UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY



TESTOVANIE PRESNOSTI DEA MODELOV

DIPLOMOVÁ PRÁCA

2024 Bc. Bernadett BERTÓKOVÁ

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

TESTOVANIE PRESNOSTI DEA MODELOV

DIPLOMOVÁ PRÁCA

Študijný program:	Ekonomicko-finančná matematika a modelovanie
Študijný odbor:	Matematika
Školiace pracovisko:	Katedra aplikovanej matematiky a štatistiky
Vedúci práce:	doc. RNDr. Mária Trnovská, PhD.

Bratislava 2024 **Bc. Bernadett BERTÓKOVÁ**





Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Bc. Bernadett Bertóková
Študijný program: ekonomicko-finančná matematika a modelovanie (Jednoodborové štúdium, magisterský II. st., denná forma)
Študijný odbor: matematika
Typ záverečnej práce: diplomová
Jazyk záverečnej práce: slovenský
Sekundárny jazyk: anglický

Názov: Testovanie presnosti DEA modelov
Testing the accuracy of DEA models

Anotácia: Klasické metódy obálkovej analýzy dát (DEA) vyhodnocujú efektivitu alebo pseudoefektivitu jednotlivých útvarov, prípadne vedia usporiadať útvary podľa skóre. Rôzne modely však dávajú rôzne hodnoty efektivity a rôzne usporiadania. Testovanie presnosti DEA modelov má preto veľký význam a výsledky testovania môžu poskytnúť užitočnú informáciu o jednotlivých modeloch.

Cieľ: Cieľom diplomovej práce je naštudovať a spracovať problematiku testovania presnosti DEA modelov, aplikovať štatistické prístupy na porovnanie DEA modelov, pričom sa zvolia vhodné funkcie na modelovanie produkčného procesu a pomocou nich sa vygenerujú experimentálne dáta, na ktoré sa aplikujú jednotlivé DEA modely. Presnosť jednotlivých modelov bude skúmaná a vyhodnotená najmä z hľadiska identifikácie efektívnosti a z hľadiska usporiadania jednotlivých útvarov.

Vedúci: doc. RNDr. Mária Trnovská, PhD.
Katedra: FMFI.KAMŠ - Katedra aplikovanej matematiky a štatistiky
Vedúci katedry: prof. RNDr. Marek Fila, DrSc.

Dátum zadania: 10.01.2023

Dátum schválenia: 11.01.2023 prof. RNDr. Daniel Ševčovič, DrSc.
garant študijného programu

.....
študent

.....
vedúci práce

Poďakovanie Touto cestou by som sa chcela poďakovať svojej vedúcej diplomovej práce, doc. RNDr. Márii Trnovskej, PhD. za jej čas, ochotu, odborné rady a užitočné pripomienky, ktoré mi pomohli pri písaní tejto práce. Ďakujem aj svojej rodine a priateľom za podporu počas celého štúdia.

Abstrakt v štátnom jazyku

BERTÓKOVÁ, Bernadett: Testovanie presnosti DEA modelov [Diplomová práca], Univerzita Komenského v Bratislave, Fakulta matematiky, fyziky a informatiky, Katedra aplikovanej matematiky a štatistiky; školiteľ: doc. RNDr. Mária Trnovská, PhD., Bratislava, 2024, 61 s.

Obálková analýza dát vyhodnocuje relatívnu efektívnosť tzv. rozhodovacích útvarov. Aplikovaním DEA modelu na vybranú skupinu útvarov vieme vypočítať hodnoty efektivít, a na základe nich usporiadať jednotlivé útvary do poradia, identifikovať efektívne útvary, a pre neefektívne navrhnúť efektívne vzory. Lenže rôzne DEA modely dávajú rôzne hodnoty efektivity, a teda rôzne usporiadania útvarov. Testovanie presnosti DEA modelov môže uľahčiť výber vhodného DEA modelu na vykonanie DEA analýzy. Avšak na testovanie presnosti DEA modelov nemôžeme použiť dáta o útvaroch z reálneho sveta, lebo nemáme žiadnu informáciu ohľadom ich skutočnej efektivity, nedokázali by sme vyhodnotiť presnosť skúmaných modelov. Preto na testovanie presnosti DEA modelov použijeme skupinu fiktívnych útvarov, ktorých údaje budeme generovať. Cieľom našej diplomovej práce je naštudovať problematiku testovania presnosti DEA modelov, kriticky zhodnotiť existujúce prístupy a navrhnúť vhodnejšiu metodológiu, ktorá umožní korektnejšie porovnávanie DEA modelov. Keďže existuje široké spektrum DEA modelov, sústredíme sa len na jednu triedu, a to modely založené na parametrickej ceste. Najprv sa pozrieme na základy teórie mikroekonómie a predstavíme translogaritmickú funkciu vybranú na modelovanie produkčného procesu. Potom uvedieme základy DEA modelovania, predstavíme vybranú triedu DEA modelov a schému testovania presnosti. Pokračujeme popisom procesu generovania dát. Následne predstavíme štatistické ukazovatele, ktoré použijeme na vyhodnotenie presnosti DEA modelov, a spomenieme dôležitosť využitia Monte Carlo simulácii v kontexte našej problematiky. Nakoniec popíšeme vykonané experimenty, a zhrnieme získané výsledky ohľadom testovania presnosti DEA modelov. Jednotlivé experimenty realizujeme pomocou nami napísaných programov v softvéri MATLAB.

Kľúčové slová: Testovanie presnosti DEA modelov, translogaritmická funkcia, proces generovania dát, CCR-O model, HDF model, DDM model

Abstract

BERTÓKOVÁ, Bernadett: Testing the accuracy of DEA models [Master Thesis], Comenius University in Bratislava, Faculty of Mathematics, Physics and Informatics, Department of Applied Mathematics and Statistics; Supervisor: doc. RNDr. Mária Trnovská, PhD., Bratislava, 2024, 61 p.

Data Envelopment Analysis evaluates the relative efficiency of the so-called decision-making units. By applying a DEA model to a group of decision-making units, we can calculate efficiency scores and based on them rank the decision-making units, identify the efficient ones, and suggest efficient benchmarks for the inefficient ones. Different DEA models give different efficiency scores for decision-making units, and therefore different rankings. Testing the accuracy of DEA models facilitates the selection of the suitable DEA model for DEA analysis. However, for testing the accuracy of DEA models we cannot use the data of real decision-making units, because we don't have any information about their true efficiency scores, we wouldn't be able to estimate the accuracy of the examined models. Therefore, we used an artificial data set to test the accuracy of DEA models. Our Master´s thesis aimed to study the issue of testing the accuracy of DEA models by critically evaluating existing approaches and proposes a more suitable methodology that allows a more correct comparison of DEA models. Since there is a wide range of DEA models, we will focus on just one class, namely the path-based DEA models. First, we state the basics of microeconomics, then we introduce the translog function which we used to model the production process. Consequently, we introduce the basics of DEA and the selected group of DEA models, then the scheme of accuracy testing. We continue with describing the data generation process. After that, we present the performance indicators we applied for evaluating the quality of DEA models and mention the importance of applying Monte Carlo simulations in the context of our issue. Finally, we describe the experiments and summarize the results of testing the accuracy of DEA models. We carry out individual experiments using scripts programmed by ourselves in MATLAB software.

Keywords: Testing the accuracy of DEA models, translog function, data generation process, CCR-O model, HDF model, DDM model

Obsah

Zoznam obrázkov **8**

Úvod **9**

1 Teória mikroekonómie a obálková analýza dát **12**

1.1 Základy teórie mikroekonómie 12

1.1.1 Cobbova-Douglasova produkčná funkcia 14

1.1.2 Translogaritmická funkcia 15

1.2 Základné pojmy a symbolika v DEA 23

1.3 DEA modely založené na parametrickej ceste 26

1.3.1 Vlastnosti vybraných DEA modelov 30

1.4 Testovanie presnosti DEA modelov 31

2 Proces generovania dát **33**

2.1 Generovanie vstupov 34

2.2 Výpočet ideálneho výstupu 36

2.3 Generovanie skutočnej efektivity 38

2.4 Transformácia ideálnych útvarov na fiktívne útvary 40

3 Vyhodnotenie presnosti DEA modelov **45**

3.1 Štatistické ukazovatele 45

3.2 Monte Carlo simulácie 47

4 Výsledky vykonaných experimentov **50**

Záver **57**

Zoznam použitej literatúry **60**

Zoznam obrázkov

1.1 Translogaritmická funkcia s parametrami $\alpha_1 = 0.75, \beta_{11} = -\frac{1}{8}$ 16
1.2 Translogaritmická funkcia s parametrami $\alpha_1 = 1.75, \beta_{11} = -1$ 17
1.3 Translogaritmická funkcia s parametrami $\alpha_1 = 0.1, \beta_{11} = 0.65$ 17
1.4 M_{CRS}: Aproximácia množiny T podľa teórie DEA 25
1.5 M_{VRS}: Aproximácia množiny T podľa teórie DEA 25
1.6 Znázornenie tvaru parametrickej cesty v prípade vybraných DEA modelov 29
1.7 Schéma testovania presnosti DEA modelov 32
2.1 Schéma procesu generovania dát 33
2.2 Funkcia hustoty pravdepodobnosti pre $Beta(8, 2.5)$ 40
2.3 Transformácia ideálnych útvarov na fiktívne útvary - CCR-O model 42
2.4 Transformácia ideálnych útvarov na fiktívne útvary - HDF model 43
2.5 Transformácia ideálnych útvarov na fiktívne útvary - DDM model 43
3.1 Ukážka vývoja smerodajných odchýlok v Monte Carlo simulácii 48
4.1 Presnosť vybraných DEA modelov 51
4.2 Zmena v presnosti modelov (Efektívne útvary) 53
4.3 Zmena v presnosti modelov (Vstupy) 53
4.4 Zmena v presnosti modelov (Korelácia medzi vstupmi) 54
4.5 Presnosť vybraných DEA modelov (štandardný prístup) 55

Úvod

S cieľom odhaliť neefektívnych producentov na trhu a šetriť využívané zdroje je vyhodnocovanie efektívnosti jednotlivých firiem ústrednou súčasťou ich rozhodovacieho manažmentu. Obálková analýza dát je oblasť operačného výskumu založená na riešení optimalizačných úloh, pomocou ktorej dokážeme vyhodnotiť spomenutú efektívnosť. Označujeme ju skratkou DEA podľa anglického názvu *Data Envelopment Analysis*.

Obálková analýza dát sa zaoberá vyhodnocovaním relatívnej efektívnosti jednotlivých útvarov v rámci sledovanej skupiny útvarov. Do jednej skupiny zaraďujeme útvary vykonávajúce tú istú činnosť, ktoré môžu byť napríklad vysoké školy, elektrárne alebo pobočky jednej poisťovne. DEA modely sú definované rôznymi úlohami matematického programovania, ktoré dokážeme efektívne riešiť využitím optimalizačných softvérov. Aplikovaním vybraného DEA modelu na skupinu útvarov vieme vypočítať hodnotu efektivity, na základe nej usporiadať jednotlivé útvary do poradia a určiť efektívne útvary. Ďalej, pre neefektívne útvary vieme určiť efektívne vzory, čiže poskytnúť informáciu o tom, ako je možné znížiť ich mieru neefektívnosti [7].

Začiatky teórie DEA sú spojené definovaním prvého DEA modelu pre konštantné výnosy z rozsahu, ktorým je model CCR predstavený Charnesom, Cooperom a Rhodesom v roku 1978. Analógiou CCR modelu pre variabilné výnosy z rozsahu je model BCC definovaný v roku 1984 a nazvaný podľa iniciálok trojice autorov Banker-Charnes-Cooper. Tieto dva základné DEA modely patria do skupiny orientovaných modelov, majú vstupne a výstupne orientované verzie, teda pri výpočte hodnoty efektivity sa sústredia buď na znižovanie množstiev vstupov, alebo na zvyšovanie množstiev výstupov jednotlivých útvarov. Ich vlastnosti sú dôkladne popísané v učebnom texte [7]. V roku 1985, kombináciou vstupne a výstupne orientovaného modelu vznikol neorientovaný model s názvom HDF, ktorý umožňuje simultánne znižovanie hodnôt vstupov a zvyšovanie hodnôt výstupov po hyperbolickej ceste [8]. Ďalším podobným modelom je DDM model, ktorý znižuje hodnoty vstupov a zvyšuje hodnoty výstupov vo vopred stanovenom smere, vyjadreného pomocou smerového vektora [3]. Ale do skupiny neorientovaných modelov patrí aj Aditívny model, predstavený v roku 1985, ku ktorej neskoršie odvodili aj jej vylepšené verzie s váhami, ako napríklad RAM model, BAM model a ďalšie [5]. V roku 2001 Tone predstavil myšlienku merania efektivity založeného na doplnkových premenných, a definoval nový

neorientovaný model s názvom SBM model [15].

Okrem uvedených modelov v súčasnej literatúre už existuje veľa rôznych DEA modelov, ktorých môžeme využiť na určenie hodnoty efektivity jednotlivých útvarov, čiže na DEA analýzu. Lenže rôzne DEA modely dávajú rôzne hodnoty efektivity pre jednotlivé útvary, a tiež rôzne výsledné usporiadania útvarov. Preto výber vhodného DEA modelu na vykonanie DEA analýzy sa stáva čoraz komplikovanejšou úlohou. Riešením tejto problematiky môže byť testovanie presnosti DEA modelov, ktorému sa vedci začali venovať približne v posledných dvadsiatich rokoch tohto storočia. Dve najrozsiahlejšie štúdie o testovaní presnosti DEA modelov sú opísané vo vedeckých článkoch [10, 19].

Testovanie presnosti DEA modelov má veľký význam, nielenže môže uľahčiť výber vhodného DEA modelu na vykonanie DEA analýzy, ale navyše môže poskytnúť aj užitočné informácie o skúmaných modeloch. Avšak na testovanie presnosti DEA modelov nemôžeme použiť dáta, čiže vstupy a výstupy vybraných útvarov z reálneho sveta, lebo nemáme žiadnu informáciu ohľadom ich skutočnej efektivity. Nevedeli by sme s akou hodnotou porovnať hodnoty efektivít získané aplikovaním jednotlivých DEA modelov na danú skupinu útvarov, teda nedokázali by sme vyhodnotiť presnosť skúmaných modelov. Práve preto, na testovanie presnosti DEA modelov použijeme skupinu fiktívnych útvarov, čiže dáta vstupujúce do experimentov budeme generovať.

Na to, aby sme vedeli správne generovať dáta, najprv potrebujeme vhodne zvoliť funkciu na modelovanie produkčného procesu. Pritom nemôžeme zabudnúť na to, že na získanie všeobecne platných výsledkov o presnosti skúmaných DEA modelov vybranou funkciou potrebujeme vedieť modelovať dostatočne rôznorodé produkčné procesy. Podľa článkov [10, 19] vhodnou voľbou, hlavne kvôli jej flexibilnosti, môže byť takzvaná translogaritmická funkcia. Treba ešte dodať, že kvôli procesu generovania dát vystupuje v problematike testovania presnosti DEA modelov aj náhodnosť. Preto vierohodnú informáciu o presnosti jednotlivých modelov získame až po aplikovaní Monte Carlo metódy na jednotlivé experimenty. Na porovnanie skúmaných DEA modelov použijeme vhodné štatistické prístupy prebrané z literatúry [10, 19], a presnosť vybraných modelov budeme skúmať najmä z hľadiska identifikácie efektívnosti, a z hľadiska usporiadania útvarov.

Cieľom našej diplomovej práce je naštudovať a spracovať problematiku testovania presnosti DEA modelov, pričom sa sústredíme na triedu DEA modelov založených na para-

metrickej ceste. Vychádzať budeme z už existujúcich prístupov k testovaniu presnosti DEA modelov, ktoré boli predstavené v článkoch [10, 19]. Následne navrhneme úpravy v procese generovania dát, ako aj pre samotnú metódu testovania presnosti DEA modelov. Cieľom je aj zdôvodnenie toho, prečo umožňuje navrhnutá nová metodológia numericky efektívnejšie a korektnejšie porovnanie DEA modelov založených na parametrickej ceste, ako doteraz používané prístupy.

Naša práca je rozdelená do štyroch kapitol. V prvej kapitole začíname so základmi teórie mikroekonómie, a predstavíme Cobbovu-Douglasovu produkčnú funkciu [2, 4]. Následne predstavíme translogaritmickú funkciu, ktorou budeme modelovať technológie s rôznymi vlastnosťami. Ukážeme, že pri vhodnom nastavení hodnôt jej parametrov má isté želané vlastnosti charakteristické pre produkčné funkcie [10, 16, 17]. Potom zavedieme v teórii obálkovej analýzy dát používanú symboliku a vysvetlíme niektoré základné pojmy [7]. Následne popíšeme tú triedu DEA modelov, ktorých presnosť sme sa rozhodli testovať. Vybranú triedu tvoria DEA modely založené na parametrickej ceste predstavené v článku [8]. Nakoniec predstavíme schému testovania presnosti DEA modelov [10, 19]. V druhej kapitole sa zaoberáme procesom generovania dát, začneme predstavením postupu generovania množstiev vstupov pre jednotlivé útvary [10, 18]. Potom popíšeme, ako je možné vypočítať hodnotu tzv. ideálneho výstupu pre jednotlivé útvary. V tretej podkapitole sa zaoberáme postupom generovania tzv. hodnoty skutočnej efektivity pre fiktívne útvary [10]. Túto hodnotu budeme následne používať pri vyhodnotení presnosti DEA modelov na porovnanie s hodnotou efektivity získanej z DEA analýzy. Nakoniec predstavíme myšlienku transformácie údajov, ktorej výsledkom je skupina fiktívnych útvarov vstupujúca do DEA analýzy v experimentoch. V tretej kapitole vymenujeme tie štatistické ukazovatele, ktoré budeme používať na vyhodnotenie presnosti vybraných DEA modelov a poukážeme na dôležitosť použitia Monte Carlo metódy v prípade našej problematiky [10]. V poslednej kapitole opíšeme vykonané experimenty a sústredíme sa na získané výsledky ohľadom testovania presnosti DEA modelov. Experimenty realizujeme v softvéri MATLAB, v ktorom naprogramujeme proces generovania dát, vybrané DEA modely a taktiež vyhodnotenie ich presnosti [20].

1 Teória mikroekonómie a obálková analýza dát

V tejto kapitole sa zaoberáme základmi teórie mikroekonómie, pričom vychádzame z učebného textu [2] a z článku [4]. Definujeme translogaritmickú funkciu a uvedieme vhodnú voľbu jej parametrov [10, 16, 17]. Predstavíme základné pojmy obálkovej analýzy dát a zavedieme symboliku [7]. Ďalej predstavíme triedu DEA modelov založených na parametrickej ceste [8]. Nakoniec poukážeme na zaužívaný prístup k testovaniu presnosti DEA modelov [10].

1.1 Základy teórie mikroekonómie

Predpokladajme, že využitím technológie, ktorou sa riadi vybraná skupina n útvarov spotrebovaním m vstupov je možné produkovať jeden výstup. Množstvá jednotlivých vstupov x_i uložíme do vektora $x = (x_1, x_2, \ldots, x_m)^T \in \mathbb{R}^m_+$. Potom danú technológiu je možné charakterizovať technologickou množinou $T \subset \mathbb{R}^{m+1}_+$, ktorá je definovaná nasledovne:

$$T = \left\{ (x, y) \in \mathbb{R}^{m+1}_+ \;\middle|\; x \text{ produkuje } y \right\}. \tag{1}$$

To znamená, že do technologickej množiny patria také dvojice (x, y) pre ktoré platí, že dané množstvo výstupu $y \in \mathbb{R}_+$ sa dá vyrobiť z množstiev vstupov x. Podľa [2] technologická množina T musí spĺňať nasledovné axiómy:

1. $(x, 0) \in T \quad \forall\, x \in \mathbb{R}^m_+$,
2. $(0_m, y) \in T \implies y = 0$,
3. $(x, y) \in T \;\wedge\; \tilde{x} \geq x \implies (\tilde{x}, y) \in T$,
4. T je konvexná a uzavretá množina.

Vymenované axiómy poukazujú na základné vlastnosti vo svete technológii, napríklad na to, že nulové množstvo výstupu y je možné získať akoukoľvek kombináciou množstiev jednotlivých vstupov x_i, ako aj na to, že s nulovými množstvami z každého vstupu $x = 0_m$ sa nedá vyrobiť kladné množstvo výstupu y. Ďalej predpokladajú, že zvýšením množstva niektorých vstupov x_i sa stále dá vyrobiť také množstvo výstupu y, aké sa dalo pred navýšením hodnôt vybraných vstupov x_i. Podľa poslednej axiómy technologická množina T

má byť konvexná a uzavretá. Konvexnosť množiny T môžeme ľahko zdôvodniť na príklade dvoch producentov A a B, ktorí produkujú v danej technológii. To znamená, že za jednotku času t producent A dokáže premieňať vstup x_A na výstup y_A, ako aj producent B vstup x_B na výstup y_B. Ak budeme za čas $\lambda t, \lambda \in [0,1]$, produkovať toľko, ako producent A, a za zvyšný čas $(1-\lambda)t$ toľko, ako producent B, premeníme vstup $\lambda x_A + (1-\lambda)x_B$ na výstup $\lambda y_A + (1-\lambda)y_B$. Keďže (x_A, y_A), ako aj (x_B, y_B) sa nachádzajú v množine T, potom $\forall\, \lambda \in [0,1]$ bude aj dvojica $(\lambda x_A + (1-\lambda)x_B, \lambda y_A + (1-\lambda)y_B)$ súčasťou množiny T, čo implikuje konvexnosť technologickej množiny T. Uzavretosť množiny T je čisto matematická záležitosť [2].

Produkčnou funkciou môžeme vyjadriť optimálny vzťah medzi vektorom vstupov x a výstupom y vo vybranej technológii, ktorou sa riadia sledovaní producenti. Dokážeme ňou modelovať rôzne výrobné procesy. Produkčnú funkciu $p(x) : \mathbb{R}_+^m \to \mathbb{R}_+$ podľa [2] definujeme nasledovne:

$$p(x) = \max\{y \mid (x,y) \in T\}. \tag{2}$$

Hodnota $p(x)$ teda vyjadruje najväčšie množstvo výstupu y, ktoré je možné vyrobiť zo vstupov x v danej technológii. Základné vlastnosti produkčnej funkcie podľa článkov [2, 4] sú nasledovné:

1. $0 \leq p(x) < \infty$,
2. $x = 0_m \implies p(x) = 0$,
3. $p(x)$ je neklesajúca v $x_i\ \forall i = 1, \ldots, m$,
4. $p(x)$ je konkávna.

Podľa základnej vlastnosti 1. produkčná funkcia $p(x)$ môže nadobúdať iba kladné hodnoty, menšie ako nekonečno. Vlastnosť 2. vyjadruje, že z vektora vstupov $x = 0_m$ nedokážeme vyrobiť nič, teda automaticky $p(x) = 0$. Neklesajúcosť v x_i znamená, že ak zvýšime hodnotu niektorého vstupu, tak hodnota $p(x)$ by nemala klesnúť. Matematicky vyjadrené, ak $\tilde{x} \geq x$, tak potom platí, že $p(\tilde{x}) \geq p(x)$. Konkávnosť produkčnej funkcie $p(x)$ vyjadruje, že postupným zvyšovaním hodnôt vstupov, by mala hodnota $p(x)$ narastať čoraz pomalšie a vyplýva z konvexnosti technologickej množiny T.

Uveďme ešte zopár potrebných vzťahov v súvislosti s produkčnou funkciou. Jedna dobrá vlastnosť produkčnej funkcie je jej homogénnosť stupňa k, ktorú definujeme nasledovne:

$$p(tx) = t^k p(x). \tag{3}$$

Hraničné produkty MP_i produkčnej funkcie (2) podľa teórie mikroekonómie definujeme spôsobom:

$$\mathrm{MP}_i = \frac{\partial p(x)}{\partial x_i},\ \forall i = 1, \ldots, m. \tag{4}$$

Pojem hraničný produkt teda zahŕňa v sebe informáciu o tom, v akej miere sa zmení hodnota $p(x)$, ak v malej miere zmeníme množstvo vstupu x_i. Keďže od dobre definovanej produkčnej funkcie očakávame, že pri zvýšení množstva vstupu x_i by hodnota $p(x)$ nemala klesnúť, hraničné produkty musia byť nezáporné, t.j. platí vzťah $\mathrm{MP}_i \geq 0\ \forall i$.

Elasticitu produkčnej funkcie vzhľadom na vstup x_i označujeme symbolom ε_i a definujeme vzťahom:

$$\varepsilon_i = \frac{\partial p(x)}{\partial x_i} \frac{x_i}{p(x)},\ \forall i = 1, \ldots, m, \tag{5}$$

čím vlastne vyjadrujeme hraničnú zmenu v hodnote $p(x)$ vzhľadom k hraničnej zmene v hodnote vstupu x_i. Keďže o vstupoch x_i sme predpokladali, že sú nezáporné a platí aj nerovnosť $p(x) \geq 0$, potom pre elasticity ε_i produkčnej funkcie bude platiť vzťah $\varepsilon_i \geq 0$. Elasticity produkčnej funkcie sú užitočné aj pri určení výnosov z rozsahu, keďže podľa [4] platí vzťah:

$$\sum_{i=1}^{m} \varepsilon_i = \begin{cases} > 1 & \leftrightarrow \mathrm{IRS} \\ = 1 & \leftrightarrow \mathrm{CRS} \\ < 1 & \leftrightarrow \mathrm{DRS} \end{cases} \tag{6}$$

Čiže ak súčet elasticít produkčnej funkcie $p(x)$ nadobudne hodnotu 1, tak hovoríme, že produkčná funkcia má konštantné výnosy z rozsahu. Ak súčet elasticít je menší/väčší ako číslo 1, v tom prípade má produkčná funkcia klesajúce/rastúce výnosy z rozsahu.

1.1.1 Cobbova-Douglasova produkčná funkcia

V literatúre existuje niekoľko typov produkčnej funkcie, napríklad Cobbova-Douglasova produkčná funkcia, produkčná funkcia s konštantnou elasticitou alebo Leontieffova produkčná funkcia [2]. Autori článku [10] spracovali prehľad o vlastnostiach procesov generovania dát využívajúcich Monte Carlo metódu v štúdiách o DEA, podľa ktorého na

modelovanie produkčného procesu bola zvolená v prvom rade Cobbova-Douglasova produkčná funkcia. Túto funkciu môžeme definovať nasledovným vzťahom:

$$p_{CD}(x) = \alpha_0 \prod_{i=1}^{m} x_i^{\alpha_i},\ \alpha_0 > 0,\ \sum_{i=1}^{m} \alpha_i \leq 1, \tag{7}$$

kde parameter α_0 podľa [12] sa nazýva efektívnosťou technológie, a vyjadruje tú časť nárastu v hodnote $p_{CD}(x)$, ktorá sa nedá vysvetliť rastom jednotlivých vstupov, ale poukazuje na technický pokrok v danom produkčnom procese. Parametre α_i charakterizujú úroveň dôležitosti jednotlivých vstupov x_i. Cobbova-Douglasova produkčná funkcia sa dostala do popredia v ekonómii aj kvôli jej jednoduchosti. Avšak pre testovanie presnosti DEA modelov už táto funkcia nie je vhodná, keďže s ňou nedokážeme modelovať dostatočne rôznorodé produkčné procesy [10]. Aby sme získali vierohodnejšie informácie o presnosti jednotlivých DEA modelov potrebujeme si zvoliť vhodnejšiu produkčnú funkciu.

1.1.2 Translogaritmická funkcia

Na to, aby sme vedeli generovať korektné dáta pre fiktívne útvary, potrebujeme vhodne zvoliť funkciu na modelovanie vybranej technológie. Pritom nemôžeme zabudnúť na to, že na získanie všeobecne platných výsledkov o presnosti skúmaných DEA modelov vybranou produkčnou funkciou potrebujeme vedieť modelovať dostatočne rôznorodé produkčné procesy. Vhodnou voľbou na účel modelovania rozličných produkčných procesov podľa článku [10], kvôli jej flexibilnosti, môže byť takzvaná translogaritmická funkcia, ktorá má nasledovný predpis:

$$p_T(x) = \alpha_0 \prod_{i=1}^{m} x_i^{\alpha_i} \prod_{i=1}^{m} x_i^{\frac{1}{2}\sum_{h=1}^{m} \beta_{ih} \ln x_h}. \tag{8}$$

Táto funkcia umožňuje zaoberať sa aj so substitučnými efektami medzi jednotlivými vstupmi, ktoré vyjadrujú tiež dôležitú vlastnosť technológii. Do vektora $x \in \mathbb{R}^m_{++}$ sú uložené hodnoty vstupov. Parameter $\alpha_0 > 0$ vyjadruje efektívnosť technológie, ktorú pre náš prípad zvolíme ako $\alpha_0 = 1$, aby sme vo väčšej miere korešpondovali s teóriou DEA modelovania. Parametre α_i vyjadrujú dôležitosť jednotlivých vstupov v danom produkčnom procese. Podľa [10] parametre dôležitosti vstupov spĺňajú nerovnosť $\alpha_i > 0$. Ďalej parametre β_{ih} poukazujú na mieru nahraditeľnosti vstupu x_i vstupom x_h. Poznamenajme, že ak by sme zvolili parametre mier nahraditeľnosti spôsobom $\beta_{ih} = 0,\ \forall i, h$ a súčet parametrov dôležitosti vstupov by spĺňal podmienku $\sum_{i=1}^{m} \alpha_i \leq 1$, tak by sme dostali známu Cobbovu-Douglasovu produkčnú funkciu.

Uveďme aj logaritmický tvar funkcie (8), ktorý vyzerá nasledovne:

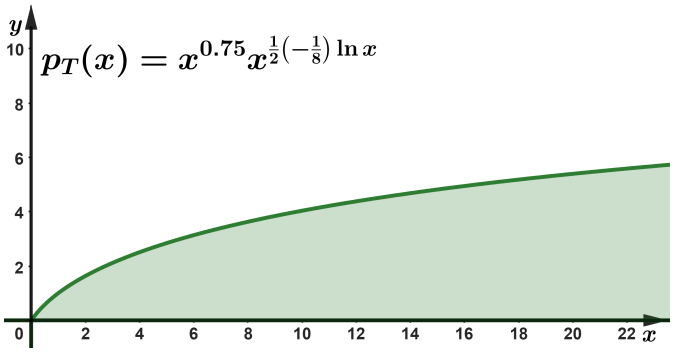
$$\ln p_T(x) = \sum_{i=1}^{m} \alpha_i \ln x_i + \frac{1}{2} \sum_{i=1}^{m} \sum_{h=1}^{m} \beta_{ih} \ln x_i \ln x_h. \tag{9}$$

Podľa [16] logaritmický tvar (9) môžeme previesť do kvadratickej formy:

$$q(z) = \alpha^T z + \frac{1}{2} z^T B z, \tag{10}$$

kde vektor z sa skladá zo zložiek $z_i = z_i(x_i) = \ln x_i$ a vektor α obsahuje parametre dôležitosti α_i jednotlivých vstupov. Symetrickú maticu B rozmeru $m \times m$ tvoria zložky β_{ih} vyjadrujúce mieru nahraditeľnosti vstupov x_i a x_h.

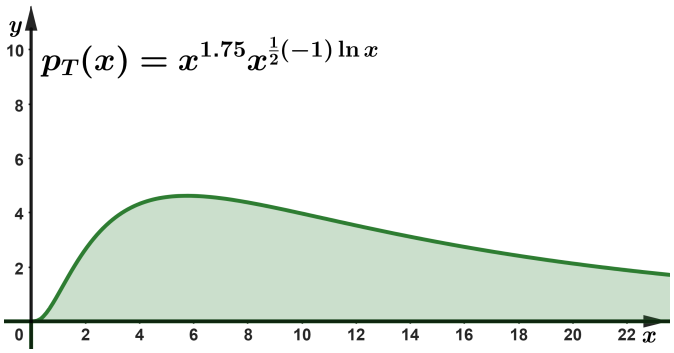
Aby sme funkciou (8) mohli modelovať produkčné procesy, potrebujeme aby spĺňala základné vlastnosti produkčnej funkcie, ktoré boli vymenované v podkapitole 1.1. To môžeme dosiahnuť vhodnou voľbou jej parametrov, ktoré musia spĺňať isté podmienky. Na Obr.1.1 sme znázornili príklad translogaritmickej funkcie $p_T(x) = x^{\alpha_1} x^{\frac{1}{2}\beta_{11} \ln x}$ závislej od jedného vstupu x s parametrami $\alpha_1 = 0.75$ a $\beta_{11} = -\frac{1}{8}$. Zelenou farbou vyznačená časť zodpovedá bodom (x, y), ktoré patria do technologickej množiny T príslušnej translogaritmickej funkcie s uvedenými hodnotami parametrov. Môžeme si všimnúť, že znázornená translogaritmická funkcia spĺňa základné vlastnosti produkčnej funkcie, t.j. má tvar produkčnej funkcie.



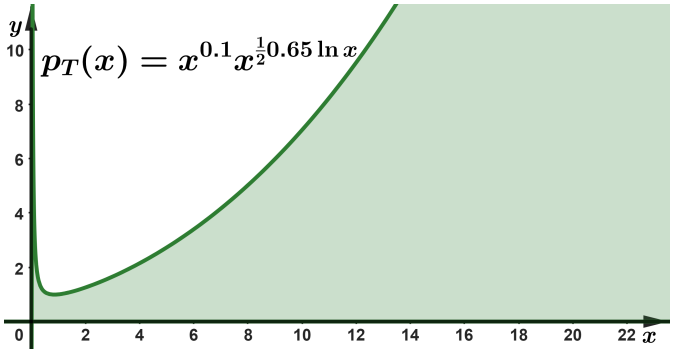
Obr. 1.1: Translogaritmická funkcia s parametrami $\alpha_1 = 0.75, \beta_{11} = -\frac{1}{8}$

Na ilustráciu dôležitosti vhodnej voľby parametrov translogaritmickej funkcie použijeme obrázky 1.2 a 1.3, na ktorých môžeme vidieť príklady translogaritmickej funkcie s nevhodnými nastaveniami jej parametrov, ktoré nespĺňajú niektoré zo základných vlastností produkčnej funkcie. Na Obr.1.2 znázornená translogaritmická funkcia s parametrami $\alpha_1 = 1.75$ a $\beta_{11} = -1$ nespĺňa vlastnosť neklesajúcosti v x_i a konkávnosti. Na Obr.1.3

znázornená translogaritmická funkcia s parametrami $\alpha_1 = 0.1$ a $\beta_{11} = 0.65$ nespĺňa žiadnu vlastnosť z vymenovaných základných vlastností produkčnej funkcie.



Obr. 1.2: Translogaritmická funkcia s parametrami $\alpha_1 = 1.75, \beta_{11} = -1$



Obr. 1.3: Translogaritmická funkcia s parametrami $\alpha_1 = 0.1, \beta_{11} = 0.65$

Predtým, ako ukážeme, že translogaritmická funkcia pri vhodnej voľbe jej parametrov spĺňa základné vlastnosti produkčnej funkcie, trochu upravíme jej predpis na tvar:

$$p_T(x) = \prod_{i=1}^{m} x_i^{\alpha_i} \prod_{i=1}^{m} x_i^{\frac{1}{2}\vartheta \sum_{h=1}^{m} \sigma_{ih} \ln x_h}, \tag{11}$$

v ktorom sme podľa myšlienky z článku [10] parameter β_{ih} nahradili súčinom $\vartheta\sigma_{ih}$, kde parameter $\vartheta > 0$ sa nazýva substitučnou schopnosťou, a parametre σ_{ih} vyjadrujú substitučnú distribúciu jednotlivých vstupov. Zvýšenie počtu parametrov funkcie (8) vedie k možnosti modelovania rozmanitejších produkčných procesov. Potom logaritmický tvar (9) v prípade funkcie (11), ako aj kvadratický tvar (10) môžeme vyjadriť vzťahmi:

$$\ln p_T(x) = \sum_{i=1}^{m} \alpha_i \ln x_i + \frac{1}{2}\vartheta \sum_{i=1}^{m} \sum_{h=1}^{m} \sigma_{ih} \ln x_i \ln x_h, \tag{12}$$

$$q(z) = \alpha^T z + \frac{1}{2}\vartheta z^T \Sigma z, \tag{13}$$

kde symetrická matica Σ obsahuje parametre substitučnej distribúcie σ_{ih}, t.j. $\vartheta\Sigma = B$.

Definujme elasticity ε_i funkcie (11):

$$\varepsilon_i = \varepsilon_i(x) = \frac{\partial \ln p_T(x)}{\partial \ln x_i} = \frac{\partial q}{\partial z_i} = \alpha_i + \vartheta \sum_{h=1}^{m} \sigma_{ih} \ln x_h,\ i = 1, \ldots, m. \tag{14}$$

Potom elasticity $\varepsilon_i(x)$ môžeme zapísať do jedného vektora $\varepsilon(x)$ nasledovne:

$$\varepsilon = \varepsilon(x) = \alpha + \vartheta\Sigma z(x). \tag{15}$$

Hraničné produkty MP_i funkcie (11) vyjadríme pomocou vzťahu $\mathrm{MP}_i = \varepsilon_i \frac{p_T(x)}{x_i}$:

$$MP_i = \varepsilon_i(x)\frac{p_T(x)}{x_i} = \frac{\partial \ln p_T(x)}{\partial \ln x_i}\frac{p_T(x)}{x_i} = \left(\alpha_i + \vartheta \sum_{h=1}^{m} \sigma_{ih} \ln x_h\right)\frac{p_T(x)}{x_i},\ i = 1, \ldots, m. \tag{16}$$

Uvedieme aj vzťah pre výnosy z rozsahu pre funkciu (11) podľa [4]:

$$\sum_{i=1}^{m} \varepsilon_i(x) = \sum_{i=1}^{m} \alpha_i + \vartheta \sum_{i=1}^{m}\sum_{h=1}^{m} \sigma_{ih} \ln x_h = \begin{cases} > 1 & \leftrightarrow \text{IRS} \\ = 1 & \leftrightarrow \text{CRS} \\ < 1 & \leftrightarrow \text{DRS} \end{cases} \tag{17}$$

V tejto práci sa zaoberáme technológiami, ktoré zodpovedajú konštantným výnosom z rozsahu, preto očakávame splnenie podmienky $\sum_{i=1}^{m} \varepsilon_i(x) = 1$ v prípade funkcie $p_T(x)$. Gradient translogaritmickej funkcie $\nabla p_T(x) = (\mathrm{MP}_1, \ldots, \mathrm{MP}_m)$ uvedieme pomocou vektorového zápisu:

$$\nabla p_T(x) = p_T(x)\varepsilon(x) \circ x^{-1},$$

kde $\circ$ označuje takzvaný Hadamardov (pozložkový) súčin a x^{-1} je vektor obsahujúci zložky $(x^{-1})_i = x_i^{-1}$, $\forall i = 1, \ldots, m$. Druhé parciálne derivácie funkcie (11) sú:

$$\frac{\partial^2 p_T(x)}{\partial x_i^2} = \frac{p_T(x)}{x_i^2}\left(\vartheta\sigma_{ii} + \varepsilon_i(x)^2 - \varepsilon_i(x)\right),$$
$$\frac{\partial^2 p_T(x)}{\partial x_i \partial x_h} = \frac{p_T(x)}{x_i x_h}\left(\vartheta\sigma_{ih} + \varepsilon_i(x)\varepsilon_h(x)\right).$$

Zaveďme nasledovné označenia:

$$\mathbb{X} := x^{-1}(x^{-1})^T,\ \text{t.j. } \mathbb{X}_{ih} = \frac{1}{x_i x_h},$$
$$\mathbb{E} := \varepsilon(x)\varepsilon(x)^T,\ \text{t.j. } \mathbb{E}_{ih} = \varepsilon_i(x)\varepsilon_h(x),$$
$$diag(\varepsilon) := \begin{bmatrix} \varepsilon_1(x) & 0 & \ldots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \ldots & 0 & \varepsilon_m(x) \end{bmatrix}_{m\times m},$$

Potom Hessovu maticu $\nabla^2 p_T(x)$ funkcie (11) môžeme vyjadriť v kompaktnom tvare:

$$\nabla^2 p_T(x) = p_T(x)\mathbb{X} \circ (\vartheta\Sigma + \mathbb{E} - diag(\varepsilon)) . \tag{18}$$

Definujme množinu D, do ktorej budú patriť množstvá vstupov jednotlivých útvarov nasledovne:

$$D = \{x \mid 1 \leq x^{\min} \leq x_i \leq x^{\max},\ i = 1, \ldots, m\},$$

kde symbolmi $x^{\min}, x^{\max}$ označujeme minimálnu/maximálnu hodnotu spomedzi množstiev vstupov x_{ij} v skúmanej skupine útvarov. Množinu D budeme považovať za definičný obor funkcie (11). Ukážeme, že pri vhodnej voľbe parametrov α_i, σ_{ih} a ϑ translogaritmická funkcia definovaná predpisom (11) vykazuje konštantné výnosy z rozsahu a spĺňa základné vlastnosti produkčnej funkcie, ktoré boli vymenované v podkapitole 1.1.

Začnime s podmienkami kladenými na parametre α_i, σ_{ih} z dôvodu modelovanie technológii zodpovedajúce konštantným výnosom z rozsahu:

Tvrdenie 1.1. *Ak parametre funkcie $p_T(x)$ spĺňajú podmienky:*

a) $\alpha_i > 0,\ \forall i = 1, \ldots, m,\ \sum_{i=1}^{m} \alpha_i = 1,$

b) $\sigma_{ih} = \sigma_{hi} \geq 0,\ \forall i, h = 1, \ldots, m, i \neq h, \sigma_{hh} = -\sum_{i=1, i\neq h}^{m} \sigma_{ih} < 0,\ \forall h = 1, \ldots, m,$

potom funkcia $p_T(x)$ je homogénna stupňa $k = \sum_{i=1}^{m} \alpha_i$, vykazuje konštantné výnosy z rozsahu, naviac matica Σ je záporne semidefinitná.

Dôkaz. Podľa článku [1] postačujúca podmienka pre homogénnosť translogaritmickej funkcie stupňa $k = \sum_{i=1}^{m} \alpha_i$ je nasledovná:

$$\vartheta \sum_{i=1}^{m} \sigma_{ih} = \vartheta \sum_{h=1}^{m} \sigma_{ih} = \vartheta \sum_{i=1}^{m} \sum_{h=1}^{m} \sigma_{ih} = 0,$$

ktorú po úvahe, že parameter substitučnej schopnosti spĺňa nerovnosť $\vartheta > 0$ môžeme zjednodušiť do tvaru:

$$\sum_{i=1}^{m} \sigma_{ih} = \sum_{h=1}^{m} \sigma_{ih} = \sum_{i=1}^{m} \sum_{h=1}^{m} \sigma_{ih} = 0. \tag{19}$$

Platnosť podmienky (19) vyplýva z voľby parametrov σ_{ih} spĺňajúce podmienky b) z Tvrdenia 1.1. Potom pri podmienke kladenej na súčet parametrov α_i bude translogaritmická funkcia homogénna stupňa $k = \sum_{i=1}^{m} \alpha_i = 1$.

Avšak pri podmienke (19) rovnosť $\sum_{i=1}^{m} \alpha_i = 1$ nedefinuje iba stupeň homogénnosti funkcie (11), ale aj jej výnosy z rozsahu, keďže pri podmienke (19) si vzťah (17) môžeme zjednodušiť na:

$$\sum_{i=1}^{m} \varepsilon_i(x) = \sum_{i=1}^{m} \alpha_i = \begin{cases} > 1 & \leftrightarrow \text{IRS} \\ = 1 & \leftrightarrow \text{CRS} \\ < 1 & \leftrightarrow \text{DRS} \end{cases} \tag{20}$$

Ukázali sme, že pri platnosti podmienok uvedených v Tvrdení 1.1. translogaritmická funkcia vykazuje konštantné výnosy z rozsahu.

Potrebujeme ešte ukázať, že matica Σ je záporne semidefinitná, t.j. jej vlastné hodnoty spĺňajú nerovnosť $\lambda_i \leq 0, i = 1, \ldots, m$. Záporná semidefinitnosť matice Σ vyplýva z Gershgorinovej vety, pomocou ktorej môžeme definovať interval pre vlastné hodnoty matice Σ označeného symbolom $[\lambda^{\min}, \lambda^{\max}]$ nasledovne [6]:

$$\lambda^{\min} = \min_i \left\{ \sigma_{hh} - \sum_{i \neq h} |\sigma_{ih}| \right\},$$

$$\lambda^{\max} = \min_i \left\{ \sigma_{hh} + \sum_{i \neq h} |\sigma_{ih}| \right\},$$

kde $\lambda^{\min}$ vyjadruje najvyššiu dolnú hranicu a $\lambda^{\max}$ najnižšiu hornú hranicu pre vlastné hodnoty λ_i. Potom pri podmienke b) z Tvrdenia 1.1. v našom prípade platia nerovnosti $\lambda^{\min} \leq 0$ a $\lambda^{\max} = 0$, z čoho vyplýva, že matica Σ má nekladné vlastné hodnoty, teda je záporne semidefinitná. □

Pokračujeme s jedným dôsledkom Tvrdenia 1.1. [17]:

Dôsledok 1.2. *Ak matica Σ je záporne semidefinitná, tak pre translogaritmickú funciu platí vzťah: $p_T(x) \leq p_{CD}(x)$.*

Dôkaz. Chceme dokázať, že platí vzťah:

$$p_T(x) \leq p_{CD}(x). \tag{21}$$

Dosadením predpisov funkcie do vzťahu (21) pre $\alpha_0 = 1$ dostaneme nerovnosť:

$$\prod_{i=1}^{m} x_i^{\alpha_i} \prod_{i=1}^{m} x_i^{\frac{1}{2}\vartheta \sum_{h=1}^{m} \sigma_{ih} \ln x_h} \leq \prod_{i=1}^{m} x_i^{\alpha_i}, \tag{22}$$

ktorá platí v prípade platnosti nerovnosti:

$$\prod_{i=1}^{m} x_i^{\frac{1}{2}\vartheta \sum_{h=1}^{m} \sigma_{ih} \ln x_h} \leq 1. \tag{23}$$

Po zlogaritmizovaní nerovnice (23) a označení zložiek vektora z ako $z_h = \ln x_h \, \forall h$, dostaneme podmienku:

$$\frac{1}{2}\vartheta z^T \Sigma z \leq 0, \tag{24}$$

ktorá pre $\vartheta > 0$ platí práve pri podmienke $\Sigma \preceq 0$, čiže ak matica Σ je záporne semidefinitná. □

Teraz využitím Dôsledku 1.2. ukážeme, že funkcia (11) spĺňa základné vlastnosti 1. a 2. produkčnej funkcie vymenované v podkapitole 1.1. Cobbova-Douglasova produkčná funkcia pri podmienke a) z Tvrdenia 1.1. a na množine D, kde $x^{\max} < \infty$ zrejme nadobúda konečné nezáporné hodnoty. Potom z Dôsledku 1.2. vyplýva, že aj translogaritmická funkcia nadobúda konečné nezáporné hodnoty za tie isté podmienky, ako Cobbova-Douglasova produkčná funkcia. Síce do definičného oboru D translogaritmickej funkcie nepatrí bod $x = 0_m$, ale môžeme ukázať, že platí vzťah $\lim_{x\to 0} p_T(x) = 0$ [17]. Platnosť tohto vzťahu vyplýva tiež z Dôsledku 1.2., keďže pre Cobbovu-Douglasovu produkčnú funkciu pri podmienke a) z Tvrdenia 1.1. platí $\lim_{x\to 0} = \lim_{x\to 0} \prod_{i=1}^{m} x_i^{\alpha_i} = 0$, a potom pri platnosti nerovnosti $0 \leq p_T(x) \leq p_{CD}(x)$ platí aj požadovaný vzťah $\lim_{x\to 0} p_T(x) = 0$.

To, že translogaritmická funkcia pri vhodnej voľbe parametra $\vartheta > 0$ spĺňa základnú vlastnosť produkčnej funkcie 3. z podkapitoly 1.1 ukážeme pomocou nasledovného tvrdenia:

Tvrdenie 1.3. *Ak funkcia $p_T(x)$ je definovaná na množine D a parameter substitučnej schopnosti spĺňa podmienku:*

$$0 < \vartheta \leq \min_i \left\{ \frac{\alpha_i}{\ln x^{max} \sum_{h=1}^{m} |\sigma_{ih}|} \right\},$$

potom translogaritmická funkcia je neklesajúca v x_i.

Aby funkcia (11) bola neklesajúca v x_i, jej hraničné produkty MP_i, a keďže vektor x a aj hodnota $p_T(x)$ sú nezáporné, tým pádom aj produkčné elasticity $\varepsilon_i(x)$ musia byť nezáporné. Vychádzajme z nerovností $\varepsilon_i(x) \geq 0$, ktoré platia pre $i = 1, \ldots, m$:

$$\alpha_i + \vartheta \sum_{h=1}^{m} \sigma_{ih} \ln x_h \geq 0, \; \forall i. \tag{25}$$

V pokračovaní budeme upravovať ľavú časť nerovností (25), pričom nemôžeme zabudnúť

na to, že pre $x \in D$ platí $\ln x_h \geq 0$:

$$\begin{aligned}
\alpha_i + \vartheta \sum_{h=1}^{m} \sigma_{ih} \ln x_h &= \alpha_i - \vartheta \sum_{h=1}^{m} -\sigma_{ih} \ln x_h \geq \\
&\geq \alpha_i - \vartheta \sum_{h=1}^{m} |-\sigma_{ih}| \ln x_h = \alpha_i - \vartheta \sum_{h=1}^{m} |\sigma_{ih}| \ln x_h \geq \\
&\geq \alpha_i - \vartheta \sum_{h=1}^{m} |\sigma_{ih}| \ln x^{\max} = \alpha_i - \vartheta \ln x^{\max} \sum_{h=1}^{m} |\sigma_{ih}|.
\end{aligned}$$

Keďže od elasticít očakávame že budú nezáporné, musia platiť nerovnosti:

$$\alpha_i - \vartheta \ln x^{\max} \sum_{h=1}^{m} |\sigma_{ih}| \geq 0, \ \forall i. \tag{26}$$

Ak tieto nerovnosti platia, tak musia platiť aj nerovnosti (25). Úpravou nerovností (26) dostaneme podmienku na parameter $\vartheta > 0$:

$$\begin{aligned}
\alpha_i &\geq \vartheta \ln x^{\max} \sum_{h=1}^{m} |\sigma_{ih}|, \quad \forall i \\
\frac{\alpha_i}{\ln x^{\max} \sum_{h=1}^{m} |\sigma_{ih}|} &\geq \vartheta, \quad \forall i \\
\vartheta &\leq \min_i \left\{ \frac{\alpha_i}{\ln x^{\max} \sum_{h=1}^{m} |\sigma_{ih}|} \right\}.
\end{aligned} \tag{27}$$

Ukázali sme, že podmienka z Tvrdenia 1.3 kladená na parameter substitučnej schopnosti zaručuje, že translogaritmická funkcia bude neklesajúca v x_i.

Nakoniec, konkávnosť translogaritmickej funkcie na množine D môžeme uviesť, ako dôsledok v tejto podkapitole vyslovených tvrdení [17]:

Dôsledok 1.4. *Ak funkcia $p_T(x)$ definovaná na množine D spĺňa podmienky:*

a) $\sigma_{ih} = \sigma_{hi} \geq 0, \ \forall i, h = 1, \ldots, m, i \neq h, \sigma_{hh} = -\sum_{i=1, i\neq h}^{m} \sigma_{ih} < 0, \ \forall h = 1, \ldots, m,$

b) vykazuje konštantné výnosy z rozsahu,

c) $\nabla p_T(x) \geq 0,$

potom platí $\nabla^2 p_T(x) \preceq 0$, teda translogaritmická funkcia je konkávna na množine D.

Dôkaz tohto dôsledku sa nachádza v článku [16]. Ukázali sme, že pri vhodnom nastavení parametrov α_i, σ_{ih} a ϑ translogaritmická funkcia definovaná predpisom (11) spĺňa základné vlastnosti produkčnej funkcie. Z dôvodu lepšej prehľadnosti zapíšeme do nasledovného tvrdenia, ktoré podmienky musia spĺňať jednotlivé parametre, aby sme mohli translogaritmickú funkciu použiť na modelovanie produkčných procesov:

Tvrdenie 1.5. *Ak parametre translogaritmickej funkcie $p_T(x)$ definovanej na množine D spĺňajú podmienky:*

a) $\alpha_i > 0,\ \forall i = 1, \ldots, m,\ \sum_{i=1}^{m} \alpha_i = 1,$

b) $\sigma_{ih} = \sigma_{hi} \geq 0,\ \forall i, h = 1, \ldots, m, i \neq h, \sigma_{hh} = -\sum_{i=1, i\neq h}^{m} \sigma_{ih} < 0,\ \forall h = 1, \ldots, m,$

c) $0 < \vartheta \leq \min_i \left\{ \frac{\alpha_i}{\ln x^{max} \sum_{h=1}^{m} |\sigma_{ih}|} \right\},$

potom translogaritmická funkcia spĺňa základné vlastnosti produkčnej funkcie.

1.2 Základné pojmy a symbolika v DEA

Obálková analýza dát je neparametrická metóda používaná na vyhodnocovanie relatívnej efektívnosti útvarov vykonávajúcich tú istú činnosť. Hodnota efektívnosti vyjadruje účinnosť jednotlivých útvarov, ktorých budeme označovať podľa anglického názvu *Decision Making Unit* skratkou DMU. Aplikovaním vybraného DEA modelu na skupinu útvarov vieme vypočítať hodnotu efektivity, na základe nej usporiadať jednotlivé útvary do poradia a určiť efektívne útvary. Ďalej, pre neefektívne útvary vieme určiť efektívne vzory, čiže poskytnúť informáciu o tom, ako je možné znížiť ich mieru neefektívnosti. Aplikácia DEA modelu na daný problém znamená riešiť sadu úloh matematického programovania, jednu úlohu pre každý jeden útvar zo sledovanej skupiny.

Uvažujme technológiu, ktorou sa riadi vybraná skupina n útvarov, a podľa ktorej spotrebovaním m vstupov je možné produkovať s výstupov. Potom vstupy jednotlivých $DMU_j,\ j = 1, \ldots, n$, zapíšeme do vektora $x_j \in \mathbb{R}^m$ a výstupy do vektora $y_j \in \mathbb{R}^s$ nasledovným spôsobom:

$$x_j = \begin{pmatrix} x_{1j} \\ \vdots \\ x_{ij} \\ \vdots \\ x_{mj} \end{pmatrix}, \quad y_j = \begin{pmatrix} y_{1j} \\ \vdots \\ y_{rj} \\ \vdots \\ y_{sj} \end{pmatrix},\ j = 1, \ldots, n,$$

pričom podľa [7] platia predpoklady $x_j \neq 0_m$ a $y_j \neq 0_s$. V prípade, že sa chceme sústrediť na jeden konkrétny útvar, tak ho budeme označovať symbolom $DMU_o = (x_o, y_o)$. V

niektorých prípadoch zapisujeme vstupy a výstupy pomocou nasledovných matíc:

$$X = \begin{pmatrix} x_{11} & \dots & x_{1j} & \dots & x_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{i1} & \dots & x_{ij} & \dots & x_{in} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{m1} & \dots & x_{mj} & \dots & x_{mn} \end{pmatrix} \in \mathbb{R}_+^{m\times n}, \qquad Y = \begin{pmatrix} y_{11} & \dots & y_{1j} & \dots & y_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ y_{r1} & \dots & y_{rj} & \dots & y_{rn} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ y_{s1} & \dots & y_{sj} & \dots & y_{sn} \end{pmatrix} \in \mathbb{R}_+^{s\times n},$$

kde zložka x_{ij} označuje množstvo i-teho vstupu a zložka y_{rj} množstvo r-teho výstupu DMU_j. V praxi nikdy nepoznáme všetky útvary, ktoré sa riadia vybranou technológiou, preto technologickú množinu T potrebujeme aproximovať pomocou tých útvarov, o ktorých sú dostupné údaje. Aproximáciu množiny T v DEA modelovaní pri konštantných výnosoch z rozsahu označujeme symbolom M_{CRS}. Algebraické vyjadrenie množiny M_{CRS} má nasledovný tvar:

$$M_{CRS} = \left\{ (x, y) \in \mathbb{R}^{m+s} \;\middle|\; X\lambda \leq x,\ Y\lambda \geq y,\ \lambda \geq 0 \right\}.$$

Množina M_{CRS} je podľa [7] uzavretý, konvexný, polyedrický kužeľ, ohraničený v smeroch znižovania vstupov x_i a zvyšovania výstupov y_r. Druhou možnosťou aproximácie množiny T v DEA modelovaní je aproximácia pri variabilných výnosoch z rozsahu označená symbolom M_{VRS}. Jej algebraické vyjadrenie má nasledovný tvar:

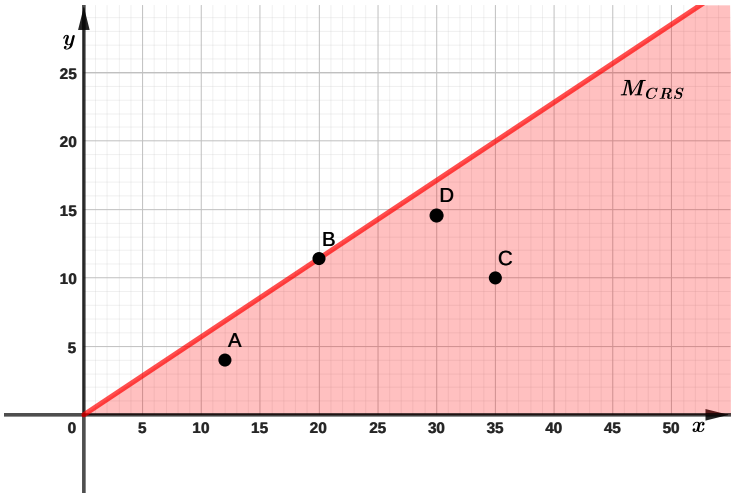
$$M_{VRS} = \left\{ (x, y) \in \mathbb{R}^{m+s} \;\middle|\; X\lambda \leq x,\ Y\lambda \geq y,\ \lambda \geq 0,\ 1_n^T\lambda = 1 \right\},$$

kde symbolom 1_n označujeme vektor jednotiek dĺžky n. Množina M_{VRS} je podľa [7] uzavretá, konvexná, polyedrická množina, ohraničená v smeroch znižovania vstupov x_i a zvyšovania výstupov y_r.

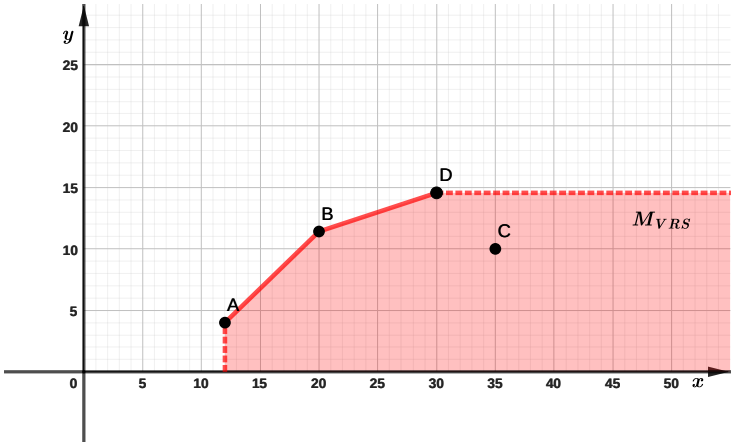
Symbolom θ_j^* budeme označovať efektivity DMU_j získané aplikovaním vybraného DEA modelu na sledovanú skupinu útvarov. Pre neefektívne útvary platí $\theta_j^* \in [0, 1)$, ak sledovaný DMU_j je efektívny, tak $\theta_j^* = 1$. V prípade konštantných výnosov z rozsahu resp. variabilných výnosov z rozsahu je DMU_o efektívny, vtedy a len vtedy, ak neexistuje taká kónická resp. konvexná kombinácia všetkých DMU_j z danej skupiny útvarov, rozdielna od DMU_o, ktorá by mala nižšie alebo rovnaké vstupy a vyššie alebo rovnaké výstupy. Efektívny vzor pre skúmaný neefektívny útvar označíme symbolom $DMU_o^* = (x_o^*, y_o^*)$, a získame ho projekciou DMU_o na hranicu množiny M_{CRS}/M_{VRS}. Pareto-efektívnosť sledovaného útvaru vzhľadom na množinu M definujeme nasledovne:

Definícia 1.6 (**Pareto-efektívnosť**). *$DMU_o = (x_o, y_o) \in M$ nazývame Pareto-efektívnym vzhľadom na množinu M vtedy, ak neexistuje $(x, y) \in M$ také, že $(x, y) \neq (x_o, y_o)$ a zároveň $x_o \geq x$ a $y_o \leq y$.*

Ukážeme na príklade technológie s jedným vstupom x, jediným výstupom y, a skupinou štyroch útvarov A, B, C a D rozdiel medzi aproximáciou technologickej množiny T pri konštantných výnosoch z rozsahu (viď. Obr.1.4) a pri variabilných výnosoch z rozsahu (viď Obr.1.5).



Obr. 1.4: M_{CRS}: Aproximácia množiny T podľa teórie DEA



Obr. 1.5: M_{VRS}: Aproximácia množiny T podľa teórie DEA

Potom podľa definície 1.6 vzhľadom na množinu M_{CRS} je Pareto-efektívny iba útvar B, a vzhľadom na množinu M_{VRS} sú Pareto-efektívne útvary A,B a D. Nesmieme zabudnúť na to, že v obálkovej analýze dát je technologická množina T aproximovaná pomocou vybra-

nej skupiny útvarov, nikdy ju nevieme presne určiť a z toho dôvodu dokážeme vypočítať iba relatívnu efektívnosť jednotlivých útvarov.

V tejto práci sa budeme zaoberať iba s technológiami nad konštantným výnosom z rozsahu, podľa ktorých spotrebovaním m vstupov je možné produkovať jeden výstup. Potom množstvá vstupov $x_i, i = 1, \dots m$ a výstupu y jednotlivých útvarov uložíme do matíc X veľkosti $m \times n$ a Y veľkosti $1 \times n$, pričom pre zložky x_{ij} matice X a pre zložky y_j matice Y platí nasledovný predpoklad:

Predpoklad 1.1. *Pre všetky* $j = 1, \dots, n$ *platí* $x_{ij} > 0, \forall i$ *a* $y_j > 0$.

Danú technológiu je možné charakterizovať technologickou množinou $T \subset \mathbb{R}_+^{m+1}$, ktorú sme definovali v kapitole 1.1.

1.3 DEA modely založené na parametrickej ceste

V tejto podkapitole predstavíme triedu DEA modelov, ktorých presnosť sme sa rozhodli testovať, pričom vedomosti čerpáme z publikácie [8]. Sú to DEA modely založené na parametrickej ceste, ktorých podľa publikácie [8] môžeme zapísať pomocou nasledovnej všeobecnej schémy v prípade konštantných výnosov z rozsahu:

$$(\mathrm{GS})_o : \min_{\theta,\lambda} \ \theta \tag{28}$$

$$s.t. X\lambda \leq x_o + (\psi^x(\theta) - 1)\, g_o^x, \tag{29}$$

$$Y\lambda \geq y_o + (\psi^y(\theta) - 1)\, g_o^y, \tag{30}$$

$$\lambda \geq 0, \tag{31}$$

kde symbolmi $g_o^x \in \mathbb{R}_+^m, g_o^y \in \mathbb{R}_+$ sú označené zložky smerového vektora $g_o = (g_o^x, g_o^y)$, ďalej symbolmi $\psi^x(\theta)$ a $\psi^y(\theta)$ sú označené hladké funkcie. Zvolený smerový vektor g_o, a rôzne definované funkcie $\psi^x(\theta), \psi^y(\theta)$ spolu určia príslušný tvar parametrickej cesty a následne aj vlastnosti jednotlivých DEA modelov patriacich do vybranej triedy modelov. Dobre definovaný smerový vektor musí spĺňať podmienku [8]:

$$g_o^x \geq 0, \ g_o^y \geq 0, \ \wedge \ g_o = (g_o^x, g_o^y) \neq 0. \tag{32}$$

Podľa publikácie [8] hladké funkcie $\psi^x(\theta), \psi^y(\theta)$ musia byť zvolené pri predpokladoch:

$$\text{a)} \quad \text{dom}(\psi^x) = (a^x, \infty),\ a^x \in \{-\infty, 0\}\ \wedge\ \text{dom}(\psi^y) = (a^y, \infty),\ a^y \in \{-\infty, 0\}, \tag{33}$$

$$\text{b)} \quad \psi^x \text{ je rastúca, konkávna funkcia}\ \wedge\ \psi^y \text{ je klesajúca, konvexná funkcia}, \tag{34}$$

$$\text{c)} \quad \psi^x(1) = \psi^y(1) = 1, \tag{35}$$

$$\text{d)} \quad \text{im}(\psi^x) = (b^x, \infty),\ b^x = -\infty \text{ ak } g_o^y = 0,\ \text{inak } b^x \in \{-\infty, 0\}\ \wedge\ \text{im}(\psi^y) = (b^y, \infty),\ b^y \in \{-\infty, 0\}, \tag{36}$$

kde symbolom $\text{dom}(\psi^x)$ je označený definičný obor a symbolom $\text{im}(\phi^x)$ obraz funkcie ψ^x, analogicky aj pre funkciu ψ^y. Uvedená všeobecná schéma (28)-(31) pokrýva nielen základné DEA modely, ktorými sú napríklad vstupne resp. výstupne orientovaný CCR model, ale aj modely založené na hyperbolickej alebo smerovej funkcie vzdialenosti.

Na testovanie presnosti sme vybrali tri DEA modely zo skupiny modelov založených na parametrickej ceste, a to výstupne orientovaný CCR model označený skratkou CCR-O, model založený na hyperbolickej funkcie vzdialenosti označený skratkou HDF, a model DDM založený na smerovej funkcie vzdialenosti. Dôvod, prečo sme si vybrali iba výstupne orientovanú verziu modelu CCR vyplýva z toho, že pri konštantných výnosoch z rozsahu hodnoty efektivít θ_j získané aplikovaním vstupne resp. výstupne orientovaného modelu CCR sú rovnaké [7]. To znamená, že aj presnosť týchto dvoch modelov bude rovnaká, preto stačí skúmať iba jeden z dvoch modelov. V Tabuľke 1.1 sme zhrnuli hodnoty zložiek smerového vektora g_o a hladkých funkcii $\psi^x(\theta), \psi^y(\theta)$ pre vybrané DEA modely:

	$\mathbf{g_o^x}$	$\mathbf{g_o^y}$	$\psi^{\mathbf{x}}(\theta)$	$\psi^{\mathbf{y}}(\theta)$
CCR-O	0	y_o	ψ^x	$\frac{1}{\theta}$
HDF	x_o	y_o	θ	$\frac{1}{\theta}$
DDM	x_o	y_o	θ	$2-\theta$

Tabuľka 1.1: Hodnoty parametrov všeobecnej schémy pre vybrané DEA modely

Keďže v prípade modelu CCR-O zložka g_o^x nadobudne nulovú hodnotu, k nemu príslušná hladká funkcia $\psi^x(\theta)$ môže byť ľubovolne zvolená, lebo po dosadení do všeobecnej schémy tento výraz aj tak zmizne. Podľa [8] tieto modely spĺňajú podmienky (32)-(36) kladené na smerový vektor g_o a hladké funkcie $\psi^x(\theta), \psi^y(\theta)$.

DEA modely založené na parametrickej ceste sú známe vlastnosťou vyplývajúcou aj z názvu triedy, totiž projekciu DMU_o^* na hranicu množiny M_{CRS} pre sledovaný útvar hľadajú po rozličných parametrických cestách príslušných vybraným DEA modelom. Označme pravé strany nerovností (29) a (30) nasledovne:

$$\phi_o^x(\theta) = x_o + (\psi^x(\theta) - 1)g_o^x,$$
$$\phi_o^y(\theta) = y_o + (\psi^y(\theta) - 1)g_o^y,$$

potom funkcia:

$$\phi_o : \theta \to (\phi_o^x(\theta), \phi_o^y(\theta)), \tag{37}$$

vyjadruje súvislú cestu v priestore vstupov a výstupu $\mathbb{R}^m \times \mathbb{R}^1$ a je parametrizovaná pomocou θ. Vlastnosti parametrickej cesty (37) sú vymenované a dokázané v článku [8], uvedieme iba tie, pomocou ktorých je možné ľahšie predstaviť geometriu vybraných DEA modelov:

a) $\phi_o(1) = (x_o, y_o) = DMU_o \in M_{CRS}$,

b) $\exists\, \theta_o^* \in [0,1] : \phi_o(\theta_o^*) \in M_{CRS} \ \wedge\ \phi_o(\theta_o) \notin M_{CRS}\ \forall\, \theta_o < \theta_o^*$.

Potom z vlastnosti b) vyplýva, že projekciu pre DMU_o môžeme označiť nasledovne:

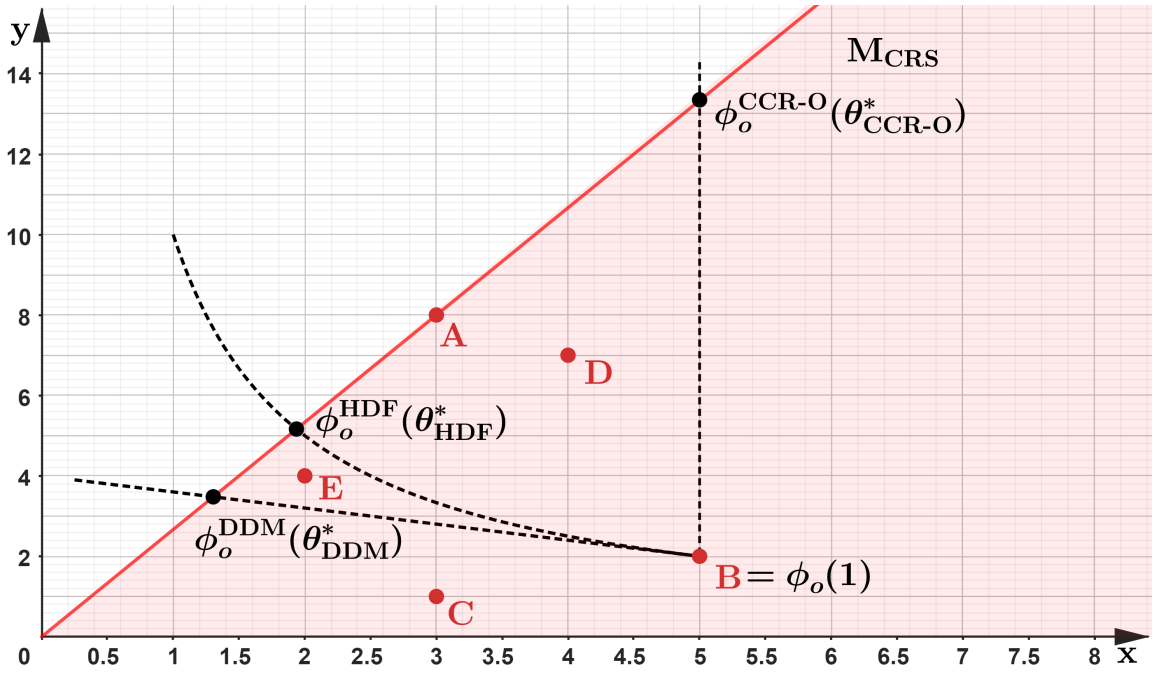
$$DMU_o^* = (x_o^*, y_o^*) = \phi_o(\theta_o^*) \in \partial M_{CRS},$$

kde θ_o^* označuje optimálne riešenie úlohy matematického programovania (28)-(31), čiže hodnotu efektivity sledovaného DMU_o. Treba ešte dodať, že v prípade modelu CCR-O hodnotu efektivity vypočítame zo vzťahu $\theta_o^* = \frac{1}{\xi_o^*}$, kde symbolom ξ_o^* je označené optimálne riešenie úlohy matematického programovania príslušnej modelu CCR-O definovanej v maximalizačnom tvare využitím vzťahu $\xi = \frac{1}{\theta}$. Geometriu DEA modelov založených na parametrickej ceste môžeme teda predstaviť ako parametrickú cestu $\phi_o(\theta)$, pre ktoré platí, že v bode $\phi_o(1)$ sa nachádza skúmaný útvar DMU_o. Následne pre klesajúce hodnoty parametra θ parametrická cesta ϕ_o pokračuje pomedzi dominantné útvary cez množinu M_{CRS} až na jej hranicu ∂M_{CRS}, čiže do bodu projekcie DMU_o^*.

Na príklade skupiny útvarov vymenovanej v Tabuľke 1.2 si ukážeme geometriu vybraných DEA modelov z triedy modelov založených na parametrickej ceste. Parametrické cesty príslušné modelom CCR-O, HDF a DDM sú znázornené na Obr.1.6 v priestore $\mathbb{R}^2$.

DMU_o	x_o	y_o
A	3	8
B	5	2
C	3	1
D	4	7
E	2	4

Tabuľka 1.2: Skupina útvarov použitá na znázornenie tvaru parametrických ciest



Obr. 1.6: Znázornenie tvaru parametrickej cesty v prípade vybraných DEA modelov

Červené body označujú miesta útvarov v množine M_{CRS}. Sústredíme sa na útvar B= $(5, 2)$ pomocou ktorého sme znázornili parametrické cesty $\phi_o^{\text{CCR-O}}, \phi_o^{\text{HDF}}$ a ϕ_o^{DDM} príslušné vybraným DEA modelom. Vidíme, že v prípade modelu CCR-O ide o špecifický prípad orientovaného radiálneho modelu, čiže definovaná parametrická cesta $\phi_o^{\text{CCR-O}}$ predstavuje maximálne radiálne predĺženie výstupu y_B pri nezmenených hodnotách vstupov x_B tak, aby sme sa dostali na hranicu množiny M_{CRS}. Parametrická cesta ϕ_o^{DDM} v prípade DDM modelu je zobrazená lúčom vo vopred určenom smere, ktorý smeruje od sledovaného útvaru B ku hranici množiny M_{CRS}. Model HDF vznikol kombináciou vstupne a výstupne orientovaných radiálnych modelov, preto pomocou cesty hyperbolického tvaru ϕ_o^{HDF} umožňuje súčasné ekviproporcionálne skracovanie vstupu x_B a predlžovanie výstupu y_B. Čierne

body $\phi_o^{\text{CCR-O}}\left(\theta^*_{\text{CCR-O}}\right), \phi_o^{\text{HDF}}\left(\theta^*_{\text{HDF}}\right)$ a $\phi_o^{\text{DDM}}\left(\theta^*_{\text{DDM}}\right)$ označujú projekcie útvaru B na hranicu množiny M_{CRS}, čiže efektívne vzory útvaru B.

1.3.1 Vlastnosti vybraných DEA modelov

Nakoniec definujeme zopár z vlastností DEA modelov založených na parametrickej ceste, ktoré podľa [8] platia pre skúmané modely CCR-O, HDF a DDM.

Definícia 1.7 (**Ohraničenosť**). *Hovoríme, že DEA model spĺňa vlastnosť ohraničenosti, ak pre hodnotu efektivity θ_j^* získanej aplikovaním DEA modelu na vybranú skupinu útvarov platí $\theta_j^* \in [0,1]$, $\forall j = 1, \ldots, n$.*

Ohraničenosť DEA modelov je jedna dobrá vlastnosť, ktorá zaručí, že hodnoty efektivít pochádzajú z rozumného intervalu $[0,1]$.

Definícia 1.8 (**Jednoznačná projekcia**). *Hovoríme, že DEA model spĺňa vlastnosť jednoznačnej projekcie, ak pre sledovaný útvar DMU_o poskytne jednoznačnú projekciu DMU_o^* na hranicu množiny M_{CRS}.*

Výhodou pri vybraných DEA modeloch je, že pre každé DMU_j poskytnú jednoznačnú projekciu $DMU_j^* = (x_j^*, y_j^*) \in \partial M_{CRS}$, aj napriek tomu, že DEA modely niekedy môžu pripúšťať viacero možností pre určenie efektívneho vzoru sledovaného útvaru.

Definícia 1.9 (**Monotónnosť**). *Hovoríme, že DEA model spĺňa vlastnosť monotónnosti, ak zvýšenie množstva ľubovolného vstupu alebo zníženie množstva ľubovolného výstupu DMU_o, pri nezmenených množstiev ostatných vstupov a výstupov DMU_o vedie k zníženiu alebo udržiavaniu hodnoty efektivity θ_o^*.*

Nakoniec definujme jednu vlastnosť DEA modelov nazývanú ako invariantnosť na zmenu jednotiek [7]:

Definícia 1.10 (**Invariantnosť na zmenu jednotiek**). *DEA model sa nazýva invariantným na zmenu jednotiek, ak po vykonaní transformácie na vstupoch a/alebo na výstupoch sa nezmení klasifikácia DMU_j na efektívne a neefektívne, a v prípade každého DMU_o hodnota optimálneho riešenia θ_o^* príslušnej úlohy matematického programovania ostáva nezmenená.*

Transformáciu vstupov môžeme zapísať nasledovne:

$$\begin{pmatrix} x_{1j} \\ \vdots \\ x_{mj} \end{pmatrix} \rightarrow \begin{pmatrix} k_1 & & 0 \\ & \ddots & \\ 0 & & k_m \end{pmatrix} \begin{pmatrix} x_{1j} \\ \vdots \\ x_{mj} \end{pmatrix} \coloneqq \begin{pmatrix} x_{1j}^{'} \\ \vdots \\ x_{mj}^{'} \end{pmatrix}, \quad \forall j = 1, \ldots, n, \tag{38}$$

kde $k_i > 0$ pre $i = 1, \ldots, m$ vyjadruje zmenu jednotky i-teho vstupu. Ak jednotku i-teho vstupu nemeníme, tak $k_i = 1$. Ak maticu transformácie označíme symbolom K, tak transformáciu vstupov môžeme zapísať zjednodušene ako:

$$x_j \rightarrow Kx_j \coloneqq x_j^{'}, \quad \forall j = 1, \ldots, n. \tag{39}$$

Transformácia výstupov vyzerá analogicky:

$$\begin{pmatrix} y_{1j} \\ \vdots \\ y_{sj} \end{pmatrix} \rightarrow \begin{pmatrix} l_1 & & 0 \\ & \ddots & \\ 0 & & l_s \end{pmatrix} \begin{pmatrix} y_{1j} \\ \vdots \\ y_{sj} \end{pmatrix} \coloneqq \begin{pmatrix} y_{1j}^{'} \\ \vdots \\ y_{sj}^{'} \end{pmatrix}, \quad \forall j = 1, \ldots, n, \tag{40}$$

kde analogicky $l_r > 0$ vyjadruje zmenu jednotky r-teho výstupu. Ak maticu transformácie označíme symbolom L, tak transformáciu vstupov môžeme zapísať zjednodušene ako:

$$y_j \rightarrow Ly_j \coloneqq y_j^{'}, \quad \forall j = 1, \ldots, n. \tag{41}$$

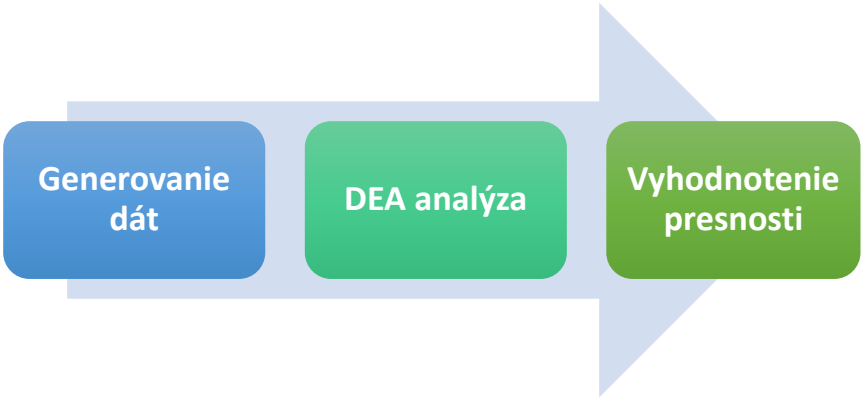
Invariantnosť na zmenu jednotiek je dôležitá vlastnosť, ktorá vyjadruje, že ak vstupné údaje jednotlivých DMU_j sú zadané v iných jednotkách, nezávisle od toho, že ktorú jednotku si zvolíme pre zjednotenie zadaných údajov, efektivity útvarov získané aplikovaním vybraného DEA modelu na zvolenú skupinu útvarov zostanú rovnaké.

Po vymenovaní výhod vybraných DEA modelov spomeňme aj jednu nevýhodu ktorá hovorí o tom, že projekcie DMU_j^* nemusia byť v každom prípade Pareto-efektívne.

1.4 Testovanie presnosti DEA modelov

Ako sme to už viackrát v tejto práci spomenuli, motiváciou zaoberať sa problematikou testovania presnosti DEA modelov môže byť uľahčenie výberu vhodného DEA modelu na vykonanie DEA analýzy, keďže v súčasnej literatúre je definovaných veľa rôznych DEA modelov. Informácia o presnosti jednotlivých DEA modelov môže vo veľkej miere pomôcť nájsť uplatnenie obálkovej analýzy dát v reálnom svete, napríklad u manažérov a politikov.

Tradičný postup testovania presnosti DEA modelov používaný v existujúcej literatúre sa dá charakterizovať pomocou všeobecnej schémy zobrazenej na Obr. 1.7 [10, 19]:

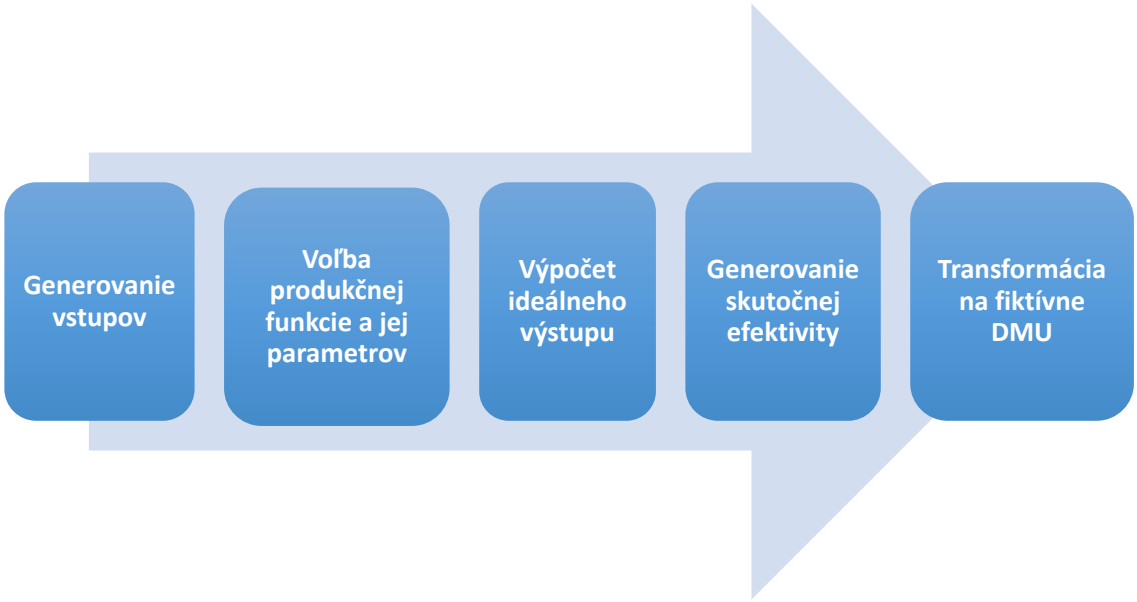


Obr. 1.7: Schéma testovania presnosti DEA modelov

Schéma sa skladá z troch blokov: generovanie dát, DEA analýza a vyhodnotenie presnosti. Keďže na testovanie presnosti DEA modelov nemôže použiť dáta o útvaroch z reálneho sveta, potrebné dáta vstupujúce do experimentov, čiže množstvá vstupov a výstupu jednotlivých útvarov treba generovať. Ako prebieha proces generovania dát podrobne predstavíme v nasledujúcej kapitole. Keď už máme vygenerované potrebné dáta, môžeme vykonať DEA analýzu, to znamená, že aplikujeme vybrané DEA modely na skupinu útvarov, a získame jednak hodnoty efektivít, ako aj výsledné poradie jednotlivých útvarov. Po získaní informácie o relatívnej efektívnosti jednotlivých útvarov môžeme vyhodnotiť presnosť skúmaných DEA modelov použitím vybraných štatistických prístupov, ktoré popíšeme v tretej kapitole našej práce.

2 Proces generovania dát

V tejto kapitole podrobne popíšeme proces generovania dát, čiže množstiev vstupov a výstupu jednotlivých fiktívnych útvarov. Akým postupom získame údaje fiktívnych útvarov je znázornené na Obr.2.1:



Obr. 2.1: Schéma procesu generovania dát

Po prvé, potrebujeme vygenerovať množstvá jednotlivých vstupov x_{ij} pre každé DMU_j. Aby sme vedeli spočítať hodnotu výstupu y_j pre fiktívne útvary, potrebujeme najprv vhodne zvoliť produkčnú funkciu. Zvolená produkčná funkcia má byť dostatočne flexibilná, aby sme jej využitím dokázali modelovať dostatočne rôznorodé produkčné procesy, a tak napodobniť jednotlivé technológie z reálneho sveta. Po voľbe produkčnej funkcie a jej parametrov dokážeme vypočítať hodnoty ideálnych výstupov DMU_j. Hodnota ideálneho výstupu vyjadruje maximálne množstvo výstupu, ktoré sa dá vyrobiť z vektora vstupov x_j v tej technológii, ktorú modelujeme vybranou produkčnou funkciou. V ďalšom kroku budeme generovať hodnotu skutočnej efektivity pre fiktívne útvary. Túto hodnotu potrebujeme práve kvôli tomu, aby sme ňou vedeli porovnať hodnotu efektivity získanej aplikovaním vybraného DEA modelu na skupinu fiktívnych útvarov, a tak otestovať presnosť vybraných DEA modelov. V poslednom kroku potrebujeme upraviť generované hodnoty vstupov a ideálneho výstupu tak, aby sme ku každému útvaru priradili generovanú

hodnotu skutočnej efektivity a získali výslednú skupinu fiktívnych útvarov vstupujúcu do DEA analýzy.

2.1 Generovanie vstupov

Začiatočným krokom procesu generovania dát je generovanie množstiev jednotlivých vstupov x_{ij}, $i = 1, \ldots, m$ pre každé DMU_j. Najprv vytvoríme maticu vstupov X veľkosti $m \times n$, ktorej zložky x_{ij} sú generované z rovnomerného rozdelenia $U \sim (10, 1000)$, čiže množstvá jednotlivých vstupov generujeme z rozsahu hodnôt $[10, 1000]$. Podľa článku [10] by bolo vhodné vykonať takú transformáciu $X \to X^G$ na matici vstupov X, aby po vykonaní transformácie korelácia medzi vstupom x_i a $x_h, i \neq h$, čiže medzi riadkami matice X^G bola rovná hodnote $\rho \in [0, 1)$. Táto úprava naznačuje, že by mal existovať nejaký vzťah medzi hodnotami vstupov jednotlivých útvarov zo skupiny útvarov zaoberajúcou sa vybranou technológiou.

Teraz odvodíme jednoduchú transformáciu, ktorá využíva Choleského rozklad kovariančných matíc, a pomocou ktorej dokážeme upraviť maticu vstupov X na maticu vstupov X^G [17]. Na začiatku máme k dispozícii maticu vstupov $X_{m \times n}$. Definujme maticu C nasledovne:

$$C = \left(I_n - \frac{1}{n} e_n e_n^T \right),$$

kde matica I_n označuje maticu identity veľkosti $n \times n$, ďalej vektor $e_n = (1, \ldots, 1)^T$ je vektor jednotiek dĺžky n. Potom kovariančnú maticu S_X môžeme vyjadriť nasledovným spôsobom:

$$X_C = XC,$$
$$S_X = \frac{1}{n} X_C X_C^T.$$

Predtým, ako uvedieme predpis pre korelačnú maticu R_X matice vstupov X, definujme maticu D_X, čo je diagonálna matica veľkosti $m \times m$ a pomocnú maticu X_S veľkosti $m \times n$ nasledovne:

$$D_X = diag\left(\sqrt{(S_X)_{ii}} \right), \ (S_X)_{ii} > 0 \ \forall i,$$
$$X_S = D_X^{-1} X_C.$$

Potom korelačnú maticu R_X môžeme vyjadriť nasledovne:

$$R_X = \frac{1}{n} X_S X_S^T = \frac{1}{n} D_X^{-1} X_C X_C^T D_X^{-1} = D_X^{-1} S_X D_X^{-1}.$$

Označme korelačnú maticu pre X^G symbolom R_{X^G}, potom R_{X^G} má jednotky na diagonále a jej mimodiagonálne zložky sa rovnajú hodnote ρ. Hľadáme takú transformáciu:

$$Z = FX, \tag{42}$$

t.j. maticu transformácie F, pre ktorú bude platiť, že korelačná matica R_Z spĺňa rovnicu:

$$R_Z = \frac{1}{n} Z_S Z_S^T = R_{X^G},$$

pričom platia tiež rovnice:

$$Z_C = ZC, \tag{43}$$

$$S_Z = \frac{1}{n} Z_C Z_C^T, \tag{44}$$

$$D_Z^{-1} = \left(diag\left(\sqrt{(S_Z)_{ii}} \right) \right)^{-1}, \tag{45}$$

$$Z_S = D_Z^{-1} Z_C. \tag{46}$$

Vychádzať budeme z definície korelačnej matice R_Z, pričom využijeme rovnice (42)-(46):

$$R_Z = \frac{1}{n} Z_S Z_S^T = \frac{1}{n} D_Z^{-1} Z_C Z_C^T D_Z^{-1} = \frac{1}{n} D_Z^{-1} ZCC^T Z^T D_Z^{-1} = \frac{1}{n} D_Z^{-1} FXCC^T X^T F^T D_Z^{-1} =$$
$$= \frac{1}{n} D_Z^{-1} F X_C X_C^T F^T D_Z^{-1} = D_Z^{-1} F S_X F^T D_Z^{-1}.$$

Ak budeme predpokladať, že variancie vstupov x_i po vykonaní transformácie (42) sa nezmenia, t.j. $D_X = D_Z$, potom pre kovariančnú maticu S_Z platí:

$$S_Z = D_Z R_Z D_Z = D_X R_Z D_X = F S_X F^T. \tag{47}$$

Zaveďme pre Choleského rozklady kovariančných matíc S_X, S_Z nasledovné označenia:

$$S_X = LL^T,$$
$$S_Z = MM^T,$$

pričom matice L, M sú dolné trojuholníkové matice rozmeru $m \times m$ s kladnými zložkami na ich diagonále [18]. Potom rovnicu (47) môžeme prepísať do tvaru:

$$MM^T = S_Z = F S_X F^T = FLL^T F^T,$$

z ktorej následne vyplýva, že pre maticu transformácie F musí platiť:

$$M = FL \;\rightarrow\; F = ML^{-1}.$$

To znamená, že ak vykonáme na matici vstupov X transformáciu (42) dostaneme upravenú maticu vstupov $Z = X^G$, pre ktoré bude platiť:

$$Z = FX = ML^{-1}X = X^G \;\rightarrow\; R_Z = R_{X^G}.$$

Teda v matici X^G sa nachádzajú také vektory vstupov x_j jednotlivých DMU_j, pre ktoré už platí, že korelácia medzi i-tym a h-tym vstupom v prípade $i \neq h$ je rovná hodnote ρ.

2.2 Výpočet ideálneho výstupu

Vhodnú voľbu produkčnej funkcie a podmienky kladené na jej parametre sme už predstavili v podkapitole 1.1.2. Ukázali sme, že ak parametre translogaritmickej funkcie spĺňajú podmienky z Tvrdenia 1.5, tak potom translogaritmickú funkciu môžeme použiť na modelovanie produkčných procesov. Na to, aby sme vedeli vypočítať ideálny výstup pre DMU_o, potrebujeme mať k dispozícii konkrétny predpis funkcie (11), čiže potrebujeme zvoliť konkrétne hodnoty pre parametre α_i, σ_{ih} a ϑ translogaritmickej funkcie. Teraz predstavíme vhodné možnosti nastavenia parametrov translogaritmickej funkcie, ktoré budú spĺňať podmienky vymenované v Tvrdení 1.5, a pomocou ktorých budeme vedieť modelovať dostatočne rôznorodé produkčné procesy.

Podľa článku [10] parametre α_i môžeme počítať dvoma spôsobmi, tak aby boli jednotlivé parametre symetrické alebo naopak asymetrické. Symetria v parametroch α_i znázorňuje technológiu, v ktorom všetky vstupy sú rovnako dôležité pri produkcii výstupu. V tomto prípade parametre α_i definujeme nasledovným spôsobom:

$$\alpha_i = \frac{1}{m},\ i = 1, \ldots, m. \tag{48}$$

V opačnom prípade, ak chceme vyjadriť, že jednotlivé vstupy nie sú rovnako dôležité v danom produkčnom procese, tak na výpočet parametrov α_i podľa publikácie [10] použijeme nasledovný vzorec:

$$\alpha_i = \frac{i + m}{1,5m^2 + 0,5m},\ i = 1, \ldots, m, \tag{49}$$

ktorý zabezpečí, že s rastúcim indexom i budú vstupy x_i čoraz dôležitejšie pre produkciu. Pri uvedenom asymetrickom nastavení platí, že nárast dôležitosti je rovnaký v každom

kroku z i do $i+1$. Ako to je ukázané v článku [10], oba vzorce spĺňajú požadované podmienky na parameter α_i.

Parametre substitučnej distribúcie σ_{ih}, ktoré informujú o miery nahraditeľnosti vstupov x_i a x_h sú symetrické, t.j. $\sigma_{ih} = \sigma_{hi}, \forall i, h = 1, \dots m$, a tiež normalizované, t.j. $|\sigma_{ih}| \leq 1$. V publikácii [10] predstavili dve možnosti spôsobu výpočtu parametrov σ_{ih}, vyjadrujúce rovnomernosť alebo nerovnomernosť substitučnej distribúcie, pričom oba možnosti spĺňajú požadované podmienky z Tvrdenia 1.5. Podľa článku [10] v prípade, že máme produkciu, v ktorom jednotlivé vstupy sú v rovnakej miere zameniteľné medzi sebou, parametre σ_{ih} definujeme nasledovne:

$$\sigma_{ih} = \frac{1}{m-1},\ i, h = 1, \dots, m, i \neq h \quad \text{a} \quad \sigma_{hh} = -1,\ h = 1, \dots, m. \tag{50}$$

Na druhej strane, ak chceme znázorniť, že vstupy sú v rôznej miere zameniteľné, tak definíciu parametrov σ_{ih} zmeníme na nerovnomernú, podľa článku [10] nasledovne:

$$\sigma_{ih} = \frac{2 - \frac{h-1}{m-1} - \frac{i-1}{m-1}}{1,5m-2},\ i, h = 1, \dots, m, i \neq h, \tag{51}$$

$$\sigma_{hh} = -\left(\frac{m\left(1,5 - \frac{h-1}{m-1}\right) - \left(2 - \frac{2(h-1)}{m-1}\right)}{1,5m-2} \right),\ h = 1, \dots, m. \tag{52}$$

Posledný parameter translogaritmickej produkčnej funkcie ϑ, vyjadrujúca substitučnú schopnosť sme sa rozhodli zvoliť nasledovne:

$$\vartheta = \min_i \left\{ \frac{\alpha_i}{\ln x^{\max} \sum_{h=1}^{m} |\sigma_{ih}|} \right\}. \tag{53}$$

Zmenu oproti prístupu predstaveného v článku [10] pri modelovaní produkčných procesov sme vykonali pri voľbe hodnoty parametra ϑ. Autori článku [10] hodnotu ϑ určili riešením úloh matematickej optimalizácie:

$$\begin{aligned}
&\min / \max\ \vartheta \\
&\alpha_i + \sum_{h=1}^{m} \sigma_{ih}\, \vartheta\ \ln x_h \geq 0,\ \forall i \\
&\sigma_{ii}\, \vartheta\ + \left(\sum_{h=1}^{m} \sigma_{ih}\, \vartheta\ \ln x_h \right)^2 - \left(\sum_{h=1}^{m} \sigma_{ih}\, \vartheta\ \ln x_h \right) \leq 0,\ \forall i,
\end{aligned}$$

kde minimálna hodnota ϑ poukazuje na vysokú substitučnú schopnosť modelovanej technológie, naopak maximálna hodnota na nízku substitučnú schopnosť. Odvodením jednoduchej podmienky kladenej na parameter substitučnej schopnosti ϑ sme vedeli obísť časovo

náročné riešenia optimalizačných úloh a dokázali sme navrhnúť numericky efektívnejšie modelovanie produkčných procesov.

Dostali sme sa k tretiemu kroku v procese generovania dát, a to k výpočte ideálnej hodnoty výstupu. Voľbou spôsobu nastavenia jednotlivých parametrov funkcie (11) dostaneme konkrétne vyjadrenie produkčnej funkcie $p_T(x)$, ktorou modelujeme produkčný proces. Dosadením generovaných hodnôt vstupov x_j^G do tejto funkcie môžeme vypočítať množstvo ideálneho výstupu $p_T\left(x_j^G\right)$ jednotlivých útvarov, ktoré označíme symbolom $y_j^G = p_T\left(x_j^G\right)$. Potom $DMU_o^G = (x_o^G, y_o^G)$ nazveme ako ideálny útvar. Prívlastok ideálny naznačuje, že všetky útvary produkujú maximálne možné množstvo výstupu, ktoré sa dá dosiahnuť modelovanou technológiou. To znamená, že ideálne útvary sa nachádzajú na krivke translogaritmickej produkčnej funkcie v množine T, ktorá je definovaná príslušnou voľbou jej parametrov.

2.3 Generovanie skutočnej efektivity

Ako sme to už v našej práci naznačili, testovanie presnosti DEA modelov môžeme vykonať iba pomocou skupiny fiktívnych útvarov. Totiž potrebujeme mať k dispozícii informáciu o takzvanej skutočnej efektivite jednotlivých útvarov, ktorú by sme v prípade útvarov z reálneho sveta nemali, a potom by sme nevedeli s akou hodnotou porovnať hodnotu efektivity získaného vykonaním DEA analýzy. Hodnotu skutočnej efektivity pre DMU_j označujeme symbolom $\hat{\theta}_j$. Informáciu o skutočnej efektivite fiktívnych útvarov využijeme jednak pri transformácii ideálnych útvarov na fiktívne, ako aj pri vyhodnocovaní presnosti vybraného DEA modelu.

Podľa existujúcej literatúry ohľadom generovania hodnôt skutočnej efektivity pre fiktívne útvary môžeme vybrať z dvoch možností [10, 11, 14]:

1. Začať voľbou počtu efektívnych fiktívnych útvarov c v danej skupine útvarov, a potom generovať využitím vhodného pravdepodobnostného rozdelenia hodnotu skutočnej efektivity pre ostatné (neefektívne) fiktívne útvary zo skupiny.

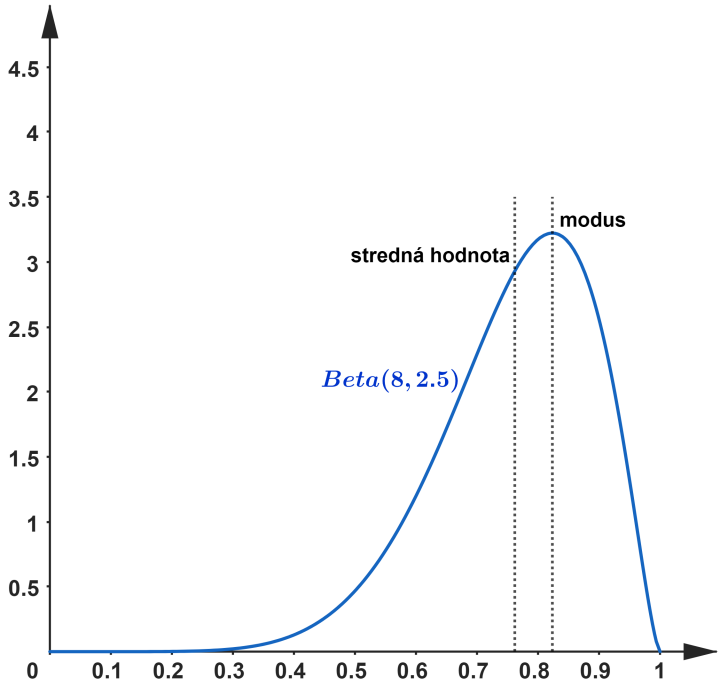
2. Generovať hodnotu skutočnej efektivity zo zvoleného pravdepodobnostného rozdelenia pre všetky fiktívne DMU_j, bez prvotnej voľby počtu efektívnych fiktívnych útvarov c.

Autori článku [10], z ktorého vychádzame, odporúčali generovať hodnotu skutočnej efektivity pomocou normálneho rozdelenia $\hat{\theta} \sim N(\mu, \sigma)$ pri rôznych nastaveniach jej parametrov a vhodnej voľbe dolnej hranice d pre hodnoty $\hat{\theta}_j \in [d, 1]$. Príkladom je generovanie hodnoty skutočnej efektivity pre DMU_j spôsobom $\hat{\theta} \sim N(0.8, 0.25^2)$ pričom $\hat{\theta}_j \in [0.4, 1]$. Voľba dolnej hranice d poukazuje na situáciu z reálneho sveta, teda na to, že útvary s veľmi nízkou efektivitou v reálnom svete zaniknú. Rôzne voľby parametrov μ, σ umožňujú testovať hypotézu, či rôzna voľba generovania hodnôt skutočnej efektivity má vplyv na presnosť skúmaných DEA modelov. To znamená, že článok [10] je príkladom druhej možnosti, keď vytvoríme skupinu fiktívnych útvarov bez prvotnej voľby počtu efektívnych fiktívnych útvarov. Potom sa môže stať, že v skupine fiktívnych útvarov sa nebude nachádzať ani jeden efektívny fiktívny útvar.

Avšak DEA modelovanie vyhodnocuje relatívnu efektívnosť útvarov v rámci sledovanej skupiny útvarov. To znamená, že nejakú časť útvarov zo skúmanej skupiny vyhodnotí ako efektívne útvary, t.j. $\theta_j^* = 1$. Preto sme sa rozhodli vybrať spomenutú prvú možnosť, teda pri generovaní hodnôt skutočnej efektivity si vopred zvolíme počet efektívnych fiktívnych útvarov c v danej skupine. Na generovanie hodnoty skutočnej efektivity $\hat{\theta}_j, j = c+1, \ldots, n$ sme zvolili spojité pravdepodobnostné rozdelenie $Beta(\alpha, \beta)$, ktoré je definované na intervale $(0, 1)$. Parametre $\alpha, \beta > 0$ zodpovedajú za tvar funkcie hustoty pravdepodobnosti rozdelenia $Beta$. Funkciou hustoty pravdepodobnosti spojitej náhodnej veličiny vieme určiť pravdepodobnosť, ktorou spadá náhodná veličina (v našom prípade $\hat{\theta}$) do určitého rozsahu hodnôt. Veľkosť pravdepodobnosti je daná plochou pod funkciou hustoty pravdepodobnosti a nad horizontálnou osou v danom rozsahu hodnôt [9]. Myslíme si, že $Beta$ rozdelenie môže byť vhodnou voľbou práve preto, lebo je definované na tom intervale, z ktorého by mali pochádzať aj hodnoty skutočnej efektivity, a je možné ľahko pomocou jej parametrov navrhnúť vhodný tvar pre funkciu hustoty pravdepodobnosti.

Na Obr.2.2 je znázornená funkcia hustoty pravdepodobnosti pre $\hat{\theta} \sim Beta(8, 2.5)$. Tvar hustoty pravdepodobnosti pre $Beta$ rozdelenie sme riadili voľbou parametrov α, β, pričom cieľom bolo navrhnúť takú voľbu parametrov, pomocou ktorej dokážeme dostatočne dobre napodobniť zloženie skupiny útvarov v reálnom svete, t.j. málo útvarov s nízkou skutočnou efektivitou a priemerná hodnota skutočných efektivít očakávaná okolo hodnoty 0.8. Chceli sme teda obísť generovanie veľmi nízkych hodnôt skutočnej efektivity, a dosiahnuť aby

hodnoty skutočnej efektivity boli hlavne okolo 0.8.



Obr. 2.2: Funkcia hustoty pravdepodobnosti pre $Beta(8, 2.5)$

Na Obr.2.2 sme vyznačili aj strednú hodnotu a modus pre $\hat{\theta} \sim Beta(8, 2.5)$, ktoré hodnoty sme vypočítali podľa článku [13]:

$$\text{stredná hodnota} = \frac{\alpha}{\alpha + \beta} = \frac{8}{8 + 2.5} \doteq 0.76,$$
$$\text{modus} = \frac{\alpha - 1}{\alpha + \beta - 2} = \frac{8 - 1}{8 + 2.5 - 2} \doteq 0.82.$$

Vidíme, že ak na generovanie skutočnej efektivity použijeme rozdelenie $Beta(8, 2.5)$, potom hodnoty $\hat{\theta}_j$ nadobudnú iba s veľmi malou pravdepodobnosťou hodnotu pod 0.4, očakávame približne $\hat{\theta}_j \in (0.3, 1)$, pričom najviac fiktívnych útvarov bude mať skutočnú efektivitu v intervale $[0.6, 0.95]$.

2.4 Transformácia ideálnych útvarov na fiktívne útvary

Proces generovania dát je v stave, keď máme vygenerované ideálne útvary, ktoré sa nachádzajú na krivke translogaritmickej funkcie, t.j. skupinu ideálnych útvarov, ktorá sa skladá podľa teórie mikroekonómie z výlučne efektívnych útvarov. V predošlej podkapitole sme opísali, akým spôsobom bude prebiehať generovanie hodnôt skutočnej efektivity,

teda máme aj informáciu o hodnote $\hat{\theta}_o$ pre každé DMU_o. Teraz je potrebné zohľadniť generované hodnoty skutočnej efektivity a vykonať transformáciu ideálnych útvarov na fiktívne útvary, pričom upravené dáta pre DMU_o budú označené ako x_o^F a y_o^F.V článku, z ktorého vychádzame definovali nasledovnú úpravu dát [10]:

$$x_o^F = x_o^G, \tag{54}$$

$$y_o^F = \hat{\theta}_o y_o^G, \tag{55}$$

čo ale znamená, že je zohľadnená iba výstupná efektivita jednotlivých útvarov. Zohľadňovanie výlučne výstupných efektivít pri transformácii ideálnych útvarov na fiktívne útvary je v rozpore s ideou neorientovaných DEA modelov, ktoré zohľadňujú vstupno-výstupné efektivity jednotlivých útvarov.

Idea novej transformácie bola inšpirovaná geometriou DEA modelov založených na parametrickej ceste. Podľa nami navrhnutého nového prístupu pri transformácii ideálnych útvarov na fiktívne je potrebné brať ohľad aj na zvolený DEA model. Tým pádom dokážeme zohľadniť vstupno-výstupné efektivity jednotlivých útvarov v prípade neorientovaných DEA modelov, a vstupné resp. výstupné efektivity v prípade orientovaných DEA modelov. Matematický popis navrhnutej transformácie je nasledovný:

$$DMU_o^F = \left(x_o^F, y_o^F\right) : \phi_o^{\mathrm{DEA}}\left(\hat{\theta}_o\right) = \left(x_o^G, y_o^G\right), \tag{56}$$

kde ϕ_o^{DEA} označuje parametrickú cestu vybraného DEA modelu. Kľúčovou myšlienkou transformácie je, že ideálne útvary budeme posúvať využitím hodnôt skutočnej efektivity do vnútra technologickej množiny po parametrickej ceste príslušnej vybranému DEA modelu, a tým pádom dokážeme zohľadniť geometriu vybraného DEA modelu.

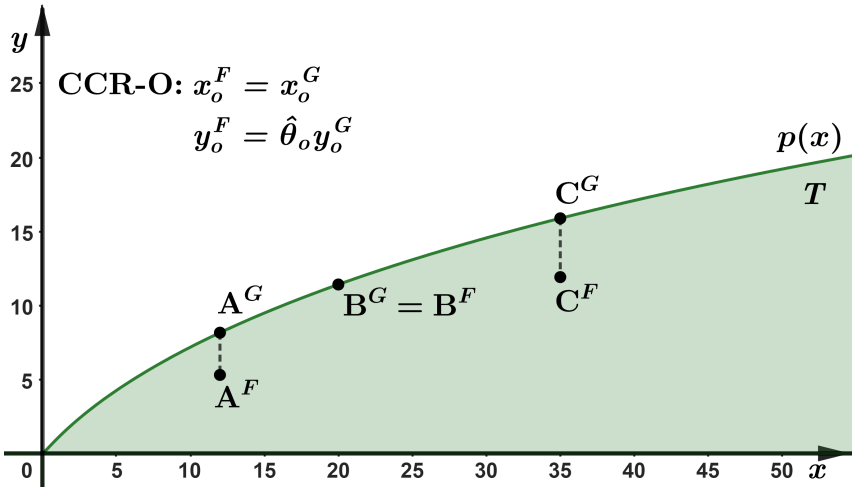
Predpokladajme, že poznáme technologickú množinu T a k nej príslušnú produkčnú funkciu $p(x) = x^{1-\frac{1}{16}\ln(x)}$, ktoré charakterizujú vybranú technológiu. Majme skupinu ideálnych útvarov A^G, B^G a C^G riadiacich sa vybranou technológiou a pre nich vygenerované hodnoty skutočnej efektivity $\hat{\theta}_o$ vymenované v Tabuľke 2.1. Na obrázkoch 2.3-2.5 sme znázornili geometrický pohľad na navrhnutú novú transformáciu (56) v prípade vybraných DEA modelov, ako aj matematický popis týchto transformácii. Ideálne útvary A^G, B^G a C^G sa nachádzajú na krivke produkčnej funkcie $p(x) = x^{1-\frac{1}{16}\ln(x)}$ v technologickej množine T. Budeme ich posúvať po parametrických cestách príslušných vybraným DEA modelom

do vnútra množiny T využitím hodnôt skutočnej efektivity $\hat{\theta}_o$, a tak dostaneme výslednú skupinu fiktívnych útvarov A^F, B^F a C^F na ktorú môžeme aplikovať DEA analýzu.

$\mathbf{DMU_o^G}$	$\hat{\theta}_o$
$\mathbf{A^G}$	0.65
$\mathbf{B^G}$	1
$\mathbf{C^G}$	0.75

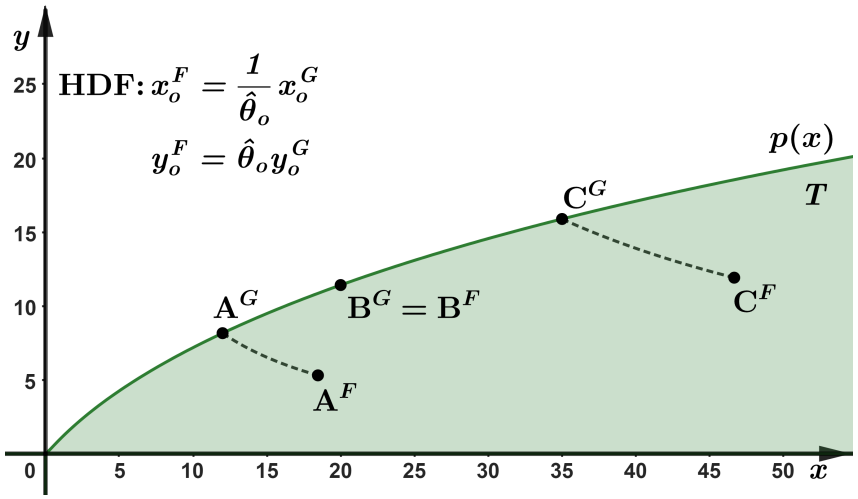
Tabuľka 2.1: Generované skutočné efektivity pre DMU_o

Na Obr.2.3 je znázornená transformácia v prípade výstupne orientovaného CCR modelu. Keďže tento model patrí medzi orientované modely, pri transformácii treba zohľadniť iba výstupnú efektivitu jednotlivých útvarov. Preto transformáciu ideálneho útvaru na fiktívny vykonáme spôsobom, že znížime hodnotu výstupu ideálneho útvaru využitím hodnoty skutočnej efektivity. Môžeme si všimnúť, že táto transformácia sa zhoduje s transformáciou (54)-(55) používanou v článku [10].



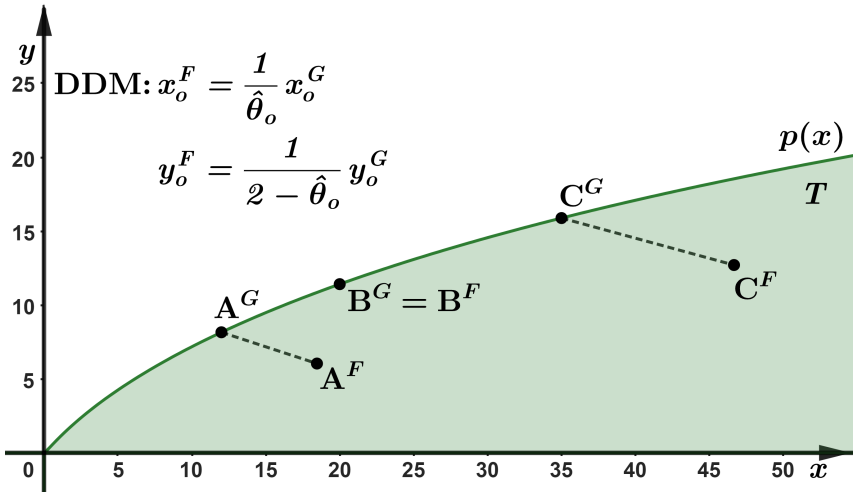
Obr. 2.3: Transformácia ideálnych útvarov na fiktívne útvary - CCR-O model

Na Obr.2.4 je znázornená transformácia v prípade hyperbolického modelu HDF. Tento model patrí medzi neorientované modely, takže pri transformácii treba zohľadniť vstupno-výstupné efektivity jednotlivých útvarov. Ideálny útvar využitím hodnoty skutočnej efektivity je posunutý do vnútra technologickej množiny po parametrickej ceste hyperbolického tvaru, pričom hodnota vstupu ideálneho útvaru je zvýšená a naopak hodnota výstupu znížená.



Obr. 2.4: Transformácia ideálnych útvarov na fiktívne útvary - HDF model

Na Obr.2.5 je znázornená transformácia v prípade DDM modelu. Tento model patrí tiež medzi neorientované modely, preto pri transformácii treba zohľadniť vstupno-výstupné efektivity jednotlivých útvarov. Ideálny útvar využitím hodnoty skutočnej efektivity je posunutý v smere určenom smerovým vektorom do vnútra technologickej množiny po parametrickej ceste tvaru úsečky, pričom hodnota vstupu ideálneho útvaru je zvýšená a hodnota výstupu znížená.



Obr. 2.5: Transformácia ideálnych útvarov na fiktívne útvary - DDM model

Výhoda predstavenej novej transformácie (56) spočíva v tom, že pri navrhnutej forme úpravy dát využitím geometrie vybraných DEA modelov dokážeme zohľadňovať vstupno-výstupné efektivity v prípade neorientovaných DEA modelov, a vstupné alebo výstupné

efektivity v prípade orientovaných DEA modelov. Preto si myslíme, že nová metodológia umožňuje korektnejšie porovnávanie DEA modelov založených na parametrickej ceste, ako štandardný prístup zohľadňujúci iba výstupnú efektivitu jednotlivých útvarov.

3 Vyhodnotenie presnosti DEA modelov

V tejto kapitole predstavíme štatistické ukazovatele používané na vyhodnotenie presnosti DEA modelov. Následne popíšeme, prečo je potrebné v problematike testovania presnosti DEA modelov využiť Monte Carlo simulácie. Vedomosti čerpáme hlavne z publikácie [10].

3.1 Štatistické ukazovatele

Presnosť vybraných DEA modelov budeme skúmať najmä z hľadiska identifikácie efektívnosti, a z hľadiska usporiadania útvarov. Pri výbere jednotlivých štatistických ukazovateľov sme vychádzali z publikácie [10]. Predtým, ako uvedieme definície vybraných štatistických ukazovateľov treba dodať, že kvôli lepšej prehľadnosti pri porovnávaní skúmaných DEA modelov každý jeden ukazovateľ bude upravený do takej formy, aby nadobúdal iba hodnoty z intervalu $[0, 1]$, pričom z pohľadu kvality sledovaného modelu najlepšia možná hodnota bude 1 a najhoršia 0.

Ako prvé, definujeme dva najpoužívanejšie štatistické ukazovatele v DEA, ktorými sú upravená forma priemernej absolútnej chyby, označená podľa anglického názvu *Mean Absolute Error* skratkou MAE, a Spearmanov korelačný koeficient (Spearmanov koeficient poradovej korelácie), označený skratkou SPEAR:

$$\mathrm{MAE} = 1 - \frac{1}{n}\sum_{j=1}^{n} |\hat{\theta}_j - \theta_j^*|, \tag{57}$$

$$\mathrm{SPEAR} = \frac{\sum_{j=1}^{n}\left(\mathrm{rank}\left(\hat{\theta}_j\right) - \overline{\mathrm{rank}\left(\hat{\theta}\right)}\right)\left(\mathrm{rank}\left(\theta_j^*\right) - \overline{\mathrm{rank}\left(\theta^*\right)}\right)}{\sqrt{\sum_{j=1}^{n}\left(\mathrm{rank}\left(\hat{\theta}_j\right) - \overline{\mathrm{rank}\left(\hat{\theta}\right)}\right)^2}\sqrt{\sum_{j=1}^{n}\left(\mathrm{rank}\left(\theta_j^*\right) - \overline{\mathrm{rank}\left(\theta^*\right)}\right)^2}}, \tag{58}$$

kde $\mathrm{rank}(\hat{\theta}_j), \mathrm{rank}(\theta_j^*)$ označujú poradové číslo j-teho fiktívneho útvaru v poradí útvarov zoradených zostupne podľa ich skutočnej efektivity, resp. podľa efektivity zrátaného aplikovaním vybraného DEA modelu. Ukazovateľ MAE priamočiaro určí kvalitu modelu spôsobom, že od najvyššej hodnoty kvality 1 je odrátaná priemerná absolútna chyba skúmaného modelu. Spearmanov korelačný koeficient je totožný s výberovým koeficientom korelácie aplikovaným na poradie útvarov, čiže vyjadruje aká je miera korelácie medzi poradím útvarov určeným podľa skutočnej efektivity fiktívnych útvarov a poradím určeným podľa efektivít získaných po aplikovaní vybraného DEA modelu na skupinu fiktívnych

útvarov. Čím vyššie budú tieto hodnoty, o to kvalitnejším je sledovaný model. Predstavené dva ukazovatele hodnotia tie očakávania, že podľa DEA modelov zrátané hodnoty efektivít by sa mali iba málo líšiť od skutočných hodnôt efektivít fiktívnych útvarov, a poradie fiktívnych útvarov zostavené podľa hodnoty efektivity získaného z DEA analýzy by mal korešpondovať s poradím zostaveným podľa skutočných hodnôt efektivít fiktívnych útvarov.

Pred predstavením ďalšieho štatistického ukazovateľa používaného na vyhodnotenie presnosti jednotlivých DEA modelov si zaveďme nasledovné označenia množín:

$$N_{\hat{\theta}} = \{j \in \{1, \ldots, n\} : \hat{\theta}_j \leq \hat{\mathrm{Q}}(\omega)\}, \tag{59}$$

$$N_{\theta^*} = \{j \in \{1, \ldots, n\} : \theta^*_j \leq \mathrm{Q}^*(\omega)\}, \tag{60}$$

$$N_{\hat{\theta} \cap \theta^*} = N_{\hat{\theta}} \cap N_{\theta^*} = \{j \in \{1, \ldots, n\} : \hat{\theta}_j \leq \hat{\mathrm{Q}}(\omega) \cap \theta^*_j \leq \mathrm{Q}^*(\omega)\}, \tag{61}$$

kde $\hat{\mathrm{Q}}(\omega)$, resp. $\mathrm{Q}^*(\omega)$, pričom $\omega = 0.15$, označujú tú hodnotu skutočnej efektivity $\hat{\theta}_j$, resp. efektivity θ^*_j, ktorá oddelí od skupiny fiktívnych útvarov dolných 15% útvarov, čiže tie najneefektívnejšie, pričom fiktívne útvary sú zostupne zoradené podľa ich skutočnej efektivity $\hat{\theta}$, resp. podľa efektivity θ^* získaného z DEA analýzy. V našom prípade, keď veľkosť skupiny fiktívnych útvarov je $n = 100$, potom hodnoty $\hat{\mathrm{Q}}(\omega), \mathrm{Q}^*(\omega)$ zodpovedajú hodnotám efektivít vystupujúcim na 86-tom mieste v jednotlivých poradiach fiktívnych útvarov. Pomocou vymenovaných množín fiktívnych útvarov môžeme ľahšie zadefinovať štatistického ukazovateľa označeného skratkou INEFF v nasledovnom tvare:

$$\mathrm{INEFF} = \frac{|N_{\hat{\theta} \cap \theta^*}|}{|N_{\hat{\theta}}|} \left(1 - \frac{\max\{|N_{\theta^*}| - |N_{\hat{\theta}}|, 0\}}{n}\right). \tag{62}$$

Štatistický ukazovateľ INEFF, ako sme to už popísali, kladie dôraz na najviac neefektívne útvary, hodnotí presnosť sledovaného DEA modelu využitím veľkosti definovaných množín (59)-(61). Podľa tohto ukazovateľa model bude o to kvalitnejší, keď veľkosti množín $N_{\hat{\theta}}$ a N_{θ^*} budú veľmi podobné (v najlepšom prípade totožné), a veľkosť prieniku týchto dvoch množín $N_{\hat{\theta}} \cap N_{\theta^*}$ bude blízke veľkosti množiny $N_{\hat{\theta}}$ (v najlepšom prípade totožné). Jednoducho povedané, kvalitný DEA model dáva také hodnoty efektivít θ^*_j, a následne také poradie fiktívnych útvarov, pri ktorom na posledných 15-tich miestach budú vystupovať práve tie fiktívne útvary, ktoré sú na posledných 15-tich miestach aj v poradí určenej podľa skutočnej efektivity fiktívnych útvarov.

Posledným štatistickým ukazovateľom, ktorého sme použili na hodnotenie presnosti vybraných DEA modelov je ukazovateľ CORRI definovaný nasledovne:

$$\text{CORRI} = \frac{1}{\gamma}\sum_{k=1}^{\gamma}\frac{|\{j \in \{1,\ldots,n\} : |\hat{\theta}_j - \theta_j^*| \leq k\delta\}|}{n}, \tag{63}$$

kde sme hodnoty parametrov zvolili podľa publikácie [10] ako $\gamma = 3$ a $\delta = 0,05$. Tento ukazovateľ určuje 3 hranice pre rozdiel hodnôt $|\hat{\theta}_j - \theta_j^*|$, a to 5%-ný, 10%-ný a 15%-ný rozdiel efektivít, následne sa môžeme pozrieť na percentuálny podiel počtu fiktívnych útvarov, ktorý sa zmestili do jednotlivých hraníc a vyrátať priemernú hodnotu z troch podielov. Keďže ukazovateľ MAE poskytuje iba informáciu o priemernej absolútnej chybe sledovaného DEA modelu, pomocou ukazovateľa CORRI môžeme sledovať schopnosť modelu umiestniť svoje odhady efektivít do určitých hraníc okolo hodnôt skutočnej efektivity fiktívnych útvarov.

Pomocou predstavených štyroch štatistických ukazovateľov hodnotíme presnosť vybraných DEA modelov, a to spôsobom, že vyrátame ich priemernú hodnotu B-Value:

$$\text{B-Value} = \frac{\text{MAE} + \text{SPEAR} + \text{INEFF} + \text{CORRI}}{4}. \tag{64}$$

Lenže, presnosť modelov vyjadrený cez priemernú hodnotu B-Value môže byť aj veľmi podobná pri skúmaných modeloch, a bolo by ťažké uvidieť nejaký rozdiel medzi nimi. Preto sa budeme pozerať aj na doplnkový ukazovateľ B-Rank:

$$\text{B-Rank} = \frac{\text{rank(MAE)} + \text{rank(SPEAR)} + \text{rank(INEFF)} + \text{rank(CORRI)}}{4}, \tag{65}$$

ktorý vyjadruje priemerné poradie skúmaného modelu pri vybraných štatistických ukazovateľoch. Táto hodnota nám môže pomôcť pri porovnávaní DEA modelov, ktorých presnosť bude veľmi podobná, keďže informuje o tom, aký je model dobrý vzhľadom k ostatným vyhodnocovaním modelom. Keďže sme sa rozhodli testovať presnosť troch DEA modelov, hodnota ukazovateľa B-Rank pochádza z intervalu $[1, 3]$, pričom nízka hodnota tohto ukazovateľa poukazuje na vyššiu kvalitu skúmaného DEA modelu.

3.2 Monte Carlo simulácie

Nakoniec treba ešte dodať, že kvôli procesu generovania dát, ktorého súčasťou je aj náhodnosť, na získanie vierohodných výsledkov o presnosti vybraných DEA modelov treba

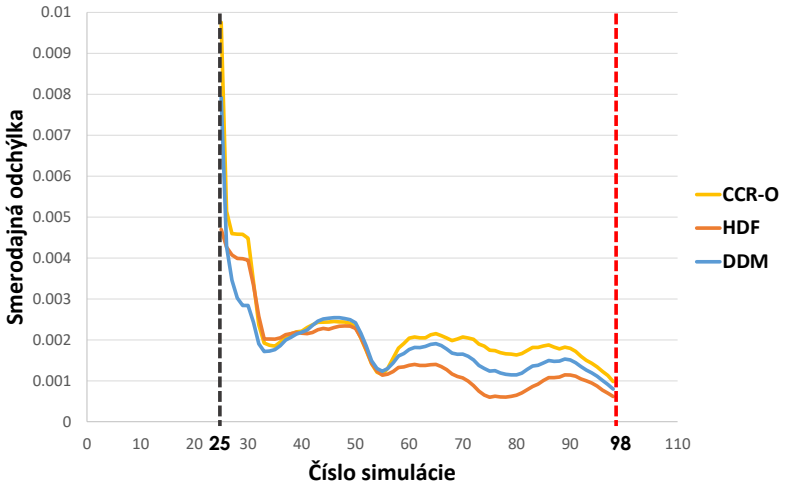
využiť Monte Carlo simulácie. To znamená, že postup testovanie presnosti DEA modelov pozostávajúci z troch blokov (Generovanie dát, DEA analýza, Vyhodnotenie presnosti) nemôžeme vykonať iba raz, ale budeme opakovať až kým nesplníme ukončovacie kritérium Monte Carlo simulácii stanovený v článku [10]. Matematický popis stanoveného kritéria na pozastavenie Monte Carlo simulácie vyzerá nasledovne:

$$\overline{\text{B-Value}}_{k-24} = \frac{1}{25}\sum_{i=k-24}^{k} \text{B-Value}_i, \quad k = 25, \ldots, K,\ K \geq 50,$$

$$\sigma_{k-24}^{\text{B-Value}} = \sqrt{\frac{1}{24}\sum_{i=k-24}^{k}\left(\text{B-Value}_i - \overline{\text{B-Value}}_{k-24}\right)^2},$$

$$\text{STOP} : \sigma_{k-24}^{\text{B-Value}} < 0.001,$$

teda celý postup zopakujeme minimálne 50-krát, pričom po 25-tom zopakovaní začneme rátať pohyblivú smerodajnú odchýlku pre hodnoty B-Value. Smerodajnú odchýlku budeme rátať vždy z posledných 25-tich realizovaných hodnôt B-Value, a keď hodnota $\sigma_{k-24}^{\text{B-Value}}$ v prípade každého skúmaného DEA modelu klesne pod hodnotu 0.001 ukončíme Monte Carlo simulácie. Potom výslednú hodnotu presnosti DEA modelu, ako aj výslednú hodnotu doplnkového ukazovateľa vyrátame ako priemer zo všetkých realizovaných hodnôt B-Value, resp. B-Rank, t.j. od 1 až po K.



Obr. 3.1: Ukážka vývoja smerodajných odchýlok v Monte Carlo simulácii

Na Obr.3.1 sme znázornili najdlhšie trvajúci prípad Monte Carlo simulácie spomedzi vykonaných experimentov na ukážku možného vývoja smerodajných odchýlok. Vidíme, že v konečnom dôsledku smerodajné odchýlky majú klesajúci charakter, konvergujú do hodnoty 0.001, ale počas trvania Monte Carlo simulácii môžu aj narastať z kroku i do $i+1$. To sa stane v prípade, keď posledná vykonaná simulácia v rade dáva značne odlišnú hodnotu B-Value, ako tie predošlé. V znázornenom prípade sme potrebovali až 98-krát zopakovať postup testovania presnosti DEA modelov na získanie vierohodných výsledkov o presnosti vybraných DEA modelov.

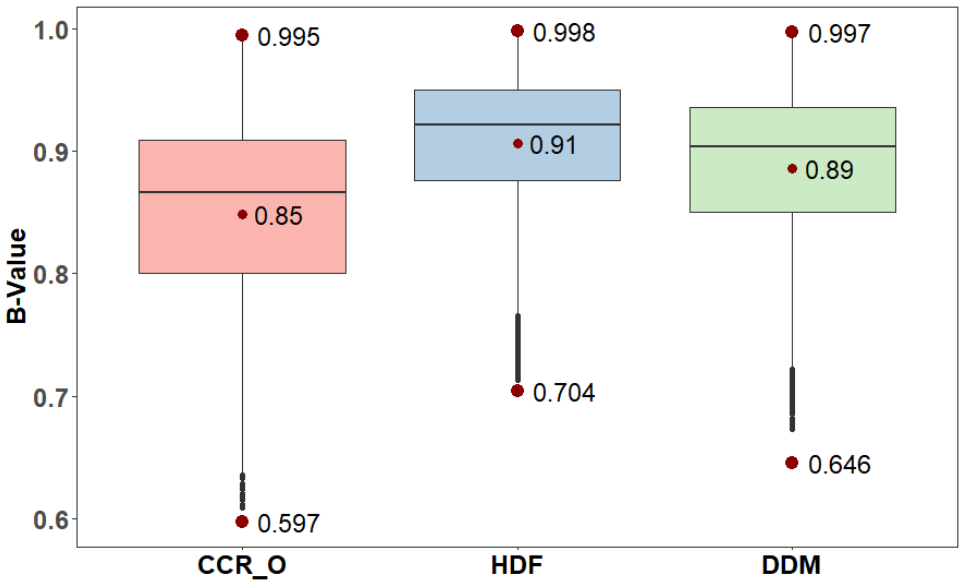
4 Výsledky vykonaných experimentov

V poslednej kapitole diplomovej práce predstavíme výsledky testovania presnosti DEA modelov získané z vykonaných experimentov. Testovali sme presnosť výstupného CCR modelu, HDF modelu a DDM modelu nad technológiou, ktorá zodpovedá konštantným výnosom z rozsahu. Celý postup testovania presnosti DEA modelov, pozostávajúci z troch hlavných blokov, ktorými sú generovanie dát, DEA analýza a vyhodnotenie presnosti sme naprogramovali v softvéri MATLAB. Zdrojové kódy sú uložené na stránke [20] v chránenom režime, autor práce je ochotný zdielať prístup k zdrojovým kódom na požiadanie. Na znázornenie analýzy výsledkov získaných z vykonaných experimentov sme využili softvér R, v ktorom sme vykresľovali obrázky uvedené v tejto kapitole.

V Tabuľke 4.1 sme zhrnuli parametre vystupujúce v procese generovania dát a ich zvolené hodnoty. Pomocou predstavených nastavení sme dokázali využitím Monte Carlo metódy simulovať 48 rôznych scenárov, pričom sme menili hodnoty nasledovných parametrov: počet vstupov, hodnota korelácie medzi vstupmi, parametre produkčnej funkcie a počet efektívnych fiktívnych útvarov. Ich kombináciami sme dokázali napodobniť rôzne technológie z reálneho sveta, napríklad také, ktoré majú vysokú koreláciu medzi jednotlivými vstupmi a rovnako dôležité vstupy, alebo naopak nízko korelované vstupy s rôznou dôležitosťou v produkčnom procese.

Parameter	**Označenie**	**Hodnota**
Počet DMU	n	100
Počet vstupov	m	3, 5
Interval pre vstupy	I	$[10, 1000]$
Korelácia medzi vstupmi	ρ	0, 0.4, 0.8
Dôležitosť vstupov	α_i	symetrická (48), asymetrická (49)
Substitučná distribúcia	σ_{ih}	rovnomerná (50), nerovnomerná (51)
Substitučná schopnosť	ϑ	podľa podmienky (53)
Počet efektívnych DMU^F	c	10, 20
Skutočná efektivita	$\hat{\theta}_j$	$\hat{\theta} \sim Beta(8, 2.5), j = c+1, \ldots, n$

Tabuľka 4.1: Tabuľka použitých parametrov



Obr. 4.1: Presnosť vybraných DEA modelov

Na Obr.4.1 je znázornený boxplot diagram vykonaných Monte Carlo simulácii. Červenými bodmi na okrajoch sú vyznačené minimálne a maximálne hodnoty B-Value v prípade skúmaných DEA modelov spomedzi všetkých simulácii. Stredný červený bod označuje priemernú hodnotu z realizovaných simulácii, t.j. priemer zo všetkých B-Value, ktorá vyjadruje výslednú presnosť DEA modelov. Čiary v strede obdĺžnikov označujú medián, to znamená, že pod aj nad čiarou sa nachádza 50% z realizovaných hodnôt B-Value. Pod jednotlivými obdĺžnikmi, ako aj nad nimi sa nachádza 25% z hodnôt B-Value. Malé čierne bodky na spodnej časti diagramu značia potenciálne odľahlé hodnoty z vykonaných simulácii. Simulačná štúdia ukazuje, že hyperbolický model HDF je najpresnejší spomedzi testovaných DEA modelov, vykazuje 91%-nú presnosť, preto by sme odporúčali použiť práve tento model na vykonanie DEA analýzy. Ostatné dva modely tiež vykazujú vysokú presnosť, a to model DDM 89%-nú presnosť, a výstupný CCR model 85%-nú presnosť.

	CCR-O	**HDF**	**DDM**
B-Rank	2.89	1.1	2.01

Tabuľka 4.2: Hodnota B-Rank pre vybrané DEA modely

Hodnoty doplnkového ukazovateľa B-Rank znázornené v Tabuľke 4.2 poukazujú tiež na vyššiu kvalitu HDF modelu, oproti ostatným skúmaným modelom.

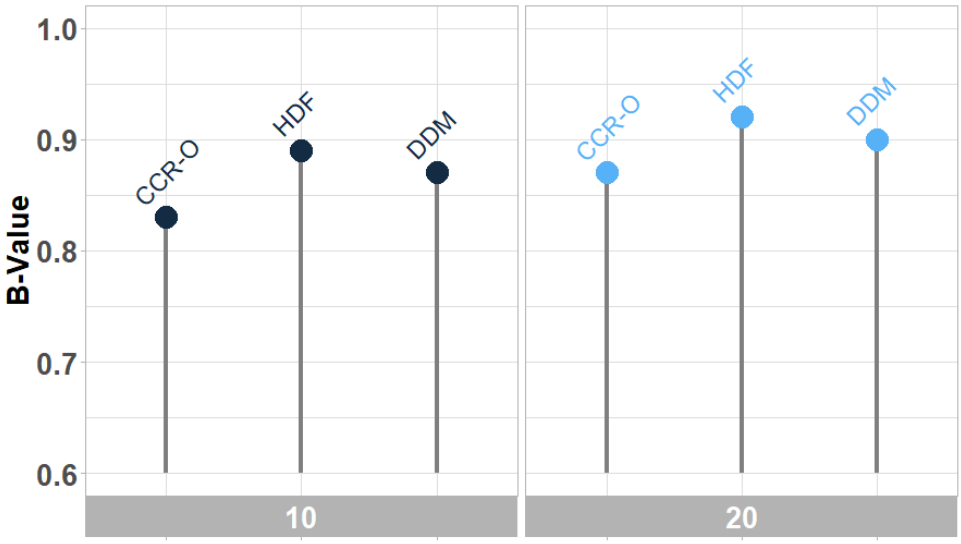
Ukazovateľ	CCR-O	HDF	DDM
MAE	0.938	0.96	0.953
(min/max)	(0.846/0.994)	(0.886/0.997)	(0.871/0.997)
SPEAR	0.883	0.939	0.924
(min/max)	(0.578/0.995)	(0.742/0.998)	(0.689/0.998)
INEFF	0.828	0.879	0.852
(min/max)	(0.5/1)	(0.563/1)	(0.5/1)
CORRI	0.746	0.85	0.817
(min/max)	(0.313/0.99)	(0.47/1)	(0.41/0.997)

Tabuľka 4.3: Charakteristika hodnôt použitých štatistických ukazovateľov

V tabuľke 4.3 sme zhrnuli priemerné, minimálne a maximálne hodnoty použitých štatistických ukazovateľov určených zo všetkých vykonaných Monte Carlo simulácii. Z priemerných hodnôt štatistických ukazovateľov vidíme, že model HDF vykazuje vyššiu kvalitu podľa každého štatistického ukazovateľa, ako ostatné dva skúmané DEA modely. Najvyššie hodnoty nadobúda štatistický ukazovateľ MAE, hodnoty blízke k hornej hranici 1, čo vyjadruje, že skúmané DEA modely vykazujú veľmi nízku priemernú absolútnu chybu poukazujúc na dostatočnú presnosť modelov. Najväčší rozdiel v kvalite skúmaných modelov môžeme vidieť pri štatistickom ukazovateľovi CORRI, kde model HDF dosiahol v priemere o 0.1 vyššiu hodnotu než model CCR-O. To znamená, že hodnoty efektivít získané aplikovaním modelu HDF sú bližšie k hodnotám skutočnej efektivity fiktívnych útvarov, ako hodnoty efektivít získané aplikovaním modelu CCR-O alebo modelu DDM.

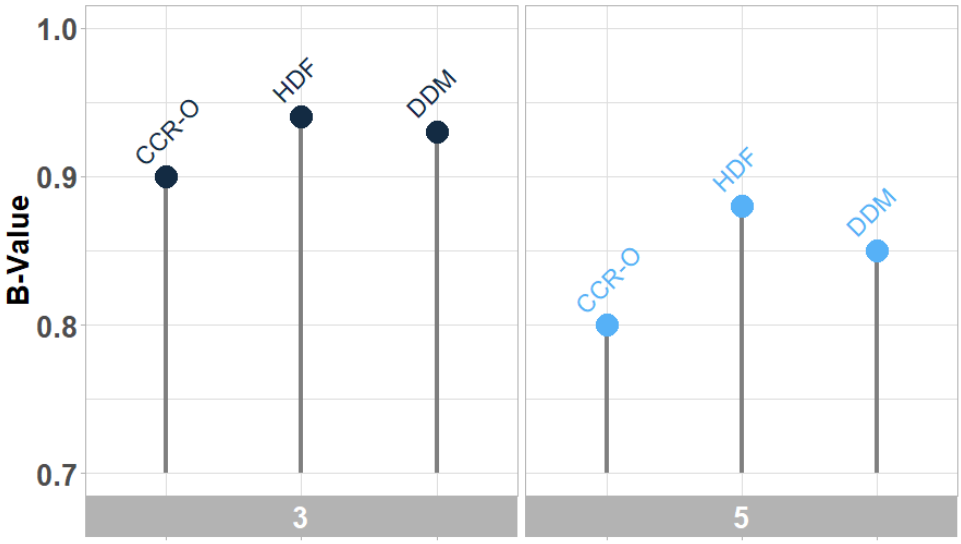
Vykonali sme aj analýzu citlivosti znázornenú pomocou obrázok 4.2-4.4. Sledovali sme, že ako sa zmení presnosť testovaných DEA modelov, ak zmeníme hodnotu vybraného parametra vystupujúceho v procese generovania dát. Analýzu citlivosti sme vykonali na nasledovných parametroch: počet efektívnych fiktívnych útvarov, počet vstupov a hodnota korelácie medzi jednotlivými vstupmi.

Na Obr.4.2 vidíme, že v prípade zvýšenia počtu efektívnych fiktívnych útvarov z 10 na 20, presnosť vybraných DEA modelov sa zvýši. Percentuálne zvýšenie presnosti je približne na úrovni 3%.

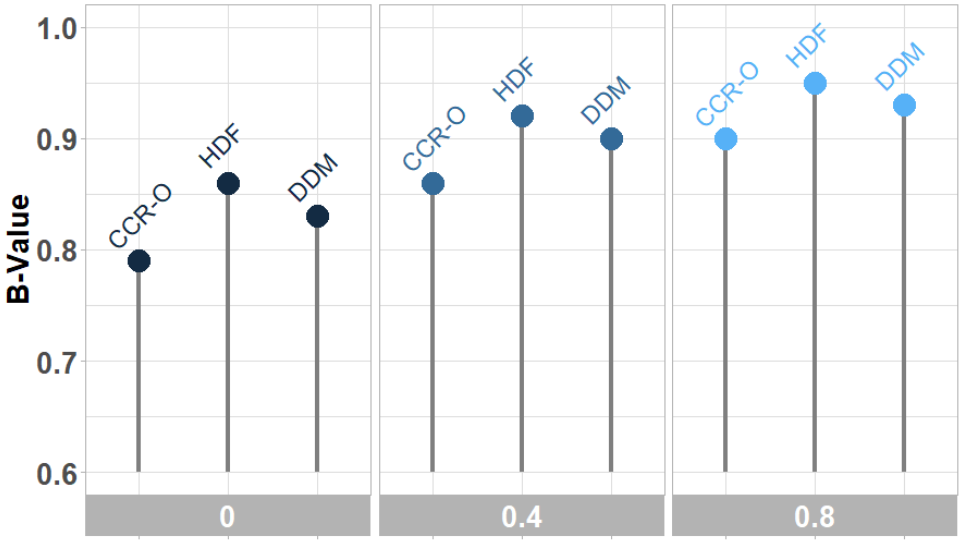


Obr. 4.2: Zmena v presnosti modelov (Efektívne útvary)

Na Obr.4.3 si môžeme všimnúť, čo sa stane, keď zvýšime počet vstupov z 3 na 5. Zvýšením počtu vstupov zvýšime aj dimenziu úloh matematického programovania príslušné vybraným DEA modelom. Vo vyššej dimenzie sa ťažšie dá aproximovať technologická množina T pomocou množiny M_{CRS}, a z toho dôvodu sa zníži presnosť jednotlivých DEA modelov. Percentuálne zníženie presnosti je v rozmedzí $6.1\% - 10.2\%$.



Obr. 4.3: Zmena v presnosti modelov (Vstupy)



Obr. 4.4: Zmena v presnosti modelov (Korelácia medzi vstupmi)

Na Obr.4.4 je znázornená zmena v presnosti skúmaných DEA modelov v závislosti od zmeny hodnoty korelácie medzi vstupmi. Pri vyšších hodnotách korelácie medzi vstupmi dostaneme štrukturovanejšiu skupinu útvarov, než pri voľbe nulovej korelácie, a práve preto zvýšením hodnoty korelácie sa zvyšuje presnosť skúmaných DEA modelov. Keď sa pozrieme na percentuálne zmeny, prvotné zvýšenie je na úrovni $5.6\% - 7\%$ a pri druhom zvýšení dochádza ešte k zvýšeniu presnosti modelov o približne $3.4\% - 4.5\%$.

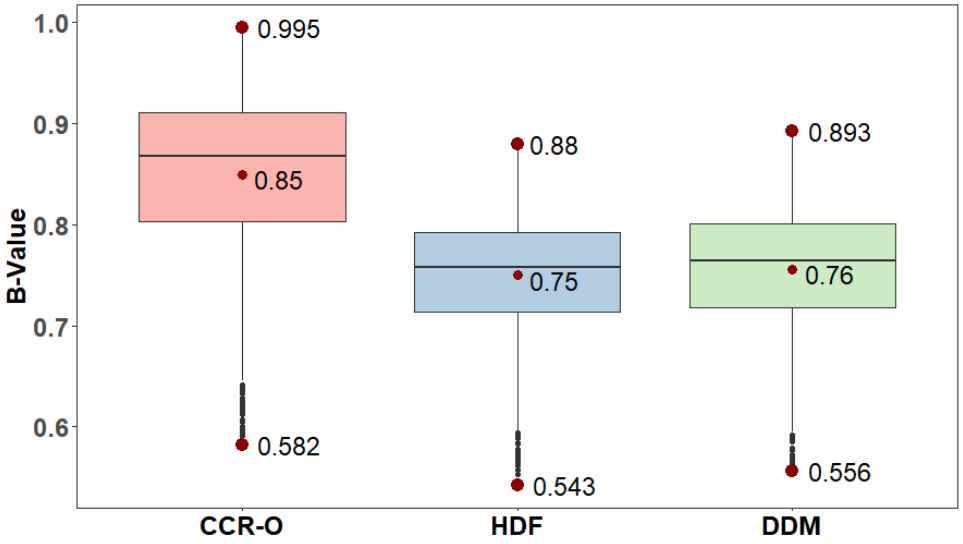
	CCR-O	**HDF**	**DDM**
$c : 10 \to 20$	3.7%	3%	3.5%
$m : 3 \to 5$	−10.2%	−6.1%	−7.3%
$\rho : 0 \to 0.4$	7%	5.6%	6.4%
$\rho : 0,4 \to 0.8$	4.5%	2.8%	3.4%

Tabuľka 4.4: Percentuálna zmena v presnosti DEA modelov

Keď sa pozrieme bližšie na tabuľku percentuálnych zmien v presnosti vybraných DEA modelov (viď. Tabuľku 4.4) môžeme si všimnúť, že model HDF nie je až taký citlivý na zmeny hodnôt parametrov, ako modely CCR-O a DDM.

Nakoniec sme ešte chceli porovnať hodnotu presnosti vybraných DEA modelov určenú použitím navrhnutej novej metodológie a štandardného prístupu vystupujúceho v

publikácii [10]. Rozdiel v dvoch prístupoch je, že v nami navrhnutej metodológie sme zohľadnili aj geometriu vybraných DEA modelov pri transformácii ideálnych útvarov na fiktívne, kým v štandardnom prístupe sa zohľadňuje iba výstupná efektivita jednotlivých útvarov vykonaním transformácie podľa schémy (54)-(55).



Obr. 4.5: Presnosť vybraných DEA modelov (štandardný prístup)

	CCR-O	**HDF**	**DDM**
B-Rank	1.48	2.5	2.02

Tabuľka 4.5: Hodnota B-Rank pre vybrané DEA modely (štandardný prístup)

Na Obr.4.5 sme znázornili boxplot diagram výsledkov Monte Carlo simulácii so štandardným prístupom. Vidíme, že ak sú zohľadnené iba výstupné efektivity, tak sa dostaneme k opačnému poradiu vybraných DEA modelov, čo sa týka ich presnosti. Model CCR-O vykazuje 85%-nú presnosť, model DDM 76%-nú a model HDF 75%-nú presnosť. Hodnoty doplnkového ukazovateľa B-Rank uvedené v Tabuľke 4.5 tiež poukazujú na to, že výstupne orientovaný CCR model má značne vyššiu kvalitu, než modely HDF a DDM.

V Tabuľke 4.6 sme opäť zhrnuli priemerné, minimálne a maximálne hodnoty štatistických ukazovateľov získaných z Monte Carlo simulácii so štandardným prístupom. Z priemerných hodnôt štatistických ukazovateľov vidíme, že CCR-O model vykazuje vyššiu kvalitu oproti ostatným skúmaným DEA modelom. Hlavne pri ukazovateľoch MAE

a CORRI nadobúdal značne vyššie hodnoty, než modely HDF a DDM, ale naopak pri ukazovateľovi INEFF nadobudol o trošku nižšiu hodnotu. Z minimálnych a maximálnych hodnôt si môžeme všimnúť, že pri štandardnom prístupe jednotlivé štatistické ukazovatele nadobúdajú hodnoty zo širších intervalov, než pri navrhnutej novej metodológie.

Ukazovateľ	**CCR-O**	**HDF**	**DDM**
MAE	0.939	0.874	0.88
(min/max)	(0.841/0.994)	(0.807/0.921)	(0.81/0.928)
SPEAR	0.884	0.883	0.883
(min/max)	(0.588/0.996)	(0.585/0.995)	(0.585/0.996)
INEFF	0.828	0.834	0.832
(min/max)	(0.389/1)	(0.49/1)	(0.495/1)
CORRI	0.749	0.409	0.429
(min/max)	(0.3/0.993)	(0.187/0.62)	(0.187/0.673)

Tabuľka 4.6: Charakteristika hodnôt použitých štatistických ukazovateľov (štandardný prístup)

Podľa vykonanej analýzy výsledkov si myslíme, že štandardný prístup k transformácii ideálnych útvarov na fiktívne útvary, zohľadňujúci výlučne výstupné efektivity jednotlivých útvarov zvýhodňuje výstupne orientovaný CCR-O model oproti neorientovaným modelom HDF a DDM. Navrhnutá nová metodológia zohľadňujúca geometriu vybraných DEA modelov, čiže vstupno-výstupné efektivity jednotlivých útvarov v prípade neorientovaných modelov HDF a DDM, a výstupnú efektivitu v prípade modelu CCR-O vedie ku korektnejšiemu porovnávaniu DEA modelov založených na parametrickej ceste.

Záver

Ciele našej diplomovej práce sa týkali spracovania problematiky testovania presnosti DEA modelov založených na parametrickej ceste. Po prvé, sme chceli naštudovať už existujúcu metodológiu testovania presnosti DEA modelov, ktorá bola predstavená v publikáciách [10, 19]. Potom, sme chceli navrhnúť efektívnejšiu metódu na generovanie experimentálnych dát a predstaviť alternatívnu metodológiu na testovanie presnosti DEA modelov, ktorá by umožnila korektnejšie porovnávanie modelov založených na parametrickej ceste. Cieľom bolo aj zdôvodniť, že navrhnuté nové prístupy sú vhodnejšie pre našu problematiku, ako tie, ktoré boli zatiaľ predstavené v literatúre.

V prvej kapitole práce sme sa zaoberali so základmi teórie mikroekonómie. Najprv sme definovali technologickú množinu a produkčnú funkciu, následne sme sa pozreli na vlastnosti produkčnej funkcie a definovali napríklad pojmy nazývané ako hraničný produkt a elasticita produkčnej funkcie. Vychádzali sme hlavne z učebného textu [2] a z publikácie [4]. Pokračovali sme predstavením známej produkčnej funkcie s názvom Cobbova-Douglasova produkčná funkcia. Následne v kapitole 1.1.2 sme popísali translogaritmickú funkciu, pričom vedomosti sme čerpali z článkov [1, 10, 16]. Translogaritmická funkcia bola zvolená na modelovanie produkčného procesu práve preto, lebo umožňuje modelovať dostatočne rôznorodé produkčné procesy. Najprv sme ju definovali, a predstavili jednotlivé parametre vystupujúce v predpise funkcie (11). Následne sme uviedli potrebné vzťahy, ako napríklad elasticity, hraničné produkty a vzťah pre výnosy z rozsahu vyjadrené pre translogaritmickú funkciu. Sformulovali sme Tvrdenia 1.1 a 1.3 o podmienkach kladených na parametre funkcie (11), pri ktorých táto funkcia spĺňa základné vlastnosti produkčnej funkcie a vykazuje konštantné výnosy z rozsahu, ktoré sme následne aj dokázali využitím Dôsledkov 1.2 a 1.4 [16, 17]. Z dôvodu lepšej prehľadnosti sme vyslovili Tvrdenie 1.5 o podmienkach kladených na parametre funkcie (11). V druhej podkapitole sme uviedli základy obálkovej analýzy dát a zaviedli známu symboliku, pričom sme vychádzali z učebného textu [7]. V tretej časti prvej kapitoly sme sa oboznámili s triedou DEA modelov založených na parametrickej ceste, ktorých presnosť sme sa rozhodli testovať. Sú to modely, ktoré sú známe vlastnosťou, že projekciu sledovaného útvaru na hranicu množiny produkčných možností hľadajú po rôzne definovaných parametrických cestách. Tvar parametrických ciest pomocou vybranej skupiny útvarov v priestore $\mathbb{R}^2$ sme znázornili na

Obr.1.6 pre vybrané DEA modely. V kapitole 1.3.1 sme zhrnuli vlastnosti skúmaných DEA modelov. Vedomosti o vybraných DEA modeloch sme čerpali z publikácie [8]. V poslednej podkapitole prvej kapitoly sme znázornili na Obr.1.7 schému testovania presnosti DEA modelov, ktorá sa skladá z troch blokov, a to generovanie dát, DEA analýza a vyhodnotenie presnosti [10].

V druhej kapitole sme dôkladne popísali proces generovania dát znázornený na Obr.2.1, pričom sme vychádzali hlavne z publikácie [10]. Po prvé sme predstavili, ako generovať množstvá jednotlivých vstupov pre DMU_j, a odvodili jednoduchú transformáciu pre maticu vstupov X na dosiahnutie hodnoty korelácie ρ medzi jej riadkami, t.j. jednotlivými vstupmi produkčného procesu. V ďalšej podkapitole sme predstavili možnosti nastavenia hodnôt parametrov translogaritmickej funkcie, a zdôvodnili sme, že nami navrhnutá voľba parametrov funkcie (11) vedie k numericky efektívnejšiemu modelovaniu produkčných procesov, ako prístup predstavený v článku [10]. Následne sme definovali výpočet ideálneho výstupu pre jednotlivé útvary. V tretej podkapitole sme predstavili dva možné prístupy ku generovaniu hodnoty skutočnej efektivity pre fiktívne útvary a zdôvodnili nami zvolený prístup. Nakoniec sme predstavili myšlienku novej transformácie generovaných ideálnych útvarov na fiktívne útvary, ktorá bola inšpirovaná geometriou DEA modelov založených na parametrickej ceste. Podľa nami navrhnutého nového prístupu pri vykonaní transformácie sa berie ohľad aj na zvolený DEA model, a tak sa zohľadňujú vstupno-výstupné efektivity jednotlivých útvarov v prípade neorientovaných DEA modelov, a vstupné resp. výstupné efektivity v prípade orientovaných DEA modelov. Na Obr.2.3-2.5 sme znázornili na príklade malej skupiny útvarov geometrický pohľad na navrhnutú novú transformáciu v prípade vybraných DEA modelov.

V tretej kapitole sme podrobne predstavili štatistické ukazovatele, ktoré sme používali na vyhodnotenie presnosti vybraných DEA modelov, pričom sme vychádzali z publikácie [10]. Následne sme poukázali na dôležitosť Monte Carlo simulácii v súvislosti s našou problematikou a predstavili ukončovacie kritérium definovaný v článku [10].

Vo štvrtej kapitole sme opísali vykonané experimenty a predstavili získané výsledky ohľadom testovania presnosti DEA modelov založených na parametrickej ceste. Experimenty sme vykonali v softvéri MATLAB, v ktorom sme naprogramovali generovanie dát, vybrané DEA modely ako aj vyhodnotenie presnosti [20]. Vykonali sme aj analýzu

citlivosti, sledovali sme zmenu v presnosti skúmaných DEA modelov pri zmene hodnoty vybraného parametra vystupujúceho v procese generovania dát. Porovnali sme navrhnutú novú metodológiu so štandardným prístupom ohľadom transformácie ideálnych útvarov na fiktívne útvary, a pozreli sme sa na dvoma spôsobmi vyrátané presnosti skúmaných DEA modelov.

Prínosmi našej práce považujeme prehľadné spracovanie problematiky testovania presnosti DEA modelov, využitie analytických vlastností translogaritmickej funkcie na korektnejšie a numericky efektívnejšie generovanie dát, a návrh novej metodológie umožňujúca korektnejšie porovnávanie DEA modelov založených na parametrickej ceste využitím geometrie týchto modelov. Možnými rozšíreniami tejto práce sú napríklad testovanie presnosti ďalších DEA modelov z vybranej triedy, alebo rozšírenie problematiky na technologické množiny s variabilnými výnosmi z rozsahu.

Zoznam použitej literatúry

[1] Boisvert, R.N.: *The translog production function: Its properties, its several interpretations and estimation problems*, Cornell University, New York, A.E.Res.82-28 (1982)

[2] Brunovský, P.: *Mikroekonómia*, učebné texty, FMFI UK, Bratislava, 2009, dostupné na internete (05.12.2023): http://www.iam.fmph.uniba.sk/skripta/brunovsky2

[3] Chambers, R.G., Chung, Y., Färe, R.: *Profit, Directional Distance Functions, and Nerlovian Efficiency*, Journal of Optimization Theory ans Applications 98 (1998), 351-364

[4] Coelli, T.J. et al.: *An Introduction to Efficiency and Productivity Analysis*, Second Edition, Springer New York (2005)

[5] Cooper, W.W. et al.: *BAM: a bounded adjusted measure of efficiency for use with bounded additive models*, Journal of Productivity Analysis 35 (2011), 85-94

[6] DeVille, L.: *Optimizing Gershgorin for symmetric matrices*, Linear Algebra and its Applications 577 (2019), 360-383

[7] Halická, M.: *DEA modely*, učebné texty, FMFI UK, Bratislava, 2015, dostupné na internete (03.01.2024): http://www.iam.fmph.uniba.sk/institute/halicka/teaching.xhtml

[8] Halická, M., Trnovská, M., Černý, A.: *A unified approach to radial, hyperbolic, and directional efficiency measurement in data envelopment analysis*, European Journal of Operational Research 312 (2024), 298-314

[9] Janková, K., Pázman, A.: *Pravdepodobnosť a štatistika*, Univerzita Komenského, Bratislava, 2011

[10] Kohl, S., Brunner, J.O.: *Benchmarking the benchmarks - Comparing the accuracy of Data Envelopment Analysis models in constant returns to scale settings*, European Journal of Operational Research 285 (2020), 1042-1057

[11] Krüger, J.J.: *A Monte Carlo study of old and new frontier methods for efficiency measurement*, European Journal of Operational Research 222 (2012), 137-148

[12] Ochotnický, P.: *Výber produkčnej funkcie pri odhade potenciálneho produktu*, Ekonomický časopis 56 (2008), č.8, 800-815

[13] Peleg, M.: *Beta distributions for particle size having a finite range and predetermined mode, mean or median*, Powder Technology 356 (2019), 790-794

[14] Perelman, S., Santín, D.: *How to generate regularly behaved production data? A Monte Carlo experimentation on DEA scale efficiency measurement*, European Journal of Operational Research 199 (2009), 303-310

[15] Tone, K.: *A slacks-based measure of efficiency in data envelopment analysis*, European Journal of Operational Research 130 (2001), 498-509

[16] Trnovská, M. et al.: *On (quasi) concavity of the trans-log production function and its implications for the accuracy analysis of the data envelopment analysis models*, working paper, FMFI UK, Bratislava, 2023

[17] Trnovská, M.: osobná konzultácia, FMFI UK, Bratislava, 2024

[18] Wang, J.-Y.: *Variance Reduction for Multivariate Monte Carlo Simulation*, The Journal of Derivatives 16 (2008), 7-28

[19] Zarrin, M., Brunner, J.O.: *Analyzing the accuracy of variable returns to scale data envelopment analysis models*, European Journal of Operational Research 308 (2023), 1286-1301

[20] *Zdrojové kódy*: https://github.com/BerniBB/TP_DEA