_{1-\alpha/2}$ s $(n - (K + 1))$ stupni volnosti (viz příloha tabulka T2), dovoluje zamítnout hypotézu $H_i : \beta_i = 0$.

Při zjištění nevýznamnosti některého z koeficientů regresní funkce ho vypustíme a analýzu provedeme znovu.

PŘÍKLAD:

Nejprve uděláme celkový F-test. Pro testovací kritérium potřebujeme znát tyto hodnoty: průměr y : $\bar{y} = 284,7542$

součet čtverců vyrovnaných hodnot \hat{y} : $Q(\hat{y}) = 409480$

$n = 24$ a $K = 3$

hladina významnosti $\alpha = 0,05$

kritická hodnota $F_{1-\alpha/2}(3, 20) = 3,859$

Hodnota testovací statistiky $F = 160,5507 > 3,859$, což znamená, že zamítáme hypotézu o nulovosti všech regresních koeficientů.

Nyní provedeme jednotlivé t -testy.

hodnota $s(b_0) = 582,7025$, hodnota $s(b_1) = 1004,9$,

hodnota $s(b_2) = 105,6245$, hodnota $s(b_3) = 0,8788$,

hladina významnosti $\alpha = 0,05$

kritická hodnota $t_{1-\alpha/2}(20) = 2,086$

Testovací statistiky nabývají hodnot:

$t_0 = 3,5746 > 2,086$ potvrzuje významnost absolutního koeficientu b_0

$t_1 = 1,3432 < 2,086$ ukazuje na malou významnost koeficientu b_1

$t_2 = 1,1048 < 2,086$ ukazuje na malou významnost koeficientu b_2

$t_3 = 1,4869 < 2,086$ ukazuje na malou významnost koeficientu b_3

Nejprve tedy vyloučíme koeficient, pro jehož hodnotu vykazuje tato statistika nejmenší hodnotu. Znovu testujeme model s kubickou regresní funkcí, ve které ale chybí kvadratický člen. Nová regresní funkce je tvaru $y = \beta_0 + \beta_1 x + \beta_2 x^3$.

Celý postup jsme zopakovali pro novou regresní funkci, všechny testy nám opět vyšli,

tedy jsme potvrdili všechny požadavky na rezidua a celkový F test nám znovu potvrdil nenulovost všech koeficientů. Znovu uvedeme pouze hodnoty jednotlivých t-testů pro novou funkci

$$t_0 = 4,5731 > 2,080 \text{ potvrzuje významnost absolutního koeficientu } b_0$$

$$t_1 = 10,8876 > 2,080 \text{ potvrzuje významnost lineárního koeficientu } b_1$$

$$t_2 = 2,4797 > 2,080 \text{ potvrzuje významnost kubického koeficientu } b_2$$

Nyní už jsme potvrdily významnost všech koeficientů a jelikož jsme potvrdili i zbývající předpoklady, budeme dále pracovat s novou funkcí

$$y = 64,6049 + 76,7352x - 0,3713x^3$$

5.5 Index determinace

Vhodnost modelu ověřujeme nejen analýzou reziduí a testy, ale i kritériem: indexem determinace. Je zřejmé, že čím je model lepší, tím menších hodnot bude nabývat součet reziduálních čtverců modelu. Naopak špatný model znamená velkou hodnotu reziduálního součtu čtverců. *Index determinace* I^2 , který nabývá hodnot $\langle 0, 1 \rangle$ vyjadřuje poměr součtu čtverců vyhlazených hodnot a skutečně naměřených.

$$I^2 = \frac{Q(\hat{y})}{Q(y)} = 1 - \frac{Q(e)}{Q(y)}$$

Číslo $100 I^2 [\%]$ vyjadřuje variabilitu y_i vysvětlenou danou regresní funkcí. Čím více se koeficient determinace blíží jedné, tím více naznačuje vhodnost zvoleného tvaru regresní funkce. Druhá odmocnina koeficientu determinace se nazývá *index korelace* I .

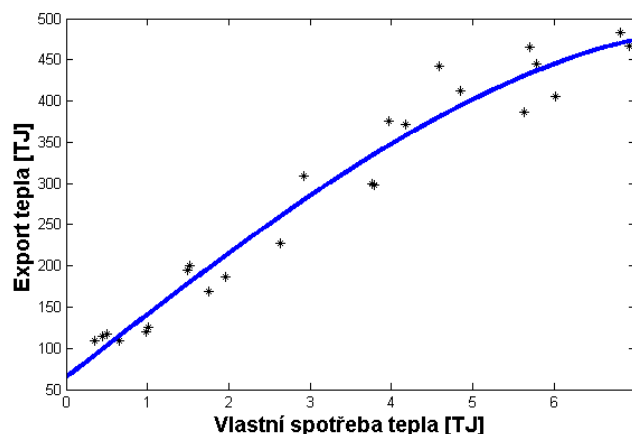
PŘÍKLAD:

Jako poslední si u tohoto příkladu ověříme velikost indexu determinace I^2 .

$$Q(e) = 18041 \text{ a } Q(y) = 426480$$

Potom $I^2 = 1 - Q(e)/Q(y) = 0,9577$. Tato vysoká hodnota indexu determinace, spolu s analýzou reziduí a dalšími testy, nám ukazuje vhodnost námi zvoleného regresního modelu

$$y = 64,6049 + 76,7352x - 0,3713x^3.$$



6 Vícerozměrná regrese

Vícerozměrná regrese znamená, že máme více než jednu proměnnou X , označíme je X_1, \dots, X_K . Narozdíl od jednorozměrné regrese po nich musíme požadovat jisté vlastnosti, které nám zaručí kvalitu vypočítaných testů a odhadů. První otázka, kterou si položíme, je: Které proměnné X_i použijeme?

6.1 Výběr proměnných

V regresních úlohách máme často více vysvětlujících proměnných, které mohou mít vliv na chování Y . Při použití všech by byl model příliš složitý, proto je naším cílem nalézt jejich nejvhodnější podmnožinu tak, aby dostatečně přesně vysvětlovala hodnoty Y . Při jejich výběru se ohlížíme, buď na známé souvislosti dané problematiky nebo na výhradně statistické ukazatele.

Algoritmus postupné regrese vybírá podmnožinu vysvětlujících proměnných následujícím postupem viz [7]:

Jako první proměnnou vybereme takové X_i , které má největší hodnotu korelačního koeficientu s Y . Poté přibereme takovou proměnnou X_j , která má největší parciální korelační koeficient s Y .

Do modelu přidáme jen takové proměnné, které dostatečně zlepši kvalitu modelu, tedy její příspěvek neklesne pod určitou úroveň. Přibírání proměnných ukončíme, jakmile žádná další proměnná významně nezlepšuje kvalitu modelu. Bohužel tento postup nenajde vhodnou regresní funkci, ale spíše vybere nejdůležitější proměnné.

Nyní máme vybrané vhodné proměnné X_i . Dalším krokem je výběr regresní funkce.

PŘÍKLAD:

Pro názorný příklad jsme si vybrali hledání závislosti účinnosti kotle na množství dávkovaného uhlí a biomasy, jejich výhřevností a na průměrném měsíčním zatížení kotle. Z provozu jsme získali měsíční data z let 2007 a 2008, naměřená data jsou v Tabulce 2.

Označení:

- X_1 ... Výhřevnost biomasy [GJ/t] (v modelu teplárny označeno α_b),
- X_2 ... Výhřevnost uhlí [GJ/t] (v modelu teplárny označeno α_u),
- X_3 ... Množství uhlí [t] (v modelu teplárny označeno \mathbf{x}_{j_u}),
- X_4 ... Množství biomasy [t] (v modelu teplárny označeno \mathbf{x}_{j_b}),
- X_5 ... Průměrné měsíční zatížení kotle [MW]
- Y ... Účinnost kotle [%] (v modelu teplárny označeno η).

Tabulka 2: Naměřená data ovlivňující účinnost kotle 3 v roce 2007 a 2008

i	X_1	X_2	X_3	X_4	X_5	Y
1	9,060422	13,452985	21702,85742	4270,198675	394,451569	87,44
2	8,6	13,755375	15673,36428	4580	388,943493	89,08
3	9,100451	13,840148	17206,50045	6431,318681	384,986587	87,08
4	10,399962	14,147268	16759,87146	5232,980769	358,321023	86,53
5	12,370974	14,458045	13224,27742	2245,27314	372,636735	83,39
6	13,685714	13,880605	547,151951	224,158572	187,972973	65,21

7	12,099978	13,779256	14740,87557	4620,991736	316,943213	88,34
8	12,3	13,732418	14664,77365	4849,837398	309,584677	88,24
9	11,799934	13,939862	15983,50863	5456,864407	353,275	88,57
10	10,880688	13,552331	18222,20841	6302,813973	383,438430	89,81
11	10,513602	13,515013	19403,79516	6384,863207	421,678886	87,31
12	9,743335	12,790582	22494,31162	6013,097651	427,540541	90,88
13	9,684819	13,091558	18800,13007	6709,469881	395,566588	91,039198
14	10,781761	13,335799	15295,29595	6269,847836	399,200588	88,931734
15	11,039162	13,134938	17256,49621	6771,709449	397,390175	90,574536
16	12,081116	13,653787	2588,585824	638,020541	366,407169	91,065612
17	11,560327	15,017486	12185,12895	6043,686935	317,648515	86,931654
18	12,534909	14,878898	10210,70276	5291,701686	313,384143	92,038778
19	13,208125	13,607032	16809,61784	7215,1046	375,469260	86,211683
20	14,227420	13,968958	16666,02481	6898,088347	383,928497	84,338076
21	13,404025	14,683747	14591,59588	6806,067634	399,437450	86,951338
22	13,260079	13,120611	17703,82476	7414,812562	399,290442	88,769857
23	12,193873	13,381786	18030,80407	6990,940646	393,332725	86,729763
24	11,327779	13,695030	18311,59156	8287,166922	396,764721	85,649320

Nejprve si vybereme důležité proměnné, ze kterých budeme později sestavovat regresní funkci, pomocí korelačních koeficientů. V našem případě vynecháváme proměnou X_5 , jelikož měla nevýznamný vliv na Y .

6.2 Tvar regresní funkce

Hledání vhodné funkce je velmi náročné, stejně jako u jednorozměrné regrese. Navíc vícerozměrná regrese má tu nevýhodu, že si data nemůžeme vykreslit ve 2D grafu a podle vizuálního posouzení odhadnout regresní funkci.

U vícerozměrné regrese se při vyšším počtu vysvětlujících proměnných nejčastěji používá zcela lineární model tvaru $Y = \beta_0 + \beta_1 X_1 + \dots + \beta_K X_K$.

PŘÍKLAD:

Po zkoušení různých typů regresních funkcí, od zcela lineární přes kvadratické, exponenciální a jejich různé kombinace, jsme po konzultacích s odborníky a po testech vhodnosti funkce, dospěli k následujícímu tvaru

$$Y = \beta_0 + \beta_1 X_1^2 + \beta_2 X_2^2 + \beta_3 X_1 X_4 + \beta_4 X_2 X_3.$$

Zavedeme si nové pomocné proměnné $Z_1 = X_1^2$, $Z_2 = X_2^2$, $Z_3 = X_1 X_4$, $Z_4 = X_2 X_3$ a řešíme novou zcela lineární regresní funkci

$$Y = \beta_0 + \beta_1 Z_1 + \beta_2 Z_2 + \beta_3 Z_3 + \beta_4 Z_4.$$

6.3 Vlastnosti dat

Jak už jsme výše uvedli, u vícerozměrné regrese máme jisté požadavky na vlastnosti dat. Ukážeme si, jak tyto požadavky ověřit, případně vyřešit jejich nesplnění.

6.3.1 Nezávislost proměnných

Vysvětlující proměnné X_1, \dots, X_K jsou nenáhodné a neexistuje mezi nimi funkční lineární závislost. Tzn. matice \mathbf{X} má kromě prvního jednotkového sloupce K lineárně nezávislých sloupců s danými hodnotami proměnných, a že počet pozorování n je větší než $K + 1$ neznámých regresních koeficientů.

Ověříme platnost:

- $h(\mathbf{X}) = K + 1 \leq n$,
- $\det(\mathbf{X}^T \mathbf{X}) \neq 0$ a matice $(\mathbf{X}^T \mathbf{X})^{-1}$ existuje,

kde $h(\mathbf{X})$ značí hodnost matice \mathbf{X} , $\det(\mathbf{X})$ označuje determinant \mathbf{X} a \mathbf{X}^{-1} inverzní matici.

PŘÍKLAD:

Nejprve si ověříme nezávislost proměnných $Z_i, i = 1, \dots, 4$. Matice \mathbf{Z} je tvořena prvním sloupcem jedniček a další sloupce tvoří proměnné Z_i .

- $h(\mathbf{Z}) = 5 \leq 23$ ✓
- $\det(\mathbf{Z}^T \mathbf{Z}) = 5,44129510^{29} \neq 0$ a matice $(\mathbf{Z}^T \mathbf{Z})^{-1}$ existuje ✓

Tedy proměnné Z_i jsou nezávislé.

6.3.2 Multikolinearita

Dalším požadavkem na data, který musíme ověřit je multikolinearita. Vektory matice \mathbf{X} musí být skutečně navzájem nezávislé (jejich párové korelační koeficienty musí být nulové nebo statisticky nevýznamné). Pokud tomu tak není, dochází k tzv. multikolinearitě, která způsobuje početní i statistické problémy.

Perfektní multikolinearita znamená, že jeden ze sloupců matice \mathbf{X} je lineární kombinací zbývajících. Pojmeme *multikolinearita* ale rozumíme i případy, kdy některé sloupce matice \mathbf{X} svírají téměř nulový úhel a jsou tedy skoro lineárně závislé ($\sum_{i=1}^K \mathbf{c}_i X_i = \delta$, kde δ je skoro nulový vektor a vektor \mathbf{c} je nenulový vektor) viz [15].

Multikolinearitu může způsobit přeurčenost regresního modelu („zbytečně“ mnoho nezávisle proměnných) nebo skutečně existující závislost mezi „původně předpokládánými nezávislými“ proměnnými.

• Ověření multikolinearity v datech

Multikolinearita způsobuje špatnou podmíněnost matice $\mathbf{X}^T \mathbf{X}$, což znamená, že:

- 1) $\det(\mathbf{X}^T \mathbf{X})$ je blízký 0.
- 2) Alespoň jedno vlastní číslo (značíme λ) této matice je blízké 0.

Multikolinearitu můžeme rozpoznat i z korelační matice \mathbf{R} , prvky této matice vypočítáme

$$r_{ij} = \frac{C(X_i, X_j)}{\sqrt{D(X_i)D(X_j)}}, \quad i, j = 1, \dots, K.$$

- 3) $\det(\mathbf{R})$ je blízký 0.
- 4) Alespoň jedno vlastní číslo této matice je blízké 0.
- 5) Některé hodnoty r_{ij} jsou vysoké (blízké ± 1).
- 6) Index podmíněnosti κ matice \mathbf{R} je vysoký ($\kappa = \sqrt{\frac{\lambda_{max}}{\lambda_{min}}}$, kde λ_{max} je maximální a λ_{min} je minimální vlastní číslo matice \mathbf{R}). Je-li $\kappa > 1000$, jde o silnou multikolinearitu.

Dalším kritériem rozpoznání je tzv. VIF-faktor (variance inflation factor). $VIF_j = \frac{1}{\widetilde{r}_{jj}}$, kde \widetilde{r}_{jj} je j -tý diagonální prvek matice \mathbf{R}^{-1} . Pokud je $VIF_j > 10$, jde o silnou multikolinearitu.

• Důsledky multikolinearity

- 1) Nestabilita odhadů β způsobená velkou citlivostí na malé změny v datech.
- 2) Vysoký koeficient determinace, i když máme nesprávný regresní model.
- 3) F-test je významný, kdežto t-testy ukazují na nevýznamnost regresních koeficientů.
- 4) Nelze odděleně sledovat vliv jednotlivých vysvětlujících proměnných na Y .

• Řešení

K odstranění nebo zmenšení vlivu multikolinearity může vést: snížení počtu nezávisle proměnných (vypouštíme takové, kterým odpovídají nízká vlastní čísla nebo vysoké hodnoty VIF-faktoru), použití jiného modelu nebo použití jiné metody výpočtu (obvykle metoda hřebenové regrese nebo metoda hlavních komponent).

PŘÍKLAD:

Data jsou nezávislá a nyní ověříme jestli nejsou multikolineární

- 1) $\det(\mathbf{Z}^T \mathbf{Z}) = 5,44129510^{29}$ není blízky 0 ✓
- 2) nejmenší vlastní číslo matice $\mathbf{Z}^T \mathbf{Z} : \lambda_{\mathbf{Z}^T \mathbf{Z}, \min} = 0,083830$ ✓
(vysoká multikolinearita se bere až od 10^{-3})
- 3) $\det(\mathbf{R}) = 0,249270$ není blízky 0 ✓
- 4) nejmenší vlastní číslo matice $\mathbf{R} : \lambda_{\mathbf{R}, \min} = 0,155923$ ✓ (jako 2)
- 5) nejvyšší hodnota prvku matice $\mathbf{R} : r_{ij, \max} = 0,472824$ ✓
- 6) index podmíněnosti matice $\mathbf{R} : \kappa_{\mathbf{R}} = 0,301005$ je malý ✓
- 7) největší VIF faktor $VIF_{j, \max} = 2,822855 < 10$ ✓

Můžeme tedy říci, že tato data nejsou multikolineární a pokračovat v další analýze, nyní si vypočítáme vyrovnané hodnoty a provedeme analýzu reziduí.

6.4 Analýza reziduí, testy koeficientů, index determinace

Po odhadnutí regresní funkce a tedy získání vyrovnaných hodnot postupujeme v analýze reziduí a dalších testech, které ověřují správnost modelu, stejně jako v jednorozměrné regresi:

Analýza reziduí

- test normality reziduí: modifikovaný Kolmogorův-Smirnovův test
- test homoskedasticity reziduí: Bartlettův test
- test neautokorelovanosti reziduí: Durbin-Watsonův test

Testy regresních koeficientů a vhodnosti funkce

- test nenulovosti všech regresních koeficientů: celkový F-test
- test nenulovosti jednotlivých regresních koeficientů: t-test
- určení míry vysvětlené variability: index determinace

Podle výsledků těchto testů rozhodneme o vhodnosti zvolené regresní funkce, jestliže testy prokážou nevhodnost, hledáme novou vhodnější regresní funkci.

PŘÍKLAD:

Nejprve jsme odstranili chybné měření (řádek 6), které by nám ovlivnilo analýzu (princip zjištění chybných měření je podrobně uvedeno v [10]). Nyní vypočítáme pomocí normálních rovnic odhady hodnot regresních koeficientů

$$\mathbf{b} = (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{Y}.$$

Odhady koeficientů v regresní funkci

$$Y = 109,302506 - 43445,176848Z_1 - 54751,437927Z_2 + 0,042672Z_3 - 0,035956Z_4.$$

Nyní získáme vyrovnané hodnoty a tedy i velikosti reziduí. Pokračujeme analýzou reziduí.

Test normality:

Vypočítáme hodnoty reziduí (jejich histogram je na Obrázku 11) a provedeme modifikovaný K-S test. Potřebná data k provedení testu:

Střední hodnota: $\mu = 0$

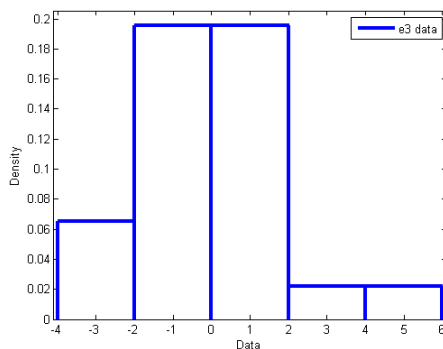
Směrodatná odchylka: $s = 1,6453$

Hladina významnosti: $\alpha = 0,05$

Kritická hodnota max odchylky empirické distribuční funkce od teoretické: $D_\alpha = 0,1798$

Hodnota testovací statistiky: $D = \max_{e_{(i)}} \left\{ \left| F(e_{(i)}) - \frac{i-1}{n} \right|, \left| F(e_{(i)}) - \frac{i}{n} \right| \right\} = 0,1465$

Jak vidíme z hodnot testovacích statistik $D_\alpha > D$ můžeme tedy potvrdit hypotézu, že rezidua získaná z regresní funkce mají normální rozdělení. ✓



Obrázek 11: Histogram reziduí

Test homoskedasticity:

Data jsme si rozdělili do čtyř skupin a provedeme Bartlettův test. Potřebná data k provedení testu:

Počet podvýběrů: $p = 4$

Hladina významnosti: $\alpha = 0,05$

α -kvantil Pearsonova rozdělení s $(p - 1)$ stupni volnosti: $\chi_{1-\alpha}^2(3) = 7,815$

Velikosti a rozptyly podvýběrů:

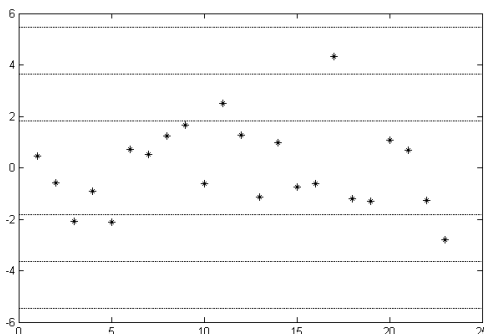
velikosti: $n_1 = n_2 = n_3 = 6, n_4 = 5$

rozptyl prvního podvýběru: $s_1^2 = 2,1420$

rozptyl druhého podvýběru: $s_2^2 = 2,5610$

rozptyl třetího podvýběru: $s_3^2 = 4,6553$

rozptyl čtvrtého podvýběru: $s_4^2 = 3,1911$
 Potom statistika $B = [(n-p) \ln s^2 - \sum_{j=1}^p (n_j - 1) \ln s_j^2] / C = 1,7382$ a tedy $B < \chi_{1-\alpha}^2(3)$,
 což znamená, že potvrzujeme hypotézu o rovnosti rozptylů (homoskedasticitu). ✓



Obrázek 12: Velikost reziduí doplněné o σ , 2σ a 3σ pásy

Test neautokorelovanosti:

Pro zjištění neautokorelovanosti reziduí využijeme Durbin-Watsonovův test. Hodnota testovací statistiky $W = \sum_{i=2}^n (e_i - e_{i-1})^2 / \sum_{i=1}^n e_i^2 = 1,8730$. Tedy $W \in \langle 1,4; 2,6 \rangle$ a my zamítáme hypotézu o autokorelovanosti reziduí. ✓

Tímto jsme dokončili celkovou analýzu reziduí, která nám potvrdila předpoklady na rezidua, což je jejich normální rozdělení s konstantním rozptylem a jejich neautokorelovanost. Nyní provedeme testy regresních koeficientů.

Celkový F-test

Nejprve si uděláme celkový F-test. Pro testovací kritérium potřebujeme znát tyto hodnoty:

průměr Y : $\bar{Y} = 88,0827$

součet čtverců vyrovnaných hodnot \hat{Y} : $Q(\hat{Y}) = 107,1504$

$n = 23$ a $K = 4$

hladina významnosti $\alpha = 0,05$

kritická hodnota $F_{1-\alpha/2}(4, 18) = 3,608$

Hodnota testovací statistiky $F = 8,0962 > 3,608$, což znamená, že zamítáme hypotézu o nulovosti všech regresních koeficientů. ✓

Jednotlivé t-testy

Data potřebná k provedení jednotlivých t-testů.

hodnota $s(b_0) = 6,2824$, hodnota $s(b_1) = 0,0183$

hodnota $s(b_2) = 0,0274$, hodnota $s(b_3) = 2,8474 \cdot 10^{-5}$

hodnota $s(b_4) = 1,2184 \cdot 10^{-5}$

hladina významnosti $\alpha = 0,05$

kritická hodnota $t_{1-\alpha/2}(18) = 2,101$

Testovací statistiky nabývají hodnot:

$t_0 = 17,3983 > 2,101$ potvrzuje významnost absolutního koeficientu b_0 ✓

$t_1 = 2,3687 > 2,101$ potvrzuje významnost koeficientu b_1 ✓

$t_2 = 2,0005 < 2,101$ ukazuje na malou významnost koeficientu b_2

$t_3 = 1,4986 < 2,101$ ukazuje na malou významnost koeficientu b_3

$t_4 = 2,9509 > 2,101$ potvrzuje významnost koeficientu b_4 ✓

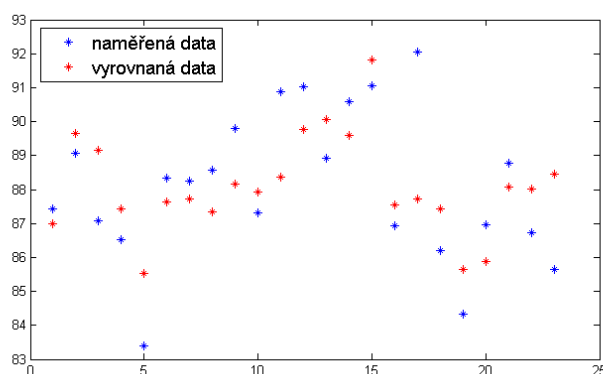
Nejprve tedy vyloučíme koeficient, pro jehož hodnotu vykazuje tato statistika nejmenší hodnotu. Znovu testujeme model s regresní funkcí v níž chybí proměnná Z_3 . Nová regresní funkce ale vykazuje podstatně horší výsledky a proto se vrátíme k předchozímu modelu, i když nám t-testy ukázali některé koeficienty jako málo významné.

Index determinace

Jako poslední si u tohoto příkladu ověříme velikost indexu determinace I^2 . Data potřebná k jeho výpočtu:

$Q(e) = 59,5558$ a $Q(y) = 107,1504$

Potom $I^2 = 1 - Q(e)/Q(y) = 0,4442$. Tato nepřilíš vysoká hodnota indexu determinace nám říká, že model není příliš vhodný. Kvůli malému počtu neopakovaných pozorování a vysokému počtu vysvětlujících proměnných se i tento výsledek dá považovat za dostačující a po konzultaci s odborníky prohlašujeme následující regresní funkci za vhodnou. Na následujícím Obrázku 13 jsou znázorněny naměřené a vyrovnané hodnoty účinnosti.



Obrázek 13: Vyrovnané \times naměřené hodnoty účinnosti kotle 3

$$Y = 109,302506 - 43445,176848Z_1 - 54751,437927Z_2 + 0,042672Z_3 - 0,035956Z_4$$

Po zpětném dosazení původních proměnných X_i dostáváme závěrečný tvar regresní funkce

$$Y = 109,302506 - 43445,176848X_1^2 - 54751,437927X_2^2 + 0,042672X_1X_4 - 0,035956X_2X_3.$$

7 Neurčitosti v modelu

Jak jsme již uvedli výše, cílem této diplomové práce je právě co nejpřesnější popsání neurčitostí v modelu. K tomuto jsme si zavedli potřebnou teorii a nyní popíšeme postup řešení a implementaci nejdůležitějších nejistot do modelu.

7.1 Účinnost kotlů

Z ekonomické stránky je účinnost kotlů důležitým faktorem, určuje totiž, jaké množství paliva se musí spálit, abychom uspokojili poptávku po teple, čili ovlivňuje provozní náklady. Z tohoto důvodu budeme věnovat faktorům ovlivňujícím účinnost zvýšenou pozornost.

K analýze účinnosti jsme získali podklady z provozu dané teplárny. Jako možné faktory, které výrazně ovlivňují účinnost, odborníci uvažují množství dávkovaných paliv (biomasa, uhlí), jejich výhřevnosti a faktor průměrného měsíčního zatížení kotle. Pozornost věnujeme kotlům 2 a 3, protože kotel 1 není po většinu roku využíván a k analýze není dostatek dat. Po dohodě s odborníky nastavíme účinnost kotle 1 na konstantní hodnotu 68%. Analýza kotle 3 je podrobně popsána v kapitole 6, kotel 2 je počítán analogicky.

Regresní analýza ukázala, že účinnost kotlů není příliš závislá na jeho průměrném měsíčním zatížení, naopak výhřevnosti paliv a součiny množství s výhřevnostmi se ukazují jako důležité. Funkce, která nám popisuje závislost při využití předchozích faktorů, vyšla následovně (kapitola 6):

kotel 2:

$$\eta_2 = 105,238638 - 22618,457024\alpha_B^2 - 80730,267772\alpha_U^2 + 0,043253\alpha_B x_B - 0,008139\alpha_U x_U,$$

kotel 3:

$$\eta_3 = 109,302506 - 43445,176848\alpha_B^2 - 54751,437927\alpha_U^2 + 0,042672\alpha_B x_B - 0,035956\alpha_U x_U.$$

Ale vzhledem k množství naměřených hodnot a ke složitosti úlohy nemusí tyto funkce, které jsme získali regresní analýzou, být reálné. Proto se vzniklou chybu snažíme eliminovat vygenerováním dalších funkcí účinnosti a tím i verifikovat získané \mathbf{x}_{opt} .

Při generování nových účinností aplikujeme WS přístup (kapitola 4.2). V regresní analýze předpokládáme, že daná funkce obsahuje chybovou složku ε , která má normální rozdělení s nulovou střední hodnotou a konstantním rozptylem $\sigma^2 = 3,3086$, což jsme ověřili (v kapitole 6.4). Nové funkce účinnosti jsme tedy získali novou regresní analýzou, kde jsme zachovali hodnoty nezávislých proměnných a hodnoty závisle proměnné (účinnosti) jsme nahradili novými hodnotami. Ty jsme získali vygenerováním náhodného výběru se stejným rozdělením jako má chybová složka a přičetli jsme je k vyrovnaným hodnotám účinnosti.

Nové rovnice účinnosti pro kotel 2:

$$\begin{aligned}\eta_{2_1} &= 106,07479 - 13549,28114\alpha_B^2 - 85719,57816\alpha_U^2 + 0,01511\alpha_B x_B - 0,01054\alpha_U x_U, \\ \eta_{2_2} &= 107,13806 - 16340,42535\alpha_B^2 - 92766,58793\alpha_U^2 + 0,02997\alpha_B x_B - 0,00926\alpha_U x_U, \\ \eta_{2_3} &= 116,55567 - 29270,41074\alpha_B^2 - 121454,31037\alpha_U^2 + 0,10239\alpha_B x_B - 0,01463\alpha_U x_U,\end{aligned}$$

$$\begin{aligned}
\eta_{24} &= 94,01270 - 29217,18855\alpha_B^2 - 27037,28284\alpha_U^2 + 0,08445\alpha_Bx_B - 0,00385\alpha_Ux_U, \\
\eta_{25} &= 99,85468 - 16731,25292\alpha_B^2 - 61481,48365\alpha_U^2 + 0,02558\alpha_Bx_B - 0,00555\alpha_Ux_U, \\
\eta_{26} &= 98,63801 - 25870,60399\alpha_B^2 - 54182,41973\alpha_U^2 + 0,03459\alpha_Bx_B - 0,00376\alpha_Ux_U, \\
\eta_{27} &= 104,58875 - 24731,56167\alpha_B^2 - 66833,65210\alpha_U^2 + 0,03334\alpha_Bx_B - 0,01141\alpha_Ux_U, \\
\eta_{28} &= 107,22921 - 25316,93052\alpha_B^2 - 88391,58787\alpha_U^2 + 0,02316\alpha_Bx_B - 0,00752\alpha_Ux_U, \\
\eta_{29} &= 99,41252 - 25130,37670\alpha_B^2 - 61500,68391\alpha_U^2 + 0,08044\alpha_Bx_B - 0,00317\alpha_Ux_U, \\
\eta_{210} &= 113,19845 - 20058,23888\alpha_B^2 - 107821,85437\alpha_U^2 + 0,01885\alpha_Bx_B - 0,01416\alpha_Ux_U.
\end{aligned}$$

Nové rovnice účinnosti pro kotel 3:

$$\begin{aligned}
\eta_{31} &= 101,74696 - 75916,71272\alpha_B^2 - 6923,46423\alpha_U^2 + 0,07598\alpha_Bx_B - 0,03135\alpha_Ux_U, \\
\eta_{32} &= 111,56863 - 43820,12751\alpha_B^2 - 60267,14022\alpha_U^2 + 0,04469\alpha_Bx_B - 0,04536\alpha_Ux_U, \\
\eta_{33} &= 108,97474 - 54599,95359\alpha_B^2 - 33798,74080\alpha_U^2 + 0,05667\alpha_Bx_B - 0,05097\alpha_Ux_U, \\
\eta_{34} &= 111,88236 - 35965,04995\alpha_B^2 - 62538,32229\alpha_U^2 + 0,05187\alpha_Bx_B - 0,04983\alpha_Ux_U, \\
\eta_{35} &= 107,43410 - 33186,57888\alpha_B^2 - 47575,83751\alpha_U^2 + 0,01889\alpha_Bx_B - 0,03443\alpha_Ux_U, \\
\eta_{36} &= 117,10730 - 47798,55237\alpha_B^2 - 71551,55817\alpha_U^2 + 0,06166\alpha_Bx_B - 0,06108\alpha_Ux_U, \\
\eta_{37} &= 121,70584 - 73041,51207\alpha_B^2 - 102905,51033\alpha_U^2 + 0,11129\alpha_Bx_B - 0,05288\alpha_Ux_U, \\
\eta_{38} &= 107,83244 - 43530,62465\alpha_B^2 - 48429,45917\alpha_U^2 + 0,06152\alpha_Bx_B - 0,03800\alpha_Ux_U, \\
\eta_{39} &= 117,46481 - 48743,43700\alpha_B^2 - 80540,12858\alpha_U^2 + 0,07373\alpha_Bx_B - 0,05363\alpha_Ux_U, \\
\eta_{310} &= 111,65517 - 40502,03347\alpha_B^2 - 69914,27673\alpha_U^2 + 0,05682\alpha_Bx_B - 0,03843\alpha_Ux_U.
\end{aligned}$$

Vygenerovali jsme si nové výše uvedené funkce účinnosti. Pro původní funkci účinnosti η vypočítanou regresní analýzou jsme v programu GAMS spočítali optimalizační úlohu, kde jsme získali hodnotu účelové funkce a především hodnotu vektoru určujícího množství dávkovaného paliva \mathbf{x}_{opt} , pro který získáme maximální hodnotu účelové funkce. Toto \mathbf{x}_{opt} jsme si zafixovali a vypočítali nové hodnoty účelové funkce tak, že jsme do optimalizační úlohy dosadili vygenerované funkce účinnosti. Hodnoty účelové funkce jsou uvedeny v Tabulce 3.

Tabulka 3: Hodnoty účelových funkcí pro jednotlivé funkce účinnosti kotlů

funkce účinnosti	hodnota účelové funkce	poměr účelové funkce
původní	787922247,16	1
1	755694212,15	0,959097
2	780779007,03	0,990934
3	783963961,48	0,994976
4	784074288,42	0,995116
5	757430335,96	0,961301
6	776670758,32	0,985720
7	780552844,23	0,990647
8	762987955,35	0,968354
9	783765481,17	0,994724
10	785403022,01	0,996803

Ve třetím sloupci je uveden poměr hodnot účelových funkcí, ve kterých jsou zapracovány nově vygenerované funkce účinnosti a hodnoty účelové funkce s použitím původní funkce účinnosti. Jak vidíme všechny tyto hodnoty přesahují 95%, můžeme tedy považovat původní funkci za dostatečně přesnou. Původní funkci účinnosti jsme verifikovali a dále tedy zahrneme do modelu pouze tuto funkci účinnosti a řešíme další neurčitosti modelu.

7.2 Velikost poptávky po teple

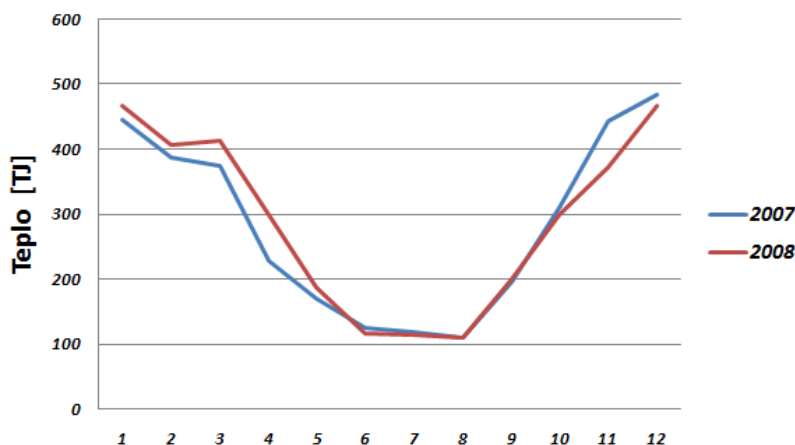
Provoz teplárny se odvíjí zejména od poptávky po teple. Pokud chce mít management teplárny představu o provozu na rok dopředu (např. z důvodu plánování odběru paliva) je třeba správně odhadnout výši poptávky. Velikost poptávky je udána v měsících. Tyto hodnoty nejsou dlouhodobě známy, ale vždy máme známou hodnotu poptávky pro aktuální měsíc, na kterou můžeme reagovat a podle velikosti poptávky dávkovat příslušné množství paliva.

Pro modelování této neurčitosti ξ_P využíváme stochastického programování. Jelikož známe aktuální velikost, na kterou reagujeme, je nejvhodnější WS přístup.

Pro generování scénářů velikosti poptávky předpokládáme znalost pravděpodobnostního rozdělení náhodné veličiny, tedy námi modelované neurčitosti. V našem případě předpokládáme v jednotlivých měsících rovnoměrné rozdělení, tedy $\xi_p^i \sim U(a_p, b_p)$, $p = 1, \dots, 12$ a $i = 1, \dots, 100$. Rovnoměrné rozdělení pravděpodobnosti přiřazuje všem hodnotám náhodné veličiny stejnou pravděpodobnost. Parametry a_p, b_p pro jednotlivé měsíce tohoto rozdělení odhadneme pomocí historických dat. K dispozici máme data z let 2007 a 2008 (Tabulka 4, Obrázek 14). Hodnoty poptávek ve stejných měsících zprůměrujeme a tuto hodnotu vezmeme jako střední hodnotu $E\xi_p = \frac{a_p + b_p}{2}$. Poté vezmeme maximální rozdíl mezi poptávkami v jednotlivých měsících a o něco větší hodnotu vezmeme jako délku všech intervalů (a_p, b_p) , na kterých generujeme odhady nových měsíčních hodnot poptávky.

Tabulka 4: Historické hodnoty poptávky po teple

	1	2	3	4	5	6
2007	444,288	386,064	374,971	227,855	169,590	125,027
2008	465,711	405,528	412,804	298,478	186,822	116,919
	7	8	9	10	11	12
2007	119,397	109,457	195,056	308,948	442,737	482,936
2008	113,923	109,778	199,763	299,236	371,711	467,121



Obrázek 14: Graf vývoje poptávky v roce 2007 a 2008

Maximální rozdíl mezi poptávkami je v měsíci listopadu 71,026, jako délku všech intervalů (a_p, b_p) jsme vzali hodnotu 80. Vygenerovali jsme si matici ξ_P typu $n \times q$, kde $q = 100$ a

$n = 12$, jejíž prvky jsou náhodné veličiny s rovnoměrným pravděpodobnostním rozdělením $\xi_p^i \sim U(a_p, b_p)$, $p = 1, \dots, 12$ a $i = 1, \dots, 100$. Máme tedy 100 hodnot ročních poptávek. Jejich realizaci považujeme za scénáře vývoje poptávky ξ_p^s , které implementujeme do stochastického optimalizačního modelu (kapitola 8).

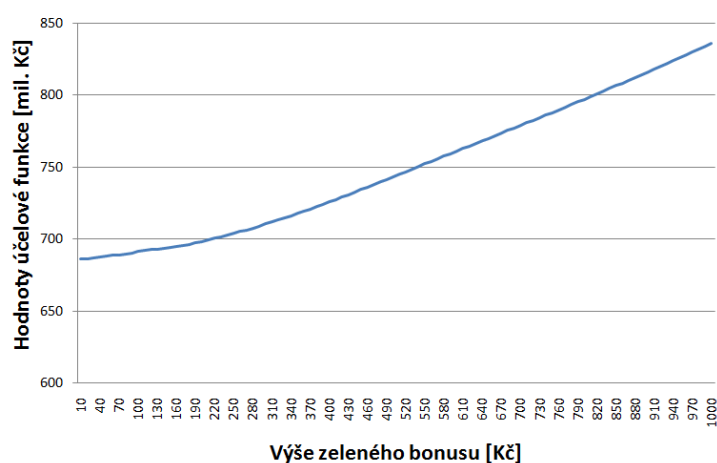
---- 3473 VARIABLE Z.L		hodnota ucelove funkce					
1	827713837,	2	825278011,	3	818931335,	4	809153028,
5	819490049,	6	810210738,	7	825124455,	8	819180608,
9	818117255,	10	820501476,	11	812254467,	12	814742111,
13	798694326,	14	829905748,	15	827450239,	16	825565636,
17	839932905,	18	819678607,	19	816671317,	20	813218914,
21	835611172,	22	816857121,	23	819229248,	24	806900699,
25	818930788,	26	820344606,	27	821845805,	28	816573293,
29	836734454,	30	827747414,	31	830886137,	32	820090984,
33	831133237,	34	812429276,	35	814607931,	36	830064630,
37	833532503,	38	822606121,	39	829491839,	40	799357457,
41	831968519,	42	839639161,	43	818920361,	44	817637626,
45	824556594,	46	821848478,	47	812304760,	48	816915763,
49	828695942,	50	828905094,	51	822275398,	52	824636989,
53	831729389,	54	826205138,	55	828507274,	56	830584362,
57	824838314,	58	841481760,	59	837478343,	60	828021578,
61	826098899,	62	826372571,	63	829672568,	64	837365514,
65	829324259,	66	821156370,	67	822591623,	68	823472281,
69	838539468,	70	828442369,	71	825406761,	72	832835691,
73	814181024,	74	825316469,	75	833312134,	76	824753223,
77	822679035,	78	828743601,	79	829737572,	80	830111821,
81	836168097,	82	822755169,	83	814117924,	84	818853769,
85	831320450,	86	829090376,	87	823428180,	88	820386816,
89	816686015,	90	840205237,	91	814080570,	92	822142662,
93	826805965,	94	805040381,	95	820582006,	96	830884220,
97	826446564,	98	833245920,	99	829341533,	100	828560216.

Toto jsou hodnoty účelové funkce pro 100 vygenerovaných poptávek, vypočítaných pomocí optimalizačního výpočtového softwaru GAMS.

7.3 Vývoj cen dotací na výrobu elektřiny z biomasy

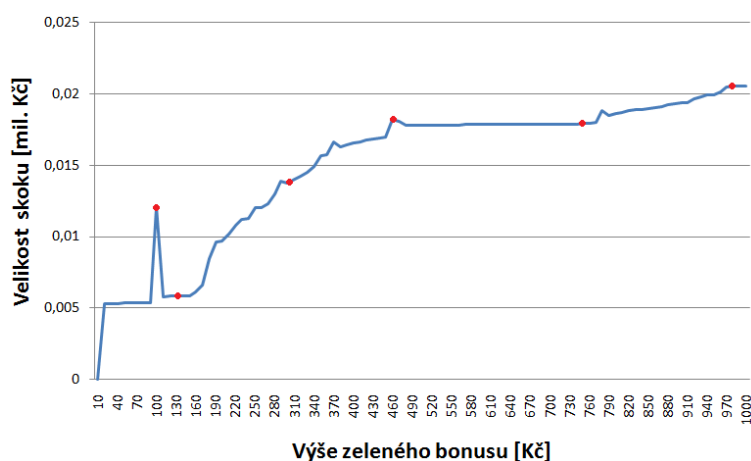
Jak jsme již uvedli v kapitole 2, náš stát se zavázal splnit do roku 2020 indikativní cíl (procentuální podíl OZE na hrubé domácí spotřebě elektřiny) ve výši 13%. Aby byl tento cíl splněn, je zde snaha o podporu OZE, např. pro nás zajímavé využívání biomasy při výrobě energií. Výše dotace získávané od státu za výrobu elektřiny z biomasy (tzv. zelené bonusy) je při tom jeden z rozhodujících parametrů, bez finanční podpory OZE by se využití biomasy za současných podmínek příliš ekonomicky nevyplatilo.

Nyní je výše dotace 750 Kč za MWh elektrické energie vyrobené z biomasy. Provedli jsme analýzu citlivosti zisku (velikosti účelové funkce) na výši zelených bonusů od státu. Naši optimalizační úlohu jsme řešili pro různé hodnoty zelených bonusů. Výši bonusů jsme uvažovali v rozmezí 0-1000 Kč, kde jsme zvyšovali hodnotu po 10 Kč. Na Obrázku 15 vidíme závislost velikosti zisku na hodnotě zelených bonusů za výrobu elektřiny z biomasy.



Obrázek 15: Závislost velikosti zisku na výši zeleného bonusu

Následující Obrázek 16 ukazuje velikost rozdílu zisku (skoku) při zvýšení dotace o 10 Kč.



Obrázek 16: Závislost rozdílu zisku na výši zeleného bonusu

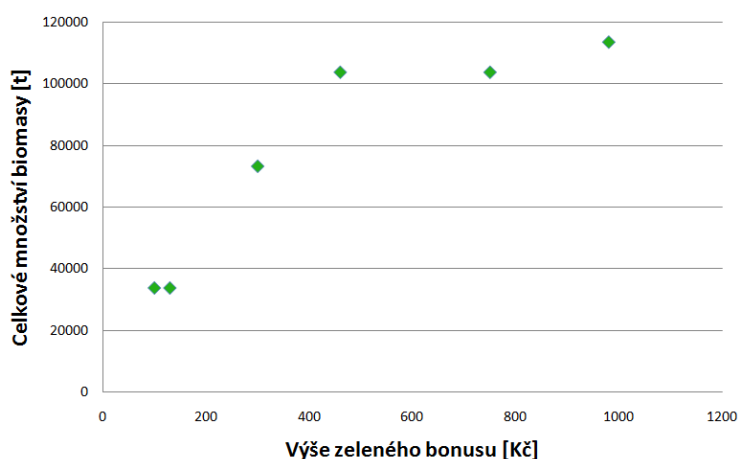
Z těchto závěrů odvodíme scénáře ξ_D pro modelování nejisté výše dotace od státu. V úvahu bereme pouze „zajímavé“ hodnoty a ty považujeme za scénáře výše dotací ξ_D^s . Na Obrázku

16 jsou scénářové hodnoty vyznačeny červenou tečkou. V Tabulce 5 uvádíme hodnoty dotací jednotlivých scénářů a pro ně hodnoty účelové funkce.

Tabulka 5: Jednotlivé scénáře výše zelených bonusů

scénář	ξ_d^1	ξ_d^2	ξ_d^3	ξ_d^4	ξ_d^5	ξ_d^6
dotace	100	130	300	460	750	980
Z	691345782	692811250	710401164	736250413	787432225	832158816

Na následujícím obrázku je uvedeno celkové množství dávkované biomasy pro jednotlivé scénáře hodnot zeleného bonusu.



Obrázek 17: Závislost množství biomasy na výši zeleného bonusu

Nyní jsme si popsali nejdůležitější neurčitosti našeho modelu (účinnost kotlů, velikost poptávky po teple, výše dotací za výrobu elektrické energie z biomasy) a nyní si zavedeme stochastický model teplárny, do něhož zapracujeme vygenerované scénáře velikosti poptávky a výše dotace od státu. V modelu také nahradíme historické hodnoty účinnosti zjištěnou a verifikovanou regresní funkcí.

8 Stochastický model teplárny

8.1 WS model teplárny

Nyní se vrátíme k deterministickému modelu z kapitoly 3, do kterého zahrneme již popsané neurčitosti z předchozí kapitoly. Náhodný parametr ξ_P značí náhodnou velikost poptávky a náhodný parametr ξ_D vývoj cen dotací na výrobu elektřiny z biomasy při spalování s uhlím (zelené bonusy). Také k výpočtům využijeme verifikované funkce účinnosti jednotlivých kotlů. Stochastický model teplárny je tvaru

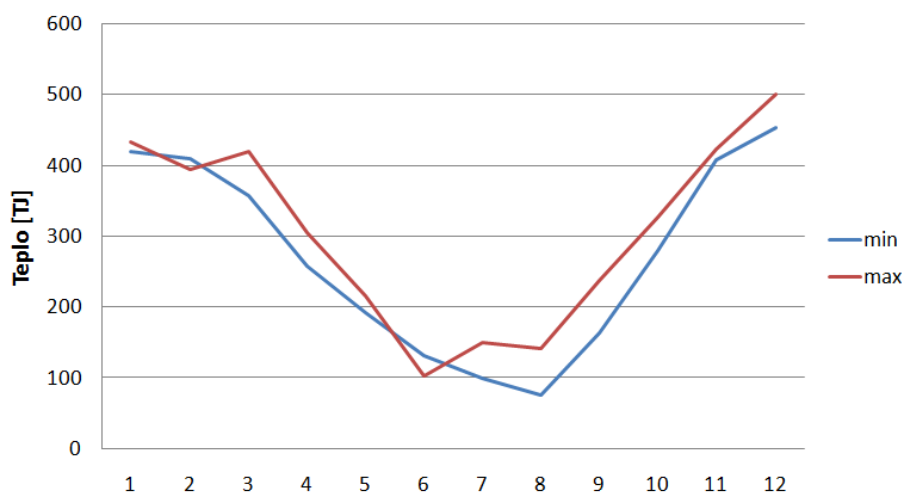
$$\begin{aligned}
 & \max_{\mathbf{x}, \mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{y}} \{d_D(\mathbf{y}(\xi_D, \xi_P), \pi, \xi_D, \xi_P) + d_G(\mathbf{y}(\xi_D, \xi_P), \mu, \xi_D, \xi_P) - c(\mathbf{x}(\xi_D, \xi_P), \rho, \xi_D, \xi_P)\} \\
 & \mathbf{E}_x \mathbf{x}(\xi_D, \xi_P) \leq \mathbf{b}_x, \\
 & \mathbf{0} \leq \mathbf{x}(\xi_D, \xi_P) \in \mathfrak{R}^{n_J \cdot n_K}, \\
 & \mathbf{b}_L \mathbf{1}^T \mathbf{x}(\xi_D, \xi_P) \leq \mathbf{x}(\xi_D, \xi_P) \leq \mathbf{b}_U \mathbf{1}^T \mathbf{x}(\xi_D, \xi_P), \\
 & \mathbf{u}(\xi_D, \xi_P) = \mathbf{f}(\mathbf{x}(\xi_D, \xi_P), \alpha, \xi_D, \xi_P), \\
 & \mathbf{0} \leq \mathbf{u}(\xi_D, \xi_P) \in \mathfrak{R}^{n_J \cdot n_K}, \\
 & \mathbf{E}_u \mathbf{u}(\xi_D, \xi_P) \leq \mathbf{b}_C, \\
 & \mathbf{v}(\xi_D, \xi_P) = \mathbf{g}(\mathbf{u}(\xi_D, \xi_P), \eta(\mathbf{x}(\xi_D, \xi_P), \alpha), \xi_D, \xi_P), \\
 & \mathbf{0} \leq \mathbf{v}(\xi_D, \xi_P) \in \mathfrak{R}^{n_J \cdot n_K}, \\
 & \mathbf{e}_v^T \mathbf{v}(\xi_D, \xi_P) = \mathbf{e}_w^T \mathbf{w}(\xi_D, \xi_P), \\
 & \mathbf{0} \leq \mathbf{w}(\xi_D, \xi_P) \in W \in \mathfrak{R}^{n_J \cdot n_K \cdot n_I}, \\
 & \mathbf{E}_w \mathbf{y}(\xi_D, \xi_P) \in \{\mathbf{H}(\mathbf{w}(\xi_D, \xi_P), \gamma, \xi_D, \xi_P), \mathbf{w} \in W\}, \\
 & \mathbf{0} \leq \mathbf{y}(\xi_D, \xi_P) \in \mathfrak{R}^{n_J \cdot n_K \cdot n_I \cdot n_M}, \\
 & \mathbf{E}_v \mathbf{y}(\xi_D, \xi_P) = \mathbf{E}_K (\mathbf{I} - \text{diag}(\text{sgn } \mathbf{e}_v))^T \mathbf{v}(\xi_D, \xi_P), \\
 & \mathbf{b}_{DL} \leq \mathbf{E}_D \mathbf{y}(\xi_D, \xi_P) \leq \mathbf{b}_{DU}.
 \end{aligned}$$

Tento model řešíme pro 100 scénářů poptávky ξ_P a pro 6 scénářů výše zelených bonusů ξ_D . Jelikož známe hodnoty náhodných parametrů ještě před naším rozhodnutím (dávkování paliv), jde o WS přístup stochastického programování (kapitola 4). Pro názornost si na následujících obrázcích ukážeme výsledky pouze pro jeden konkrétní scénář cen dotací za výrobu elektřiny pomocí OZE a pro všechny scénáře poptávky.

8.2 Analýza výsledků

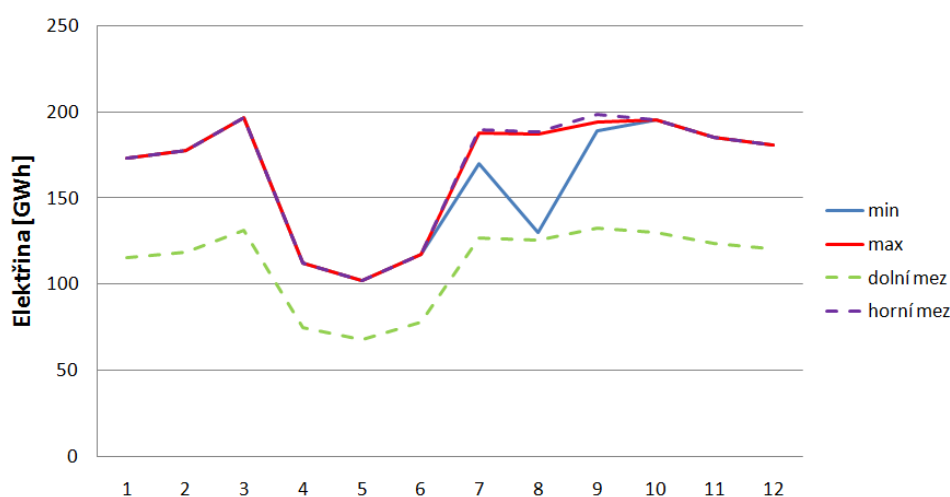
Pro ukázkou jsme vybrali pro nejbližší budoucnost nejpravděpodobnější ze scénářů dotací a to nynější výši zeleného bonusu 750 Kč/GWh vyrobené elektřiny z biomasy. Pro prezentaci výsledků jsme si vybrali 2 poptávky. Tu které měla v součtu minimální a maximální hodnotu poptávaného tepla za rok a pro tyto poptávky jsme srovnali plán provozu teplárny, tj. výrobu tepla a elektřiny, dávkování všech typů paliv a množství mařeného tepla v průběhu celého roku.

Na Obrázku 18 jsou vykreslené vybrané scénáře poptávky po teple, a to scénář s minimálním a maximálním ročním úhrnem poptávaného tepla, jejichž vygenerování jsme si popsali v kapitole 7.2.



Obrázek 18: Velikost maximální a minimální poptávky po teple

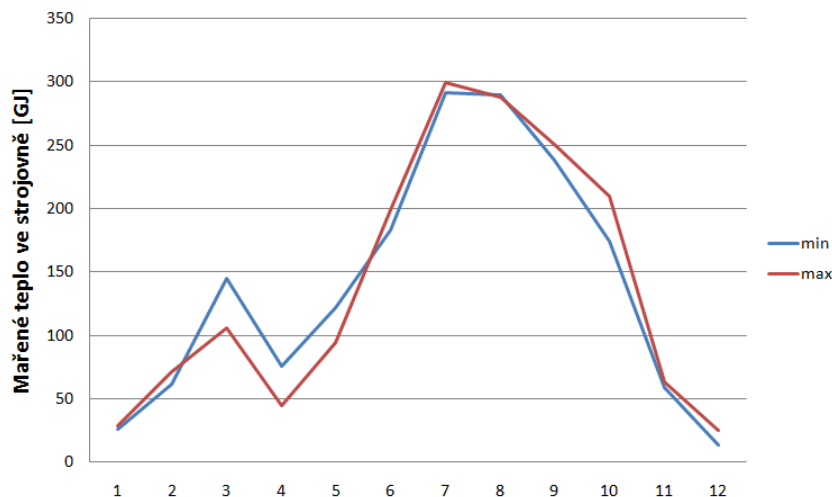
Pro tyto poptávky jsme na Obrázku 19 vykreslili množství vyrobené elektřiny. Jak vidíme ve většině měsíců je množství elektřiny stejné pro minimální i maximální poptávku a pohybuje se na horní možné hranici (meze představují možnost uplatnění vyrobené elektřiny v síti). Ale je zajímavé, že v létě (7. a 8. měsíc) výroba elektřiny pro minimální poptávku výrazně klesne téměř až na dolní mez. To nastává vlivem malé poptávky po teple v daném období. Rozdíl je pravděpodobně způsoben nízkým využitím protitlakové turbíny. Nízký odběr tepla umožňuje pouze omezené využití protitlakové turbíny a tím pádem klesá výroba elektřiny na této turbíně. To je dáno zapojením turbín ve strojovně (Obrázek 4). Tedy výše vyrobené elektřiny je omezená množstvím dodávaného tepla do CZT.



Obrázek 19: Export elektřiny při minimální a maximální poptávce po teple

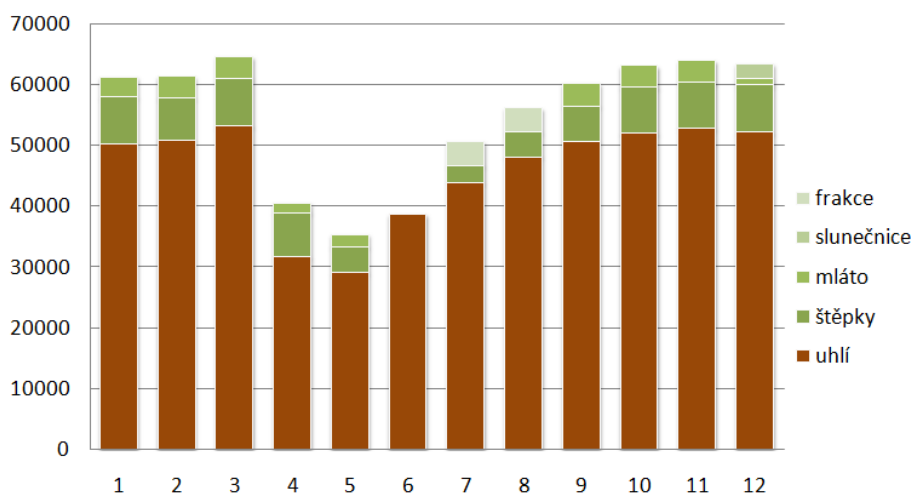
Jako další si v souvislosti s minimální a maximální poptávkou uvedeme na Obrázku 20 výši mařeného tepla ve strojovně. Z technologie kondenzačního stupně je nutné část páry

mařit bez dalšího efektivního využití. Toto maření se provádí v kondenzátoru (Obrázek 4). Z grafu je vidět, že v letních měsících, kdy je poptávka po teple nižší se více maří, tedy využívá se více kondenzační turbína pro výrobu elektřiny. Rozdíl je pravděpodobně způsoben nízkým využitím protitlakové turbíny. Nízký odběr tepla umožňuje pouze omezené využití protitlakové turbíny a tím pádem klesá výroba elektřiny na této turbíně.

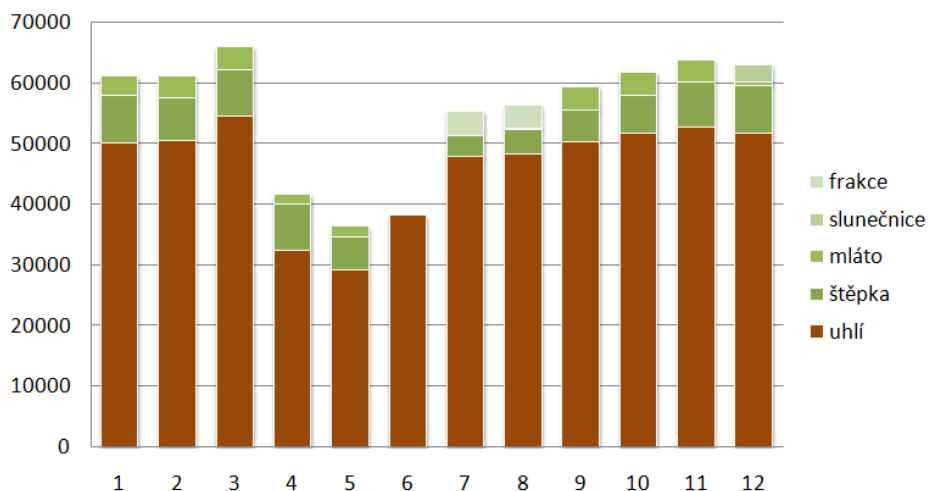


Obrázek 20: Množství mařeného tepla pro minimální a maximální poptávku

Obrázky 21 a 22 znázorňují dávkování paliv opět pro minimální a maximální poptávku po teple z Obrázku 18. Rozdíl mezi maximálním a minimálním ročním úhrnem poptávaného tepla je 400,227 TJ. Při porovnání celkového objemu spáleného paliva se množství uhlí zvýší o 4004,184 t a množství biomasy o 385,545 t oproti množství paliv dávkovaných pro minimální poptávku. V obou případech se dávkuje veškeré dostupné množství dřevní štěpky, mláta a podsítné frakce. Pro maximální poptávku se dávkuje v prosinci více slunečnicového šrotu o 385,545 t. Je to způsobeno tím, že z finančního hlediska se nejprve dávkuje levnější typy biomasy (mláto, štěpka, frakce) a až po jejich vyčerpání se využívají dražší (slunečnice).

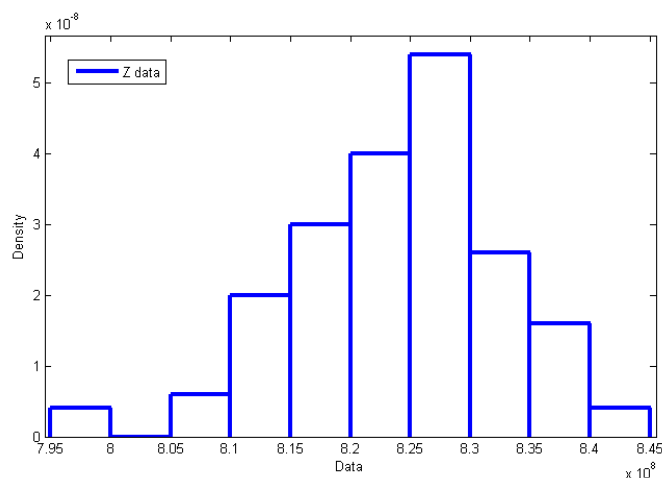


Obrázek 21: Množství paliv při minimální poptávce



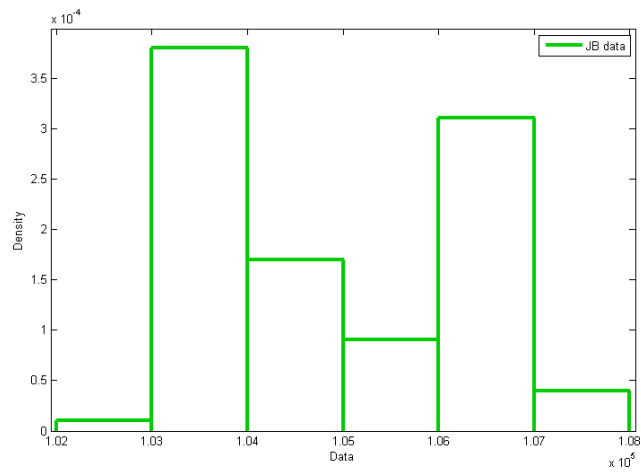
Obrázek 22: Množství paliv při maximální poptávce

Na Obrázku 23 je vykreslen histogram hodnot účelové funkce Z pro 100 vygenerovaných poptávek. Z tohoto histogramu je vidět, že nemůžeme dopředu usuzovat o pravděpodobnostním rozdělení výsledku (možné normální rozdělení), jestliže známe pravděpodobnostní rozdělení vstupních veličin (rovnoměrné rozdělení poptávek). Tento výsledek nemusí mít obecnou platnost. Pro zjištění by byla potřeba důkladnější analýza, která není v tuto chvíli, pro naši potřebu nutná.



Obrázek 23: Histogram hodnot účelové funkce Z

Dále jsme sestrojili histogram množství spálené biomasy (Obrázek 24). Zjistili jsme, že množství vyrobeného tepla nemá výrazný vliv na dávkování biomasy. Jak vidíme všechny hodnoty se pohybují v rozmezí 102 až 107 tisíc tun pro všechny scénáře poptávek. Tento výsledek je velice důležitý, protože můžeme předpokládat pro neznámou poptávku optimální dávkování v tomto poměrně úzkém rozmezí. Z tohoto předpokladu si teplárna může dopředu naplánovat objednávku daného množství biomasy a tím získat výhodnou pořizovací cenu.



Obrázek 24: Histogram množství dávkované biomasy pro 100 poptávek

9 Závěr

Cílem práce bylo přispět k ekonomicky efektivnímu využití biomasy při spalování s uhlím pro reálnou teplárnu, jejímž požadavkem bylo optimální využití stávající palivové základny a naplánování ročního provozu vzhledem k měsíční poptávce po teple.

Tato práce navazuje na deterministický model, který je uveden v kapitole 3. Pro přesnější popis modelu teplárny, který obsahuje jisté neurčitosti jako každý reálný proces, jsme si zavedli potřebné pojmy (stochastické programování - kapitola 4, regresní analýza - kapitoly 5 a 6). V řešeném problému se vykytovaly tyto neurčitosti:

- velikost poptávky po teple,
- výše zeleného bonusu,
- účinnost kotlů.

Popis a modelování těchto neurčitostí jsou podrobně popsány v kapitole 7, které v následující kapitole implementujeme do stochastického modelu reálné teplárny. Používáme Wait-and-See přístup stochastického programování, jelikož velikost poptávky po teple a výši zeleného bonusu v budoucnu neznáme, ale aktuální hodnoty jsou známy a tedy na ně můžeme reagovat ještě před naším rozhodnutím a upravit ho. Rozhodnutí teplárny spočívá ve správném dávkování paliv, kdy zohledňujeme jejich dostupné množství, nákupní cenu, výhřevnost, ale také technologické omezení, za cílem maximálního zisku. Zisk teplárny se skládá z prodeje vyrobeného tepla a elektřiny a dále ze zelených bonusů za výrobu elektřiny z OZE.

Tento stochastický model jsme implementovali do optimalizačního softwaru GAMS, který nám v této fázi určí při známé poptávce přesné dávkování paliv, které nám zaručí maximální možný zisk.

Z analýzy výsledků jsme zjistili, že pro všechny vygenerované poptávky po teple (kapitola 7.2) a zelený bonus ve výši 750 Kč, je optimální množství dávkované biomasy v rozmezí 102 000 - 108 000 tun (Obrázek 24). Z toho můžeme vyvodit závěr, že teplárna může plánovat stabilní smluvní podmínky na pořizování biomasových paliv i při neznámé poptávce po teple. Z porovnání mezních poptávek po teple (kapitola 8.2) a jim odpovídajících poměrů dávkování biomasy a uhlí, jsme zjistili zanedbatelné rozdíly. Z podrobného zkoumání jsme usoudili, že při vyšších měsíčních poptávkách v zimním období se zvyšuje poměr biomasy.

Při výrobě elektřiny se pro všechny poptávky držíme na horní hranici po většinu roku (Obrázek 19), tato hranice nám představuje možnost uplatnění vyrobené elektřiny v síti. Pod tuto hranici klesáme pouze při nízké poptávce po teple, což nastává většinou v letních měsících.

Reference

- [1] *Směrnice EU 2009/28/ES [PDF]*. 2009, [cit. 2011-02-10]. <[http://www.mzp.cz/ris/ais-risdb-ec-table.nsf/C006392F4EC5C218C12575D70049B057/\\$file/32009L0028.pdf](http://www.mzp.cz/ris/ais-risdb-ec-table.nsf/C006392F4EC5C218C12575D70049B057/$file/32009L0028.pdf)>.
- [2] BUFKA, Aleš, et al. *Zpráva o plnění indikativního cíle výroby elektřiny z obnovitelných zdrojů za rok 2008*. Ministerstvo průmyslu a obchodu ČR, 2009. 29 s.
- [3] TOUŠ, Michal; POPELA, Pavel; PAVLAS, Martin; STEHLÍK, Petr; DRÁPELA, Tomáš. *A stochastic programming approach to optimization of combustion plant utilizing coal and biomass*. In proceedings Mendel 2010.
- [4] HÁJEK, Zdeněk. *Náhrada fosilních paliv ve velkých energetických zdrojích*. Brno: Vysoké učení technické v Brně, Fakulta strojního inženýrství, 2010. 92 s.
- [5] ANDĚL, Jiří. *Základy matematické statistiky*. 2. opr. vyd. Praha: Matfyzpress, 2007. 358 s. ISBN 978-807-3780-012.
- [6] HEBÁK, Petr; HUSTOPECKÝ, Jiří; MALÁ, Ivana. *Vícerozměrné statistické metody [2]*. 1. vyd. Praha: Informatorium, 2005. 239 s. ISBN 80-733-3036-9.
- [7] HENDL, Jan. *Přehled statistických metod zpracování dat: analýza a metaanalýza dat*. 1. vyd. Praha: Portál, 2004. 583 s. ISBN 80-717-8820-1.
- [8] MELOUN, Milan; MILITKÝ, Jiří. *Statistické zpracování experimentálních dat*. Praha: PLUS, 1994. 839 s. ISBN 80-852-9756-6.
- [9] KARPÍŠEK, Zdeněk. *Statistická analýza*. Brno: Vysoké učení technické v Brně, Fakulta strojního inženýrství, 2008. 117 s.
- [10] KARPÍŠEK, Zdeněk. *Regresní diagnostika*. Brno: Vysoké učení technické v Brně, Fakulta strojního inženýrství. 4 s.
- [11] POPELA, Pavel. *Stochastic programming*. Malta: University of Malta, Faculty of Science. 2004. 72 s.
- [12] KLIMEŠ, Lubomír. *Stochastic programming algorithms*. Brno: Vysoké učení technické v Brně, Fakulta strojního inženýrství, 2010. 95 s.

- [13] KALL, Peter and WALLACE, Stein W. *Stochastic Programming*. John Wiley & Sons, Chichester, 1994.
- [14] ŠEDIVÁ, Blanka. *Rozhodování o normalitě dat [PDF]*. [cit. 2011-03-11]. <<http://home.zcu.cz/~sediva/stav/kap05.pdf>>.
- [15] MELOUN, Milan. *Multikolinearita a polynomy [PDF]*. [cit. 2011-03-17]. <<http://meloun.upce.cz/docs/lecture/chemometrics/slidy/64multikol.pdf>>.
- [16] MELOUN, Milan. *Kritika metody v regresním tripletu [PDF]*. [cit. 2011-03-11]. <<http://meloun.upce.cz/docs/lecture/chemometrics/slidy/66metoda.pdf>>.
- [17] TVRDÝ, Lubor. *Korelační a regresní analýza*. VŠB-TU Ostrava, Ekonomická fakulta. 50 s.
- [18] JAROŠOVÁ, Eva; KRÁL, Jan. *Ověřování předpokladů normality [PDF]*. [cit. 2011-04-02]. Národní informační středisko pro podporu jakosti. <http://isq.cz/npj/2006/05%20-%2012%20-%20Testy_normality.pdf>.
- [19] HELLEBRANDOVÁ, Iveta. *Regrese s $AR(p)$ chybami*. Brno: Masarykova univerzita, Přírodovědecká fakulta, 2009. 40 s.
- [20] KARPÍŠEK, Zdeněk. *Matematika IV: statistika a pravděpodobnost*. 3. dopl. vyd. Brno: CERM, 2007. 170 s. ISBN 978-802-1433-809.

Příloha

Statistické tabulky

T1 Kritické hodnoty Lillieforsova testu D_α

n	$\alpha = 0,01$	$\alpha = 0,05$	$\alpha = 0,1$	$\alpha = 0,15$	$\alpha = 0,2$
4	0,417	0,381	0,352	0,319	0,300
5	0,405	0,337	0,315	0,299	0,285
6	0,364	0,319	0,294	0,277	0,265
7	0,348	0,300	0,276	0,258	0,247
8	0,331	0,285	0,261	0,244	0,233
9	0,311	0,271	0,249	0,233	0,223
10	0,294	0,258	0,239	0,224	0,215
11	0,284	0,249	0,230	0,217	0,206
12	0,275	0,242	0,223	0,212	0,199
13	0,268	0,234	0,214	0,202	0,190
14	0,261	0,227	0,207	0,194	0,183
15	0,257	0,220	0,201	0,187	0,177
16	0,250	0,213	0,195	0,182	0,173
17	0,245	0,206	0,189	0,177	0,169
18	0,239	0,200	0,184	0,173	0,166
19	0,235	0,195	0,179	0,169	0,163
20	0,231	0,190	0,174	0,166	0,160
25	0,200	0,173	0,158	0,147	0,142
30	0,187	0,161	0,144	0,136	0,131
n	$1,031/\sqrt{n}$	$0,886/\sqrt{n}$	$0,805/\sqrt{n}$	$0,768/\sqrt{n}$	$0,736/\sqrt{n}$

T2 Kvantily t_P Studentova rozdělení $S(k)$

$k \backslash P$	0,95	0,975	0,99	0,995	0,999	0,9995
1	6,314	12,706	31,821	63,656	318,289	636,578
2	2,920	4,303	6,965	9,925	22,328	31,600
3	2,353	3,182	4,541	5,841	10,214	12,924
4	2,132	2,776	3,747	4,604	7,173	8,610
5	2,015	2,571	3,365	4,032	5,894	6,869
6	1,943	2,447	3,143	3,707	5,208	5,959
7	1,895	2,365	2,998	3,499	4,785	5,408
8	1,860	2,306	2,896	3,355	4,501	5,041
9	1,833	2,262	2,821	3,250	4,297	4,781
10	1,812	2,228	2,764	3,169	4,144	4,587
11	1,796	2,201	2,718	3,106	4,025	4,437
12	1,782	2,179	2,681	3,055	3,930	4,318
13	1,771	2,160	2,650	3,012	3,852	4,221
14	1,761	2,145	2,624	2,977	3,787	4,140
15	1,753	2,131	2,602	2,947	3,733	4,073
16	1,746	2,120	2,583	2,921	3,686	4,015
17	1,740	2,110	2,567	2,898	3,646	3,965
18	1,734	2,101	2,552	2,878	3,610	3,922
19	1,729	2,093	2,539	2,861	3,579	3,883
20	1,725	2,086	2,528	2,845	3,552	3,850
21	1,721	2,080	2,518	2,831	3,527	3,819
22	1,717	2,074	2,508	2,819	3,505	3,792
23	1,714	2,069	2,500	2,807	3,485	3,768
24	1,711	2,064	2,492	2,797	3,467	3,745
25	1,708	2,060	2,485	2,787	3,450	3,725
26	1,706	2,056	2,479	2,779	3,435	3,707
27	1,703	2,052	2,473	2,771	3,421	3,689
28	1,701	2,048	2,467	2,763	3,408	3,674
29	1,699	2,045	2,462	2,756	3,396	3,660
30	1,697	2,042	2,457	2,750	3,385	3,646
35	1,690	2,030	2,438	2,724	3,340	3,591
40	1,684	2,021	2,423	2,704	3,307	3,551
45	1,679	2,014	2,412	2,690	3,281	3,520
50	1,676	2,009	2,403	2,678	3,261	3,496
60	1,671	2,000	2,390	2,660	3,232	3,460
70	1,667	1,994	2,381	2,648	3,211	3,435
80	1,664	1,990	2,374	2,639	3,195	3,416
90	1,662	1,987	2,368	2,632	3,183	3,402
100	1,660	1,984	2,364	2,626	3,174	3,390
120	1,658	1,980	2,358	2,617	3,160	3,373
140	1,656	1,977	2,353	2,611	3,149	3,361
160	1,654	1,975	2,350	2,607	3,142	3,352
180	1,653	1,973	2,347	2,603	3,136	3,345
200	1,653	1,972	2,345	2,601	3,131	3,340
300	1,650	1,968	2,339	2,592	3,118	3,323
500	1,648	1,965	2,334	2,586	3,107	3,310
1000	1,646	1,962	2,330	2,581	3,098	3,300
∞	1,645	1,960	2,326	2,576	3,090	3,290

Poznámka: Pro $0 < P \leq 0,5$ použijeme vztah $t_P = -t_{1-P}$.

T3 Kvantily χ^2_P Pearsonova rozdělení $\chi^2(k)$

$k \backslash P$	0,005	0,01	0,025	0,05	0,95	0,975	0,99	0,995
1	0,000	0,000	0,001	0,004	3,841	5,024	6,635	7,879
2	0,010	0,020	0,051	0,103	5,991	7,378	9,210	10,597
3	0,072	0,115	0,216	0,352	7,815	9,348	11,345	12,838
4	0,207	0,297	0,484	0,711	9,488	11,143	13,277	14,860
5	0,412	0,554	0,831	1,145	11,070	12,832	15,086	16,750
6	0,676	0,872	1,237	1,635	12,592	14,449	16,812	18,548
7	0,989	1,239	1,690	2,167	14,067	16,013	18,475	20,278
8	1,344	1,647	2,180	2,733	15,507	17,535	20,090	21,955
9	1,735	2,088	2,700	3,325	16,919	19,023	21,666	23,589
10	2,156	2,558	3,247	3,940	18,307	20,483	23,209	25,188
11	2,603	3,053	3,816	4,575	19,675	21,920	24,725	26,757
12	3,074	3,571	4,404	5,226	21,026	23,337	26,217	28,300
13	3,565	4,107	5,009	5,892	22,362	24,736	27,688	29,819
14	4,075	4,660	5,629	6,571	23,685	26,119	29,141	31,319
15	4,601	5,229	6,262	7,261	24,996	27,488	30,578	32,801
16	5,142	5,812	6,908	7,962	26,296	28,845	32,000	34,267
17	5,697	6,408	7,564	8,672	27,587	30,191	33,409	35,718
18	6,265	7,015	8,231	9,390	28,869	31,526	34,805	37,156
19	6,844	7,633	8,907	10,117	30,144	32,852	36,191	38,582
20	7,434	8,260	9,591	10,851	31,410	34,170	37,566	39,997
21	8,034	8,897	10,283	11,591	32,671	35,479	38,932	41,401
22	8,643	9,542	10,982	12,338	33,924	36,781	40,289	42,796
23	9,260	10,196	11,689	13,091	35,172	38,076	41,638	44,181
24	9,886	10,856	12,401	13,848	36,415	39,364	42,980	45,558
25	10,520	11,524	13,120	14,611	37,652	40,646	44,314	46,928
26	11,160	12,198	13,844	15,379	38,885	41,923	45,642	48,290
27	11,808	12,878	14,573	16,151	40,113	43,195	46,963	49,645
28	12,461	13,565	15,308	16,928	41,337	44,461	48,278	50,994
29	13,121	14,256	16,047	17,708	42,557	45,722	49,588	52,335
30	13,787	14,953	16,791	18,493	43,773	46,979	50,892	53,672
31	14,458	15,655	17,539	19,281	44,985	48,232	52,191	55,002
32	15,134	16,362	18,291	20,072	46,194	49,480	53,486	56,328
33	15,815	17,073	19,047	20,867	47,400	50,725	54,775	57,648
34	16,501	17,789	19,806	21,664	48,602	51,966	56,061	58,964
35	17,192	18,509	20,569	22,465	49,802	53,203	57,342	60,275
36	17,887	19,233	21,336	23,269	50,998	54,437	58,619	61,581
37	18,586	19,960	22,106	24,075	52,192	55,668	59,893	62,883
38	19,289	20,691	22,878	24,884	53,384	56,895	61,162	64,181
39	19,996	21,426	23,654	25,695	54,572	58,120	62,428	65,475
40	20,707	22,164	24,433	26,509	55,758	59,342	63,691	66,766
41	21,421	22,906	25,215	27,326	56,942	60,561	64,950	68,053
42	22,138	23,650	25,999	28,144	58,124	61,777	66,206	69,336
43	22,860	24,398	26,785	28,965	59,304	62,990	67,459	70,616
44	23,584	25,148	27,575	29,787	60,481	64,201	68,710	71,892
45	24,311	25,901	28,366	30,612	61,656	65,410	69,957	73,166

**T3 Kvantily χ^2_P Pearsonova rozdělení $\chi^2(k)$
(pokračování)**

$k \backslash P$	0,005	0,01	0,025	0,05	0,95	0,975	0,99	0,995
46	25,041	26,657	29,160	31,439	62,830	66,616	71,201	74,437
47	25,775	27,416	29,956	32,268	64,001	67,821	72,443	75,704
48	26,511	28,177	30,754	33,098	65,171	69,023	73,683	76,969
49	27,249	28,941	31,555	33,930	66,339	70,222	74,919	78,231
50	27,991	29,707	32,357	34,764	67,505	71,420	76,154	79,490
51	28,735	30,475	33,162	35,600	68,669	72,616	77,386	80,746
52	29,481	31,246	33,968	36,437	69,832	73,810	78,616	82,001
53	30,230	32,019	34,776	37,276	70,993	75,002	79,843	83,253
54	30,981	32,793	35,586	38,116	72,153	76,192	81,069	84,502
55	31,735	33,571	36,398	38,958	73,311	77,380	82,292	85,749
56	32,491	34,350	37,212	39,801	74,468	78,567	83,514	86,994
57	33,248	35,131	38,027	40,646	75,624	79,752	84,733	88,237
58	34,008	35,914	38,844	41,492	76,778	80,936	85,950	89,477
59	34,770	36,698	39,662	42,339	77,930	82,117	87,166	90,715
60	35,534	37,485	40,482	43,188	79,082	83,298	88,379	91,952
61	36,300	38,273	41,303	44,038	80,232	84,476	89,591	93,186
62	37,068	39,063	42,126	44,889	81,381	85,654	90,802	94,419
63	37,838	39,855	42,950	45,741	82,529	86,830	92,010	95,649
64	38,610	40,649	43,776	46,595	83,675	88,004	93,217	96,878
65	39,383	41,444	44,603	47,450	84,821	89,177	94,422	98,105
66	40,158	42,240	45,431	48,305	85,965	90,349	95,626	99,330
67	40,935	43,038	46,261	49,162	87,108	91,519	96,828	100,554
68	41,714	43,838	47,092	50,020	88,250	92,688	98,028	101,776
69	42,493	44,639	47,924	50,879	89,391	93,856	99,227	102,996
70	43,275	45,442	48,758	51,739	90,531	95,023	100,425	104,215
71	44,058	46,246	49,592	52,600	91,670	96,189	101,621	105,432
72	44,843	47,051	50,428	53,462	92,808	97,353	102,816	106,647
73	45,629	47,858	51,265	54,325	93,945	98,516	104,010	107,862
74	46,417	48,666	52,103	55,189	95,081	99,678	105,202	109,074
75	47,206	49,475	52,942	56,054	96,217	100,839	106,393	110,285
80	51,172	53,540	57,153	60,391	101,879	106,629	112,329	116,321
85	55,170	57,634	61,389	64,749	107,522	112,393	118,236	122,324
90	59,196	61,754	65,647	69,126	113,145	118,136	124,116	128,299
95	63,250	65,898	69,925	73,520	118,752	123,858	129,973	134,247
100	67,328	70,065	74,222	77,929	124,342	129,561	135,807	140,170
110	75,550	78,458	82,867	86,792	135,480	140,916	147,414	151,948
120	83,852	86,923	91,573	95,705	146,567	152,211	158,950	163,648
130	92,223	95,451	100,331	104,662	157,610	163,453	170,423	175,278
150	109,142	112,668	117,985	122,692	179,581	185,800	193,207	198,360
200	152,241	156,432	162,728	168,279	233,994	241,058	249,445	255,264
500	422,303	429,387	439,936	449,147	553,127	563,851	576,493	585,206
1000	888,563	898,912	914,257	927,594	1074,68	1089,53	1106,97	1118,95

T4 Kvantily F_P Fisherova – Snedecorova rozdělení $F(k_1, k_2)$ pro $P = 0,975$

$k_1 \backslash k_2$	1	2	3	4	5	6	7	8	9	10
1	647,793	799,482	864,151	899,599	921,835	937,114	948,203	956,643	963,279	968,634
2	38,506	39,000	39,166	39,248	39,298	39,331	39,356	39,373	39,387	39,398
3	17,443	16,044	15,439	15,101	14,885	14,735	14,624	14,540	14,473	14,419
4	12,218	10,649	9,979	9,604	9,364	9,197	9,074	8,980	8,905	8,844
5	10,007	8,434	7,764	7,388	7,146	6,978	6,853	6,757	6,681	6,619
6	8,813	7,260	6,599	6,227	5,988	5,820	5,695	5,600	5,523	5,461
7	8,073	6,542	5,890	5,523	5,285	5,119	4,995	4,899	4,823	4,761
8	7,571	6,059	5,416	5,053	4,817	4,652	4,529	4,433	4,357	4,295
9	7,209	5,715	5,078	4,718	4,484	4,320	4,197	4,102	4,026	3,964
10	6,937	5,456	4,826	4,468	4,236	4,072	3,950	3,855	3,779	3,717
11	6,724	5,256	4,630	4,275	4,044	3,881	3,759	3,664	3,588	3,526
12	6,554	5,096	4,474	4,121	3,891	3,728	3,607	3,512	3,436	3,374
13	6,414	4,965	4,347	3,996	3,767	3,604	3,483	3,388	3,312	3,250
14	6,298	4,857	4,242	3,892	3,663	3,501	3,380	3,285	3,209	3,147
15	6,200	4,765	4,153	3,804	3,576	3,415	3,293	3,199	3,123	3,060
16	6,115	4,687	4,077	3,729	3,502	3,341	3,219	3,125	3,049	2,986
17	6,042	4,619	4,011	3,665	3,438	3,277	3,156	3,061	2,985	2,922
18	5,978	4,560	3,954	3,608	3,382	3,221	3,100	3,005	2,929	2,866
19	5,922	4,508	3,903	3,559	3,333	3,172	3,051	2,956	2,880	2,817
20	5,871	4,461	3,859	3,515	3,289	3,128	3,007	2,913	2,837	2,774
21	5,827	4,420	3,819	3,475	3,250	3,090	2,969	2,874	2,798	2,735
22	5,786	4,383	3,783	3,440	3,215	3,055	2,934	2,839	2,763	2,700
23	5,750	4,349	3,750	3,408	3,183	3,023	2,902	2,808	2,731	2,668
24	5,717	4,319	3,721	3,379	3,155	2,995	2,874	2,779	2,703	2,640
25	5,686	4,291	3,694	3,353	3,129	2,969	2,848	2,753	2,677	2,613
26	5,659	4,265	3,670	3,329	3,105	2,945	2,824	2,729	2,653	2,590
27	5,633	4,242	3,647	3,307	3,083	2,923	2,802	2,707	2,631	2,568
28	5,610	4,221	3,626	3,286	3,063	2,903	2,782	2,687	2,611	2,547
29	5,588	4,201	3,607	3,267	3,044	2,884	2,763	2,669	2,592	2,529
30	5,568	4,182	3,589	3,250	3,026	2,867	2,746	2,651	2,575	2,511
35	5,485	4,106	3,517	3,179	2,956	2,796	2,676	2,581	2,504	2,440
40	5,424	4,051	3,463	3,126	2,904	2,744	2,624	2,529	2,452	2,388
45	5,377	4,009	3,422	3,086	2,864	2,705	2,584	2,489	2,412	2,348
50	5,340	3,975	3,390	3,054	2,833	2,674	2,553	2,458	2,381	2,317
55	5,310	3,948	3,364	3,029	2,807	2,648	2,528	2,433	2,355	2,291
60	5,286	3,925	3,343	3,008	2,786	2,627	2,507	2,412	2,334	2,270
70	5,247	3,890	3,309	2,975	2,754	2,595	2,474	2,379	2,302	2,237
80	5,218	3,864	3,284	2,950	2,730	2,571	2,450	2,355	2,277	2,213
90	5,196	3,844	3,265	2,932	2,711	2,552	2,432	2,336	2,259	2,194
100	5,179	3,828	3,250	2,917	2,696	2,537	2,417	2,321	2,244	2,179
120	5,152	3,805	3,227	2,894	2,674	2,515	2,395	2,299	2,222	2,157
150	5,126	3,781	3,204	2,872	2,652	2,494	2,373	2,278	2,200	2,135
250	5,085	3,744	3,169	2,837	2,618	2,459	2,338	2,243	2,165	2,100
500	5,054	3,716	3,142	2,811	2,592	2,434	2,313	2,217	2,139	2,074
∞	5,024	3,689	3,116	2,786	2,566	2,408	2,288	2,192	2,114	2,048

**T4 Kvantily F_P Fisherova – Snedecorova rozdělení $F(k_1, k_2)$ pro $P = 0,975$
(pokračování)**

$k_2 \backslash k_1$	12	15	20	24	30	40	60	100	250	∞
1	976,725	984,874	993,081	997,272	1001,40	1005,60	1009,79	1013,16	1016,22	1018,26
2	39,415	39,431	39,448	39,457	39,465	39,473	39,481	39,488	39,494	39,498
3	14,337	14,253	14,167	14,124	14,081	14,036	13,992	13,956	13,924	13,902
4	8,751	8,657	8,560	8,511	8,461	8,411	8,360	8,319	8,282	8,257
5	6,525	6,428	6,329	6,278	6,227	6,175	6,123	6,080	6,041	6,015
6	5,366	5,269	5,168	5,117	5,065	5,012	4,959	4,915	4,876	4,849
7	4,666	4,568	4,467	4,415	4,362	4,309	4,254	4,210	4,170	4,142
8	4,200	4,101	3,999	3,947	3,894	3,840	3,784	3,739	3,698	3,670
9	3,868	3,769	3,667	3,614	3,560	3,505	3,449	3,403	3,361	3,333
10	3,621	3,522	3,419	3,365	3,311	3,255	3,198	3,152	3,109	3,080
11	3,430	3,330	3,226	3,173	3,118	3,061	3,004	2,956	2,912	2,883
12	3,277	3,177	3,073	3,019	2,963	2,906	2,848	2,800	2,755	2,725
13	3,153	3,053	2,948	2,893	2,837	2,780	2,720	2,671	2,626	2,595
14	3,050	2,949	2,844	2,789	2,732	2,674	2,614	2,565	2,519	2,487
15	2,963	2,862	2,756	2,701	2,644	2,585	2,524	2,474	2,427	2,395
16	2,889	2,788	2,681	2,625	2,568	2,509	2,447	2,396	2,349	2,316
17	2,825	2,723	2,616	2,560	2,502	2,442	2,380	2,329	2,280	2,247
18	2,769	2,667	2,559	2,503	2,445	2,384	2,321	2,269	2,220	2,187
19	2,720	2,617	2,509	2,452	2,394	2,333	2,270	2,217	2,167	2,133
20	2,676	2,573	2,464	2,408	2,349	2,287	2,223	2,170	2,120	2,085
21	2,637	2,534	2,425	2,368	2,308	2,246	2,182	2,128	2,077	2,042
22	2,602	2,498	2,389	2,332	2,272	2,210	2,145	2,090	2,039	2,003
23	2,570	2,466	2,357	2,299	2,239	2,176	2,111	2,056	2,004	1,968
24	2,541	2,437	2,327	2,269	2,209	2,146	2,080	2,024	1,972	1,935
25	2,515	2,411	2,300	2,242	2,182	2,118	2,052	1,996	1,942	1,906
26	2,491	2,387	2,276	2,217	2,157	2,093	2,026	1,969	1,915	1,878
27	2,469	2,364	2,253	2,195	2,133	2,069	2,002	1,945	1,891	1,853
28	2,448	2,344	2,232	2,174	2,112	2,048	1,980	1,922	1,867	1,829
29	2,430	2,325	2,213	2,154	2,092	2,028	1,959	1,901	1,846	1,807
30	2,412	2,307	2,195	2,136	2,074	2,009	1,940	1,882	1,826	1,787
35	2,341	2,235	2,122	2,062	1,999	1,932	1,861	1,801	1,743	1,702
40	2,288	2,182	2,068	2,007	1,943	1,875	1,803	1,741	1,680	1,637
45	2,248	2,141	2,026	1,965	1,900	1,831	1,757	1,694	1,631	1,586
50	2,216	2,109	1,993	1,931	1,866	1,796	1,721	1,656	1,592	1,545
55	2,190	2,083	1,967	1,904	1,838	1,768	1,692	1,625	1,559	1,511
60	2,169	2,061	1,944	1,882	1,815	1,744	1,667	1,599	1,532	1,482
70	2,136	2,028	1,910	1,847	1,779	1,707	1,628	1,558	1,488	1,436
80	2,111	2,003	1,884	1,820	1,752	1,679	1,599	1,527	1,455	1,400
90	2,092	1,983	1,864	1,800	1,731	1,657	1,576	1,503	1,428	1,371
100	2,077	1,968	1,849	1,784	1,715	1,640	1,558	1,483	1,407	1,347
120	2,055	1,945	1,825	1,760	1,690	1,614	1,530	1,454	1,374	1,310
150	2,032	1,922	1,801	1,736	1,665	1,588	1,502	1,423	1,340	1,271
250	1,997	1,886	1,764	1,697	1,625	1,546	1,457	1,374	1,282	1,201
500	1,971	1,859	1,736	1,669	1,596	1,515	1,423	1,336	1,235	1,137
∞	1,945	1,833	1,708	1,640	1,566	1,484	1,388	1,296	1,183	1,000

T4 Kvantily F_P Fisherova – Snedecorova rozdělení $F(k_1, k_2)$ pro $P = 0,995$

$k_1 \backslash k_2$	1	2	3	4	5	6	7	8	9	10
1	16212,5	19997,4	21614,1	22500,8	23055,8	23439,5	23715,2	23923,8	24091,5	24221,8
2	198,503	199,012	199,158	199,245	199,303	199,332	199,361	199,376	199,390	199,390
3	55,552	49,800	47,468	46,195	45,391	44,838	44,434	44,125	43,881	43,685
4	31,332	26,284	24,260	23,154	22,456	21,975	21,622	21,352	21,138	20,967
5	22,785	18,314	16,530	15,556	14,939	14,513	14,200	13,961	13,772	13,618
6	18,635	14,544	12,917	12,028	11,464	11,073	10,786	10,566	10,391	10,250
7	16,235	12,404	10,883	10,050	9,522	9,155	8,885	8,678	8,514	8,380
8	14,688	11,043	9,597	8,805	8,302	7,952	7,694	7,496	7,339	7,211
9	13,614	10,107	8,717	7,956	7,471	7,134	6,885	6,693	6,541	6,417
10	12,827	9,427	8,081	7,343	6,872	6,545	6,303	6,116	5,968	5,847
11	12,226	8,912	7,600	6,881	6,422	6,102	5,865	5,682	5,537	5,418
12	11,754	8,510	7,226	6,521	6,071	5,757	5,524	5,345	5,202	5,085
13	11,374	8,186	6,926	6,233	5,791	5,482	5,253	5,076	4,935	4,820
14	11,060	7,922	6,680	5,998	5,562	5,257	5,031	4,857	4,717	4,603
15	10,798	7,701	6,476	5,803	5,372	5,071	4,847	4,674	4,536	4,424
16	10,576	7,514	6,303	5,638	5,212	4,913	4,692	4,521	4,384	4,272
17	10,384	7,354	6,156	5,497	5,075	4,779	4,559	4,389	4,254	4,142
18	10,218	7,215	6,028	5,375	4,956	4,663	4,445	4,276	4,141	4,030
19	10,073	7,093	5,916	5,268	4,853	4,561	4,345	4,177	4,043	3,933
20	9,944	6,987	5,818	5,174	4,762	4,472	4,257	4,090	3,956	3,847
21	9,829	6,891	5,730	5,091	4,681	4,393	4,179	4,013	3,880	3,771
22	9,727	6,806	5,652	5,017	4,609	4,322	4,109	3,944	3,812	3,703
23	9,635	6,730	5,582	4,950	4,544	4,259	4,047	3,882	3,750	3,642
24	9,551	6,661	5,519	4,890	4,486	4,202	3,991	3,826	3,695	3,587
25	9,475	6,598	5,462	4,835	4,433	4,150	3,939	3,776	3,645	3,537
26	9,406	6,541	5,409	4,785	4,384	4,103	3,893	3,730	3,599	3,492
27	9,342	6,489	5,361	4,740	4,340	4,059	3,850	3,687	3,557	3,450
28	9,284	6,440	5,317	4,698	4,300	4,020	3,811	3,649	3,519	3,412
29	9,230	6,396	5,276	4,659	4,262	3,983	3,775	3,613	3,483	3,376
30	9,180	6,355	5,239	4,623	4,228	3,949	3,742	3,580	3,451	3,344
35	8,976	6,188	5,086	4,479	4,088	3,812	3,607	3,447	3,318	3,212
40	8,828	6,066	4,976	4,374	3,986	3,713	3,509	3,350	3,222	3,117
45	8,715	5,974	4,892	4,294	3,909	3,638	3,435	3,276	3,149	3,044
50	8,626	5,902	4,826	4,232	3,849	3,579	3,376	3,219	3,092	2,988
55	8,554	5,843	4,773	4,181	3,800	3,531	3,330	3,173	3,046	2,942
60	8,495	5,795	4,729	4,140	3,760	3,492	3,291	3,134	3,008	2,904
70	8,403	5,720	4,661	4,076	3,698	3,431	3,232	3,076	2,950	2,846
80	8,335	5,665	4,611	4,028	3,652	3,387	3,188	3,032	2,907	2,803
90	8,282	5,623	4,573	3,992	3,617	3,352	3,154	2,999	2,873	2,770
100	8,241	5,589	4,542	3,963	3,589	3,325	3,127	2,972	2,847	2,744
120	8,179	5,539	4,497	3,921	3,548	3,285	3,087	2,933	2,808	2,705
150	8,118	5,490	4,453	3,878	3,508	3,245	3,048	2,894	2,770	2,667
250	8,021	5,412	4,382	3,812	3,444	3,183	2,987	2,833	2,709	2,607
500	7,950	5,355	4,330	3,763	3,396	3,137	2,941	2,789	2,665	2,562
∞	7,879	5,298	4,279	3,715	3,350	3,091	2,897	2,744	2,621	2,519

**T4 Kvantily F_P Fisherova – Snedecorova rozdělení $F(k_1, k_2)$ pro $P = 0,995$
(pokračování)**

$k_1 \backslash k_2$	12	15	20	24	30	40	60	100	250	∞
1	24426,7	24631,6	24836,5	24937,1	25041,4	25145,7	25253,7	25339,4	25413,9	25466,1
2	199,419	199,434	199,449	199,449	199,478	199,478	199,478	199,478	199,507	199,507
3	43,387	43,085	42,779	42,623	42,466	42,310	42,150	42,022	41,906	41,829
4	20,705	20,438	20,167	20,030	19,892	19,751	19,611	19,497	19,394	19,325
5	13,385	13,146	12,903	12,780	12,656	12,530	12,402	12,300	12,206	12,144
6	10,034	9,814	9,589	9,474	9,358	9,241	9,122	9,026	8,938	8,879
7	8,176	7,968	7,754	7,645	7,534	7,422	7,309	7,217	7,132	7,076
8	7,015	6,814	6,608	6,503	6,396	6,288	6,177	6,087	6,006	5,951
9	6,227	6,032	5,832	5,729	5,625	5,519	5,410	5,322	5,242	5,188
10	5,661	5,471	5,274	5,173	5,071	4,966	4,859	4,772	4,692	4,639
11	5,236	5,049	4,855	4,756	4,654	4,551	4,445	4,359	4,279	4,226
12	4,906	4,721	4,530	4,431	4,331	4,228	4,123	4,037	3,958	3,904
13	4,643	4,460	4,270	4,173	4,073	3,970	3,866	3,780	3,700	3,647
14	4,428	4,247	4,059	3,961	3,862	3,760	3,655	3,569	3,490	3,436
15	4,250	4,070	3,883	3,786	3,687	3,585	3,480	3,394	3,314	3,260
16	4,099	3,920	3,734	3,638	3,539	3,437	3,332	3,246	3,166	3,111
17	3,971	3,793	3,607	3,511	3,412	3,311	3,206	3,119	3,039	2,984
18	3,860	3,683	3,498	3,402	3,303	3,201	3,096	3,009	2,929	2,873
19	3,763	3,587	3,402	3,306	3,208	3,106	3,000	2,913	2,832	2,776
20	3,678	3,502	3,318	3,222	3,123	3,022	2,916	2,828	2,747	2,690
21	3,602	3,427	3,243	3,147	3,049	2,947	2,841	2,753	2,671	2,614
22	3,535	3,360	3,176	3,081	2,982	2,880	2,774	2,685	2,602	2,546
23	3,474	3,300	3,116	3,021	2,922	2,820	2,713	2,624	2,541	2,484
24	3,420	3,246	3,062	2,967	2,868	2,765	2,658	2,569	2,486	2,428
25	3,370	3,196	3,013	2,918	2,819	2,716	2,609	2,519	2,435	2,377
26	3,325	3,151	2,968	2,873	2,774	2,671	2,563	2,473	2,389	2,330
27	3,284	3,110	2,927	2,832	2,733	2,630	2,522	2,431	2,346	2,287
28	3,246	3,073	2,890	2,794	2,695	2,592	2,483	2,392	2,307	2,247
29	3,211	3,038	2,855	2,759	2,660	2,557	2,448	2,357	2,270	2,210
30	3,179	3,006	2,823	2,727	2,628	2,524	2,415	2,323	2,237	2,176
35	3,048	2,876	2,693	2,597	2,497	2,392	2,282	2,188	2,099	2,036
40	2,953	2,781	2,598	2,502	2,401	2,296	2,184	2,088	1,997	1,932
45	2,881	2,709	2,527	2,430	2,329	2,222	2,109	2,012	1,918	1,851
50	2,825	2,653	2,470	2,373	2,272	2,164	2,050	1,951	1,855	1,786
55	2,779	2,608	2,425	2,327	2,226	2,118	2,002	1,902	1,804	1,733
60	2,742	2,570	2,387	2,290	2,187	2,079	1,962	1,861	1,761	1,689
70	2,684	2,513	2,329	2,231	2,128	2,019	1,900	1,797	1,694	1,618
80	2,641	2,470	2,286	2,188	2,084	1,974	1,854	1,748	1,643	1,563
90	2,608	2,437	2,253	2,155	2,051	1,939	1,818	1,711	1,602	1,520
100	2,583	2,411	2,227	2,128	2,024	1,912	1,790	1,681	1,570	1,485
120	2,544	2,373	2,188	2,089	1,984	1,871	1,747	1,636	1,521	1,431
150	2,506	2,335	2,150	2,050	1,944	1,830	1,704	1,590	1,471	1,374
250	2,446	2,275	2,089	1,989	1,882	1,765	1,636	1,516	1,387	1,274
500	2,402	2,230	2,044	1,943	1,835	1,717	1,584	1,460	1,319	1,184
∞	2,358	2,187	2,000	1,898	1,789	1,669	1,533	1,402	1,245	1,000