

ZÁPADOČESKÁ UNIVERZITA V PLZNI
FAKULTA ELEKTROTECHNICKÁ

Dizertační práce

POKROČILÉ METODY ANALÝZY A
OPTIMALIZACE MATEMATICKÝCH MODELŮ

Iveta PETRÁŠOVÁ

ANOTACE A KLÍČOVÁ SLOVA

Práce je zaměřena na řešení vybraných optimalizačních problémů s využitím metod pro aproximaci původního systému vhodným náhradním modelem, na kterém lze mnohdy optimalizaci provést rychleji a efektivněji. Vzhledem ke stoupající náročnosti počítačových simulací, které jsou v elektrotechnice při návrzích zařízení používány, roste zájem o využití různých metod, které umožňují redukovat výpočetní náročnost modelu. V dnešní době je stále větší pozornost věnována vhodným aproximačním modelům v návaznosti na vhodný návrh experimentu. Tyto metody zpravidla nacházejí uplatnění v prvotních fázích návrhu zařízení nebo u expertních systémů. Jejich řešení je provedeno jednak technikami implementovanými v komerčním programu, jednak vlastními algoritmy zabudovanými v programových balících Agros Suite a Artap, jež jsou dlouhodobě vyvíjeny na pracovišti katedry elektrotechniky a počítačového modelování.

KLÍČOVÁ SLOVA

Počítačové modelování, optimalizační úloha, návrh experimentu, náhradní model, neuronové sítě, strojové učení

ANNOTATION AND KEYWORDS

The work is focused on solving selected optimization problems using methods for approximation of the original system by a simpler equivalent model, on which optimization can often be performed faster and more efficiently. Due to the increasing complexity of computer simulations used in electrical engineering designs, there is a growing interest in the use of various methods that allow to reduce the computational complexity of the model. Nowadays, more and more attention is being paid to suitable approximation models in conjunction with appropriate experimental design. These methods usually find application in the early stages of plant design or in expert systems. Their solution is carried out both by techniques implemented in a commercial code and by own algorithms embedded in the Agros Suite and Artap software packages, which have been developed at the department.

KEYWORDS

Computer modeling, optimization methods, design of experiments, surrogate models, neural networks, machine learning

PODĚKOVÁNÍ

V první řadě bych ráda poděkovala prof. Ing. Pavlu Karbanovi, Ph.D. za cenné rady, připomínky a zároveň za ochotu a čas při vedení práce. Velké díky patří i prof. Ing. Ivovi Doleželovi, CSc., za odborné konzultace a pomoc při řešení problémů spojených se vznikem práce.

Plzeň 2021

Iveta Petrášová

SEZNAM SYMBOLŮ

FYZIKÁLNÍ POLE

φ	V	elektrický skalární potenciál
U	V	napětí
I	A	proud
R	Ω	rezistivita
L	H	indukčnost
E	$V \cdot m^{-1}$	intenzita elektrického pole
D	$C \cdot m^{-2}$	elektrická indukce
H	$A \cdot m^{-1}$	intenzita magnetického pole
B	T	magnetická indukce
J	$A \cdot m^{-2}$	proudová hustota
A	$Wb \cdot m^{-1}$	magnetický vektorový potenciál
E_{em}	N	elektromotorická síla
M	$A \cdot m^2$	elektromagnetický moment
P_{em}	W	výkon motoru
P_J	W	Jouleovy ztráty
P_{Fe}	W	ztráty v železe
P_m	W	mechanické ztráty
T	K, °C	teplota
R_t	$W \cdot K^{-1}$	tepelný odpor
ρ	$C \cdot m^{-3}$	objemová hustota náboje

MATERIÁLOVÉ PARAMETRY

ε	$F \cdot m^{-1}$	permitivita
μ	$H \cdot m^{-1}$	permeabilita
σ	$S \cdot m^{-1}$	elektrická vodivost
λ	$W \cdot m^{-1} \cdot K^{-1}$	tepelná vodivost
ρ	$kg \cdot m^{-3}$	měrná hmotnost
$\rho \cdot c_p$	$J \cdot m^{-3} \cdot K^{-1}$	měrné teplo

OSTATNÍ

t	s	čas
ω	$rad \cdot s^{-1}$	úhlová rychlost
m	kg	hmotnost
η	%	účinnost

SEZNAM ZKRATEK

OBEČNÉ ZKRATKY

FEM	Finite Element Method
SBO	Surrogate Based Optimization
DoE	Design of Experiments
MKP	metoda konečných prvků
PDR	parciální diferenciální rovnice
ODR	obyčejná diferenciální rovnice

ZKRATKY ALGORITMŮ A VÝPOČETNÍCH METOD

LHS	Latin Hypercube Sampling
EA	evoluční algoritmy
NSGA-II	Nondominated Sorting Genetic Algorithm II
AGEMOEA	Adaptive Geometry Estimation based MOEA
MOEA	Multi Objective Evolution Algorithm
GMM	Gaussian Mixture Models
EM	Expectation - Maximization
BOBYQA	Bound Optimization BY Quadratic Approximation
MoE	Mixture of Experts
MLP	Multi Layer Perceptron
RNN	Recurrent Neural Network
LSTM	Long Short-Term Memory
GRU	Gated Recurrent Unit
KPLSK	Kernel Partial Least Squares Regression by Kriging

OBSAH

i	ÚVOD DO PROBLEMATIKY	2
1	MOTIVACE	3
1.1	Členění práce	3
1.2	Vymezení řešeného problému	3
1.3	Optimalizace s využitím náhradního modelu	7
2	ZÁKLADNÍ POJMY	9
3	SOUČASNÝ STAV POZNÁNÍ	17
3.1	Vzorkování návrhového prostoru	17
3.2	Náhradní modely	19
3.3	Pokročilé optimalizační techniky	21
3.4	Optimalizační techniky s využitím náhradních modelů	22
4	HLAVNÍ CÍLE PRÁCE	24
ii	OPTIMALIZACE S VYUŽITÍM NÁHRADNÍCH MODELŮ	25
1	VZORKOVÁNÍ NÁVRHOVÉHO PROSTORU	26
1.1	Statické metody	27
1.1.1	Rovnoměrné rozdělení	27
1.1.2	Náhodné rozdělení	27
1.1.3	Faktoriální	28
1.1.4	Latin-Hypercube Sampling (LHS)	29
1.1.5	Haltonova sekvence	30
1.2	Adaptivní metody	31
2	NÁHRADNÍ MODELY	36
2.1	Strojové učení	36
2.2	Ověření na testovacích datech	36
2.3	Metody regrese a klasifikace	37
2.3.1	Polynomiální regrese	37
2.3.2	Gaussovské procesy	38
2.3.3	Hluboké neuronové sítě	40
2.3.4	Rozhodovací stromy	44
3	OPTIMALIZAČNÍ METODY	47
3.1	Definice optimalizační úlohy	47
3.2	Klasifikace optimalizačních úloh	50
3.3	Metody optimalizace	52
3.3.1	Metody využívající informace o změně funkce	55
3.3.2	Stochastické metody	63
3.3.3	Ostatní	65
iii	SOFTWAREOVÁ IMPLEMENTACE	70
1	ARTAP	72
2	AGROS SUITE	74
3	COMSOL MULTIPHYSICS	76

iv	TESTOVACÍ PŘÍKLADY S ANALYTICKÝM ŘEŠENÍM	78
1	ANALYTICKÝ MODEL ODDĚLOVACÍHO TRANSFORMÁTORU	80
1.1	Analytický model	80
1.2	Ilustrativní příklad	82
1.3	Optimalizace	83
1.3.1	Jednokriteriální optimalizace	83
1.3.2	Vícekriteriální optimalizace	84
1.4	Numerické řešení	84
2	ANALYTICKÝ MODEL STEJNOSMĚRNÉHO MOTORU S PM	87
2.1	Analytický model	87
2.2	Ilustrativní příklad	89
2.3	Optimalizace	89
2.3.1	Vzorkování návrhového prostoru	90
2.3.2	Jednokriteriální optimalizace	91
2.4	Numerické řešení	93
3	ANALYTICKÝ MODEL SÉRIOVÉHO RL OBVODU	99
3.1	Analytický model	99
3.2	Trénování náhradního modelu	100
3.3	Predikce proudu pro zvolené napětí zdroje	102
v	ILUSTRATIVNÍ PŘÍKLADY	105
1	LAMELOVÝ PŘÍVOD	106
1.1	Matematický model a simulace úlohy	107
1.2	Optimalizace plného modelu	108
1.2.1	Monokriteriální optimalizace	108
1.2.2	Vícekriteriální optimalizace	108
1.3	Optimalizace s využitím náhradního modelu	110
2	LASEROVÝ OHŘEV	115
2.1	Matematický model	116
2.1.1	Teplotní pole produkované laserovým paprskem	116
2.1.2	Keyhole efekt	117
2.1.3	Nárůst objemu materiálu	117
2.2	Počítačový model	117
2.2.1	Laserové svařování	118
2.2.2	Laserová depozice kovů	120
2.3	Optimalizace	121
2.3.1	Laserové svařování	121
2.3.2	Laserová depozice kovů	124
3	VYPALOVÁNÍ NEVODIVÝCH VRSTEV	127
3.1	Matematický model	128
3.2	Numerické řešení	129
3.3	Náhradní model	129
3.4	Řízení procesu pomocí PID regulátoru	132
vi	ZÁVĚR A VÝHLED DO BUDOUCNA	135
1	ZÁVĚR A DALŠÍ VÝVOJ	136

Část I

ÚVOD DO PROBLEMATIKY

MOTIVACE

ZNALOST fyzikálních polí patří v současnosti k nezbytným předpokladům pro návrh zařízení nejen v elektrotechnice. Jejich modelování a simulace se využívají v širokém spektru vědních disciplin počínaje klasickou technikou, přes astronomii až po molekulární biologii. Matematické modely popisující mnohé fyzikální jevy jsou již vesměs v dostupné literatuře velmi dobře popsány a vhodné numerické algoritmy a jejich řešení jsou zabudovány v komerčních i volně šiřitelných programových balících. V současné době nejpoužívanější metoda konečných prvků se stala součástí takových kódů, jako je například COMSOL Multiphysics, ANSYS nebo CST Suite. Mezi volně šiřitelnými softwarovými produkty, které ji rovněž využívají, se nacházejí například knihovny deal.II, FEniCS, MFEM, Hermes, FEM a programy FEMM a Agros Suite, který je vyvíjen na katedře elektrotechniky a počítačového modelování Fakulty elektrotechnické ZČU již více než 10 let. S jejím využitím lze úspěšně vyřešit poměrně širokou škálu úloh z oblasti elektromagnetického pole, teplotního pole, strukturální mechaniky nebo akustického pole.

1.1 ČLENĚNÍ PRÁCE

Samotná práce je rozčleněna do pěti hlavních částí, kdy v první části jsou stručně vysvětleny základní pojmy návrhu experimentu, náhradních modelů a optimalizačních úloh, které jsou podrobněji popsány v druhé teoretické části. Dále je uvedeno pojednání o současném stavu poznání, z něhož vyplývají hlavní cíle práce. Třetí část práce se věnuje použitému softwaru. Daná problematika je prezentována na testovacích příkladech se známým analytickým řešením ve čtvrté části a následně je pak implementována na složitějších ilustrativních příkladech z oblasti elektrotechniky v závěrečné páté části.

1.2 VYMEZENÍ ŘEŠENÉHO PROBLÉMU

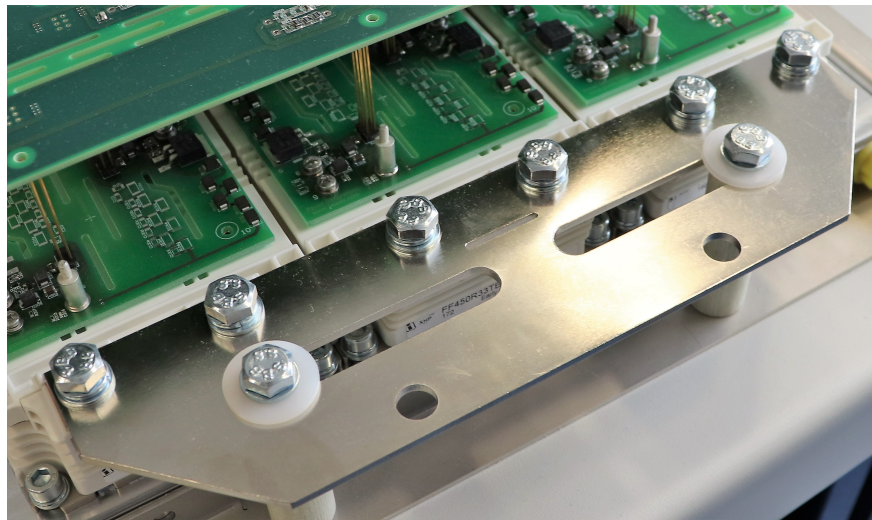
V současnosti se stále častěji nahrazují časově náročné a drahé fyzikální experimenty rychlejšími počítačovými simulacemi, které umožňují daleko širší experimentování s chováním systému a jeho reakce na různé kombinace vnějších (vstupních) faktorů. Vzhledem ke stoupajícím nárokům na simulace stále složitějších fyzikálních problémů stoupá i snaha o jejich co nejefektivnější řešení.

K zajištění požadovaných vlastností systému je možné optimalizační proces neřešit na plném a výpočetně náročném FEM modelu, ale využít k tomuto náhradní modely, jejichž řešení se přibližuje cílové funkci plného modelu a jsou sestaveny z dat, která jsou získána vzorkováním chování původního systému. Tyto modely jsou sestavené na základě chování plného modelu v několika vzorových bodech návrhového prostoru. Naznačená problematika bude ilustrována na několika typických příkladech.

Problém ohřevu přívodních lamel frekvenčního měniče

Pro názornost bude uveden příklad návrhu konkrétního elektrotechnického zařízení, z něhož vyplyne i motivace této práce. Jedná se o ohřev přívodních lamel frekvenčního měniče (viz obr. 1), jejichž provozní teplota nesmí přesáhnout předepsanou mez (ta bývá u mědi obvykle 150 °C).

Průtokem elektrického proudu se přívodní lamely měniče nerovnoměrně ohřívají Jouleovými ztrátami. Vzhledem k tomu, že fyzikální parametry materiálu lamel (elektrická vodivost, tepelná vodivost, specifické teplo a hustota) se mění s teplotou, je zřejmé, že proudové i teplotní pole se zde navzájem ovlivňují. To teoreticky znamená, že v každém časovém kroku by se měly příslušné změny zmíněných materiálových parametrů respektovat, což by však (uspořádání je trojrozměrné) vyžadovalo nepřiměřeně dlouhou dobu výpočtu.



Obr. 1: Lamelový přívod frekvenčního měniče

V podobných případech se proto akceptuje zjednodušení spočívající v předpokladu, že se v definovaných teplotních intervalech tyto parametry nemění (mění se tedy skokově od jednoho teplotního intervalu k dalšímu). Tento předpoklad je velmi výhodný, poněvadž se doba výpočtu zkrátí až o 90 %, přičemž vznikající chyba je většinou na úrovni několik promile a jen zřídka přesáhne 2–3 %. Přesnost výsledků tedy zůstane za tohoto zjednodušení zcela přijatelná.

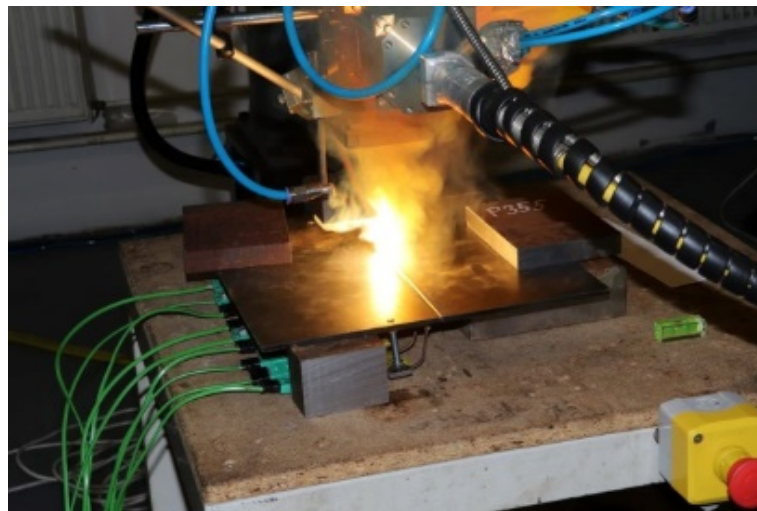
Tím je vyřešena tzv. dopředná část úlohy, kdy jsou na základě znalosti vstupních údajů stanoveny výstupní parametry zařízení a jeho charakteristiky, v tomto případě rozložení ustálené (jsou-li napájecí proudy stejnosměrné) nebo kvaziustálené (jsou-li tyto proudy periodické) teploty v lamelách. Může se ovšem snadno stát, že lokálně tato teplota překročí povolenou hodnotu. V takovém případě je žádoucí tvar lamely a dílčí přívody upravit tak, aby v žádném místě kritická teplota překročena nebyla. Tato úloha se nazývá zpětná a řeší se vhodnou optimalizační technikou.

Uvedená technika (zpravidla se jedná o iterační proces) povolí v zadaném rozsahu vstupních parametrů nalézt jedno nebo více optimálních či suboptimálních řešení, která vyhovují

zadaným požadavkům. Z nich lze zvolit definitivní uspořádání lamel přívodu, jež může být ovlivněno i na základě doplňkových kritérií (například množství materiálu či jeho cena). Během iteračního procesu lze navíc testovat vliv nejistot vstupních parametrů, robustnost řešení, využití náhradních modelů atd.

Identifikace materiálových parametrů pro vybrané elektrotepelné problémy

Jedním ze závažných problémů ovlivňujících přesnost numerických modelů různých procesů elektroohřevu jsou nejistoty ve vlastnostech příslušných kovových materiálů závislých na teplotě [1]. Tyto vlastnosti jsou dobře známy pro čisté kovy, ale v případě slitin jsou často známé pouze přibližně nebo dokonce neznámé. Obecně závisí na jejich chemickém složení či homogenitě jejich struktur. Dokonce i stejné slitiny vyrobené různými výrobci mohou vykazovat nestejně parametry. Další parametry, jako je zvolená metoda, diskretizace atd., lze během simulačního procesu poměrně snadno upravit. V případě nejistot v materiálových charakteristikách nelze jednoznačně určit přesnost výsledného modelu a srovnání reality s její simulací by pak nedávalo velký smysl.

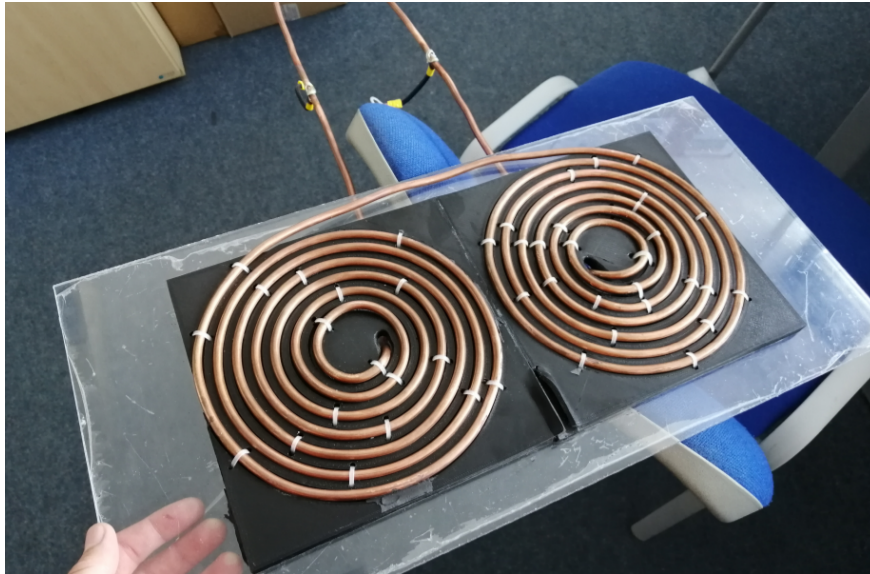


Obr. 2: Příklad laserového svařování

V případě modelování laserového svařování (viz obr. 2) a laserové depozice jde o poměrně komplikovanou problematiku s velkým množstvím úskalí. Při vytváření modelu pro svařování je nutné počítat s parametry oblasti svaru, která je charakterizována určitou hloubkou, šířkou a mikrostrukturou. Je také nutné určit rychlost posuvu svařovaných těles a výkon dodávaný do ozářeného místa laserovým paprskem. Při laserové depozici je zase nutné počítat s navýšením objemu materiálu během procesu. Optimální nastavení těchto veličin představuje velmi komplikovanou inverzní úlohu. Plný model, který by aproximoval realitu s dostatečnou přesností, by zahrnoval interakci mezi elektromagnetickým polem, teplotním polem, fázovými změnami (nejprve změna z pevného skupenství na kapalné a poté částečně i na plynné), polem rychlosti proudění roztaveného kovu, atd. 3D řešení problému by bylo samo o sobě velmi časově náročné a optimalizační proces takto komplikovaného modelu by byl v podstatě neproveditelný. V tomto případě je tedy nutné najít dostatečně přesný zjednodušený model (založený buď na redukcí řádu modelu, nebo na principu technik náhradních modelů), který bude zpracován o několik řádů rychleji, viz [IF.2].

Vypalování tenkých elektricky nevodivých vrstev nepřímým indukčním ohřevem

Poslední ilustrativní problém je zaměřen na vypalování tenkých elektricky nevodivých vrstev. Proces vypalování, jehož účelem je chránit kovový povrch před nepříznivými vlivy prostředí, je realizován nepřímým indukčním ohřevem kovového substrátu, který předává teplo tenké vrstvě vedením. Reálné řešení experimentu je zobrazeno na obrázku 3.



Obr. 3: Řešený problém ohřevu vypalování

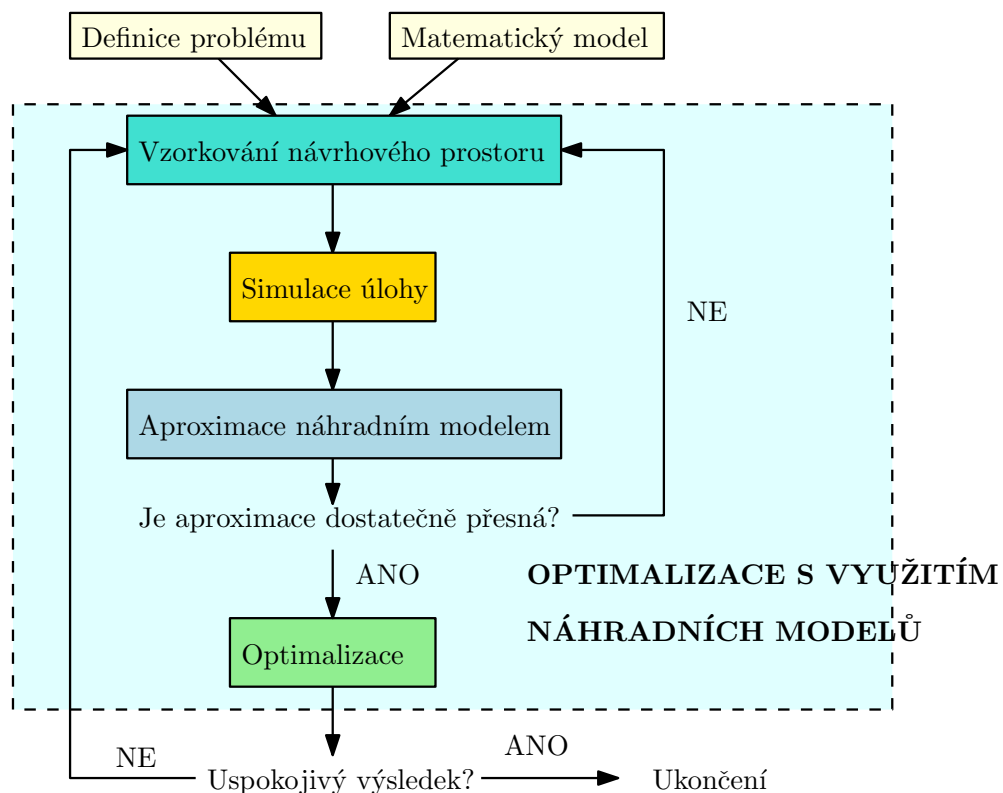
Jako ochranné prostředky se velmi často používají barvy, laky a podobné tenké nátěry. Kvalitní barvy s dlouhou životností se obvykle vypalují po nanesení tenké vrstvy, tj. pokrytý povrch se zahřeje na definovanou teplotu, kterou je třeba po předepsanou dobu udržovat.

Vzhledem k této podmínce je kontinuální ohřev substrátu nevhodný. Proces je rozdělen na interval, kdy je napájení zapnuto na čas Δt_{on} a vypnuto na čas Δt_{off} , přičemž délka intervalů je řízena s ohledem na požadovanou teplotu vypalování. Požadavky na dobu vypalování pro udržení konkrétní teploty udává výrobce barvy a úkolem je navrhnout parametry ohřevu - amplitudu proudu, frekvenci a dobu spínání [2].

U problémů tohoto typu často není možné přímo měřit teplotu v celé ohřívané ploše, ale pouze v jejích okrajových oblastech. Teplotu v nedostupných místech je pak nutné zjistit z modelu. Řešení plného modelu je však časově velmi náročné a nelze jej provádět v reálném čase. Vzhledem k nutnosti velmi jemné diskretizace výpočetní sítě ochranné vrstvy (aby konečné prvky byly geometricky souměřitelné) narůstá počet stupňů volnosti na několik milionů a výpočet jedné varianty pak trvá i na velmi dobrém počítači několik hodin. Standardní on-line řízení procesu je prakticky nemožné, proto je pro spolehlivé řízení procesu nutno zvolit jinou alternativu, například použití digitálního dvojčete, kdy je možné model nahradit vhodnou evoluční neuronovou sítí a na základě sledování teploty v několika měřených bodech vyhodnocovat její rozložení v celé oblasti. Tento přístup by mohl výrazně zpřesnit realistický odhad teploty a tedy i kvalitu procesu.

1.3 OPTIMALIZACE S VYUŽITÍM NÁHRADNÍHO MODELU

Při řešení složitějších multifyzikálních prostorových problémů je tedy nutné počítat s tím, že již samotné numerické řešení dopředné úlohy bude časově velice náročnou záležitostí, mnohdy dokonce jen stěží proveditelnou, a problematiku optimalizace pak nebude možné řešit vůbec. Takovou úlohu je potom zapotřebí zjednodušit, jak je to jen možné, avšak výsledky dopředného řešení musí ležet v předepsané toleranci. V dnešní době se za tímto účelem dostávají do popředí techniky návrhu experimentu, které lze velmi dobře využít s náhradními modely a následně i s vhodnými optimalizačními algoritmy.



Obr. 4: Princip optimalizace s využitím náhradního modelu

Hlavní motivací této práce je využití standardních technik numerického výpočtu pro realizaci řešení složitějších optimalizačních problémů spolu s aplikací náhradních modelů. Tyto postupy doposud ještě nejsou ve spojení s počítačovým modelováním v elektrotechnice obvyklé, přičemž se ale jeví pro budoucnost oboru jako vysoce perspektivní. V zahraniční literatuře se označují jako Surrogate-based optimization—SBO [3].

K řešení SBO využíváme kromě matematického modelu a plně numerické simulace

1. metody návrhů experimentu,
2. náhradní modely
3. a klasické optimalizační metody.

Princip celého procesu je znázorněn na obr. 4. Nejprve je nutno aproximovat plný model daného problému vhodným modelem náhradním. Za tím účelem je nejprve zapotřebí shromáždit dostatek vstupních dat pro nácvič náhradního modelu, k čemuž jsou hojně využí-

vány metody pro návrh experimentu [4]. Jakmile jsou vytvořeny náhradní modely, pak lze k hledání nového optimálního návrhu využít dostupné optimalizační algoritmy. V praxi se k tomuto často využívají například algoritmy genetické, kterým bude věnována kapitola 2.3.4. Pokud je vytvořen náhradní model, je jeho předpověď vzhledem k časové a výpočetní náročnosti mnohem efektivnější než zdlouhavé a náročné výpočetní numerické simulace [5, 6]. Výpočty s náhradními modely tedy představují poměrně efektivní techniku, která je schopná využít vzorkovaná data k vytvoření alternativní reality na základě dostatečně dobré predikce tak, aby byla schopna alternovat drahé a výpočetně náročné numerické simulace na základě určení chování modelu v několika vzorových bodech návrhového prostoru. Pro vytvoření dostatečně přesného náhradního modelu je zapotřebí provést

1. výběr vzorových bodů pro simulaci odezev systému,
2. návrh metody vytvoření náhradního modelu
3. a vyhodnocení přesnosti náhradního modelu.

Pokud není známo dostatečné množství vzorků, je potřeba začít s počátečním návrhem experimentu, natrénovat znovu náhradní model a vyhodnotit jeho přesnost, kterou můžeme zlepšit například přidáním dalších vzorků do cvičné sady dat a celý postup následně opakovat. Tato technika s využitím náhradních modelů bude nyní prezentována na jednoduchém ilustrativním příkladu včetně vysvětlení základních pojmů.

ZÁKLADNÍ POJMY

Základní princip pro optimalizační problémy, jejichž řešení je založeno na aproximační realitě, bude uveden v následující kapitole. Zároveň jsou v této kapitole stručně vysvětleny základní pojmy (podrobněji v teoretické části ii), na které je tato práce zaměřena.

Matematický model a simulace úlohy

Pro simulaci úlohy je nejprve potřeba popsat chování systému, viz obr. 4, tzn. vyřešit korektní matematický model. Matematický model aproximuje realitu souborem rovnic a vztahů s určitými parametry, jež jsou zformulovány na dané definiční oblasti.

1. Vzorkování návrhového prostoru

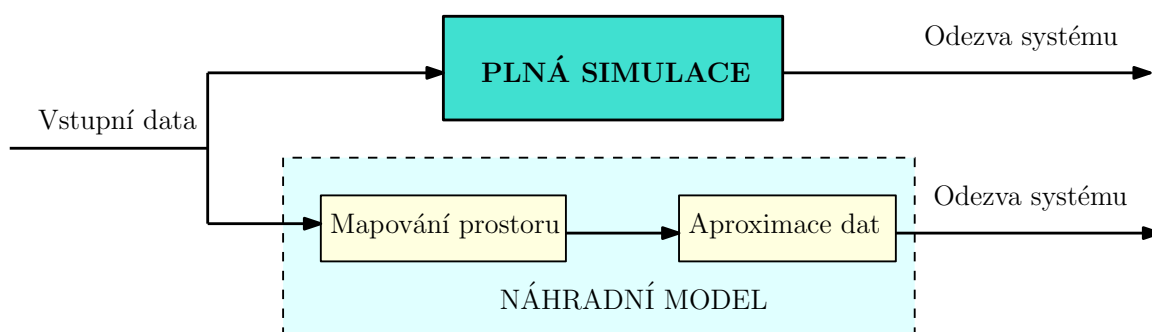
Pro sestavení náhradního modelu je nutné znát reakci systému nebo modelu v několika vzorových bodech pomocí simulací či experimentů. Pro návrh experimentu je nutno vybrat vhodnou metodu k navzorkování návrhového prostoru. Jedná se o postup s cílem maximalizovat dané množství informací získaných z omezeného počtu **vzorových bodů** umístěných v náhradním návrhovém prostoru [5]. **Návrhový prostor** je prostorem, který zahrnuje všechny dostupné hodnoty a kombinace vstupních parametrů, které poskytují odpovídající řešení v rámci své specifikace.

V současnosti existuje poměrně velké množství metod pro návrh experimentu (v zahraniční literatuře označován jako DoE - **D**esign of **E**xperiments). Obecně je lze rozdělit na

- **klasické** - zaměřené především na reálné fyzické experimenty a
- **moderní** - charakterizované využitím dalších technik a prací s daty na vstupu i na výstupu. Lze je dělit na metody
 - statické, jež pracují na základě vyplňování vstupního prostoru v daném rozsahu dle specifikace jednotlivých metod
 - a adaptivní, které kromě kvality vstupních dat vyhodnocují i kvalitu dat na výstupu.

2. Náhradní model

Náhradní model (metamodel, aproximační model) je empirický výraz (analytický nebo numerický), který kvantifikuje vztahy mezi nejdůležitějšími vstupy a výstupy systému. Jinými slovy se jedná o náhradu složité reality (komplikované numerické modely, reálné fyzické systémy, apod.) výpočetně méně náročným modelem.



Obr. 5: Princip náhradních modelů

Pokud se očekává, že výsledek bude drahý, časově náročný, či jinak obtížně měřitelný, lze místo něj použít právě vhodný náhradní model. Účelem tohoto modelu je co nejpřesněji předpovědět vlastnosti původní reality (tzv. podkladového modelu) při daleko nižších nákladech.

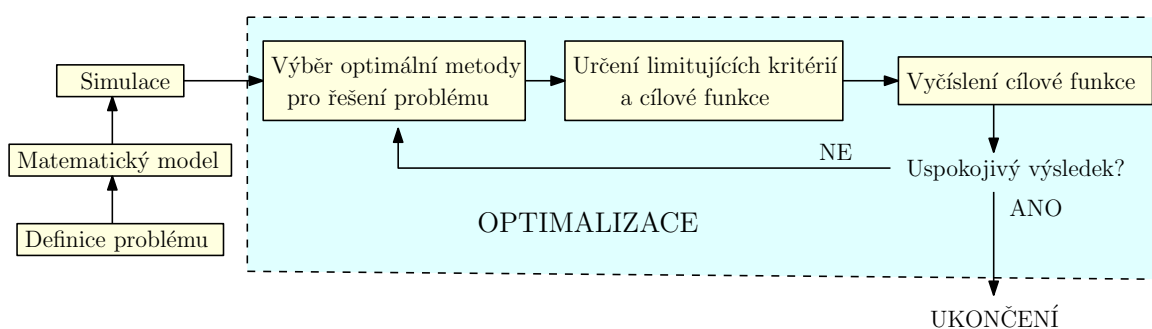
Pro tvorbu náhradního modelu je zásadní dle [7]

- znalost vlastností reality v několika vzorových bodech návrhového prostoru,
- forma náhradního modelu - matematický popis vztahů mezi vstupy a výstupy,
- metoda náhradního modelu - postup pro odvození nejlepších parametrů pro vzorkovaná vstupní a výstupní data,
- vyhodnocení přesnosti náhradního modelu.

Princip aproximace náhradním modelem lze rozdělit do několika základních kroků, což je velmi dobře patrné z obr. 5.

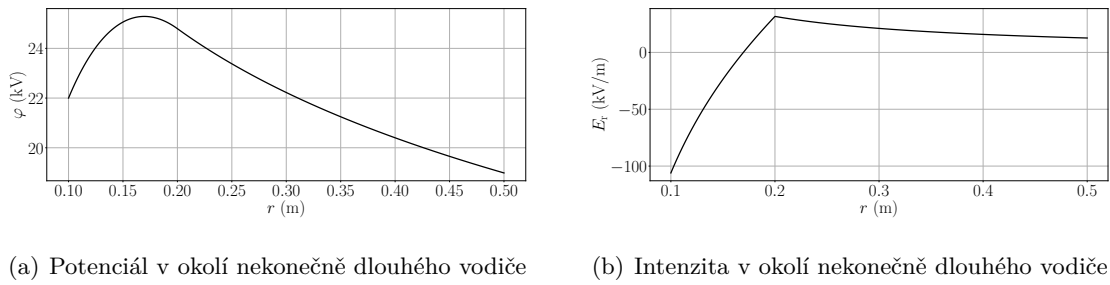
3. Optimalizace

Optimalizační proces je snahou o nalezení co nejuspěšnějšího řešení (extrému) daného problému za předem daných omezujících kritérií, k čemuž využívá soubor technik, metod, algoritmů a dalších procedur, které byly k tomuto účelu vyvinuty.



Obr. 6: Princip optimalizačního procesu

Předpokládejme tedy optimalizační problém, jenž má řešení $OF(x)$. Toto řešení nazýváme *cílovou (kritériální) funkcí*.



Obr. 8: Potenciál a intenzita elektrického pole v okolí nekonečně dlouhého vodiče

Jednorozměrné obecné řešení této rovnice v radiálním směru lze zapsat jako

$$\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial \varphi}{\partial r} \right) = -\frac{\rho}{\varepsilon}.$$

Po integraci získáme obecné řešení pro potenciál ve tvaru

$$\varphi = -\frac{\rho r^2}{4\varepsilon} + k_1 \ln r + k_2, \quad (1)$$

kde k_1 a k_2 jsou konstanty. Výsledné řešení potenciálu je řešením ve dvou podoblastech, kdy pro první z nich lze psát

$$\varphi_1 = -\frac{\rho r^2}{4\varepsilon_0} + k_{11} \ln r + k_{12}$$

a pro druhou platí

$$\varphi_2 = k_{21} \ln r + k_{22}.$$

Intenzitu elektrického pole lze v jednotlivých prostředích vyjádřit vztahy

$$E_{r,1} = \frac{\partial \varphi_1}{\partial r} = -\frac{\rho r}{2\varepsilon_0} + \frac{k_{11}}{r}$$

a

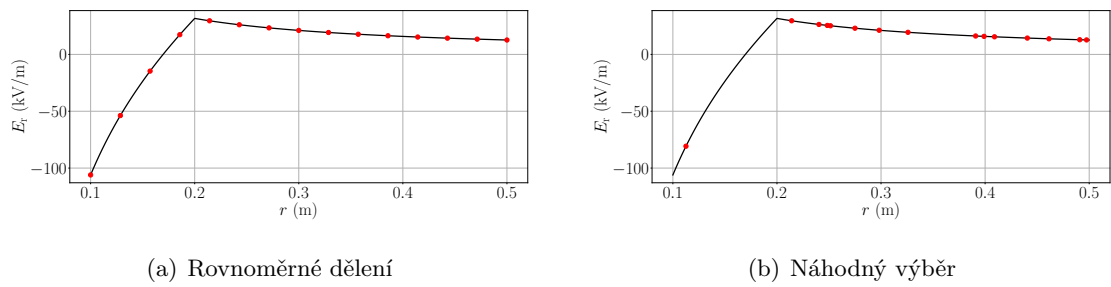
$$E_{r,2} = \frac{\partial \varphi_2}{\partial r} = \frac{k_{21}}{r}.$$

Integrační konstanty k_{11} , k_{12} , k_{21} a k_{22} se získají z okrajové podmínky, podmínky spojitosti potenciálu a podmínky na rozhraní. Po dosazení získáme soustavu rovnic ve tvaru

$$\begin{bmatrix} \ln a & 1 & 0 & 0 \\ 0 & 0 & \ln c & 1 \\ -\ln b & -1 & \ln b & 1 \\ -\frac{1}{b} & 0 & \frac{1}{b} & 0 \end{bmatrix} \cdot \begin{bmatrix} k_{11} \\ k_{12} \\ k_{21} \\ k_{22} \end{bmatrix} = \begin{bmatrix} U + \frac{\rho a^2}{4\varepsilon_0} \\ 0 \\ \frac{\rho b^2}{4\varepsilon_0} \\ \frac{\rho b}{2\varepsilon_0} \end{bmatrix}. \quad (2)$$

Konstanty k_{11} , k_{12} , k_{21} a k_{22} jsou pak řešením této soustavy.

Průběh elektrického potenciálu v závislosti na poloměru je znázorněn na obr. 8 a). Odpovídající graf závislosti intenzity elektrického pole je vidět pak na obr. 8 b).

Obr. 9: Použité metody návrhu experimentu, počet vzorových bodů $N = 15$

1. Vzorkování návrhového prostoru

Ukázka využití jednotlivých metod pro vzorkování prostoru pro výsledné hodnoty elektrické intenzity je zobrazena na obr. 9. Jak je z obrázku patrné, v tomto případě je nejvhodnější rovnoměrné rozdělení, viz obr. 9 a). Naproti tomu jako naprosto nevhodný se jeví náhodný výběr, viz obr. 9 b). Vzorové body mají mezi sebou jen velmi malý rozptyl, takže náhradní model sice poměrně přesně aproximuje hustě navzorkovanou oblast, avšak na úkor oblastí ostatních, kde jsou viditelné neakceptovatelné odchylky.

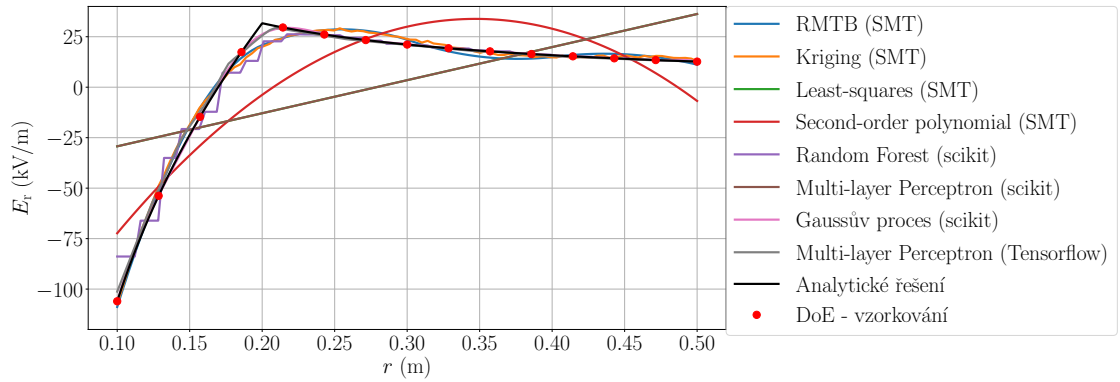
2. Náhradní model

Na obr. 10 jsou znázorněny možnosti aproximace návrhového prostoru předchozího ilustrativního příkladu pomocí jednotlivých metod z několika volně dostupných knihoven. Na obr. 10 a) je znázorněno proložení rovnoměrně navzorkovaného prostoru a na obr. 10 b) náhodně navzorkovaného prostoru. Pro každé jednotlivé proložení jsou pak na obr. 11 znázorněny odchylky jednotlivých metod původního analytického řešení intenzity elektrického pole. Z obrázku je patrné, jak je důležitá vhodná volba bodů v prostoru. U reálných vysokodimenzionálních problémů se používá kromě rovnoměrného dělení také vzorkování s využitím Latin Hypercube Sampling nebo Haltonovy sekvence.

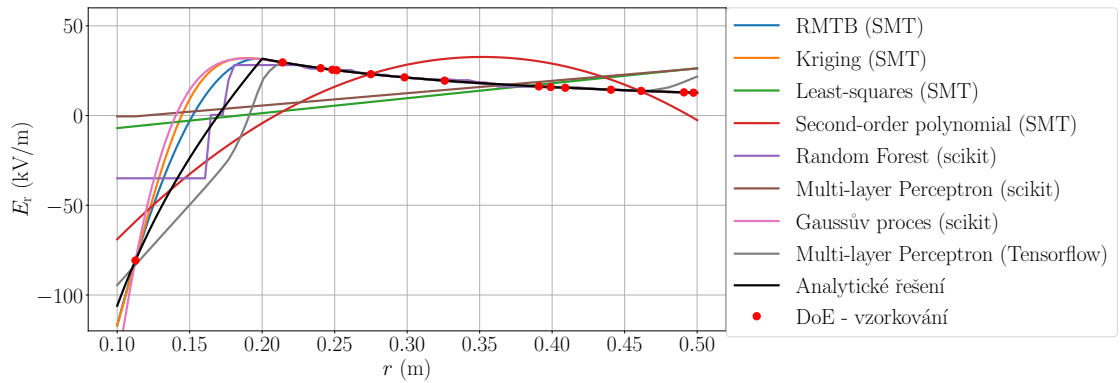
V tabulce 1 je uvedena průměrná odchylka intenzity elektrického pole náhradního modelu od analytického řešení ve tvaru

$$E_{\text{odch}} = \sqrt{\frac{1}{N} \sum_{i=1}^N (E - E_a)^2}.$$

kde E_a je analytické řešení a E představuje intenzitu pole z náhradního modelu. Z tabulky je zřejmé, že nevhodnější metodou proložení je Gaussův proces a hluboká neuronová síť.



(a) Aproximace pro rovnoměrně navzorkovaný prostor



(b) Aproximace pro náhodně navzorkovaný prostor

Obr. 10: Aproximace navzorkovaného prostoru

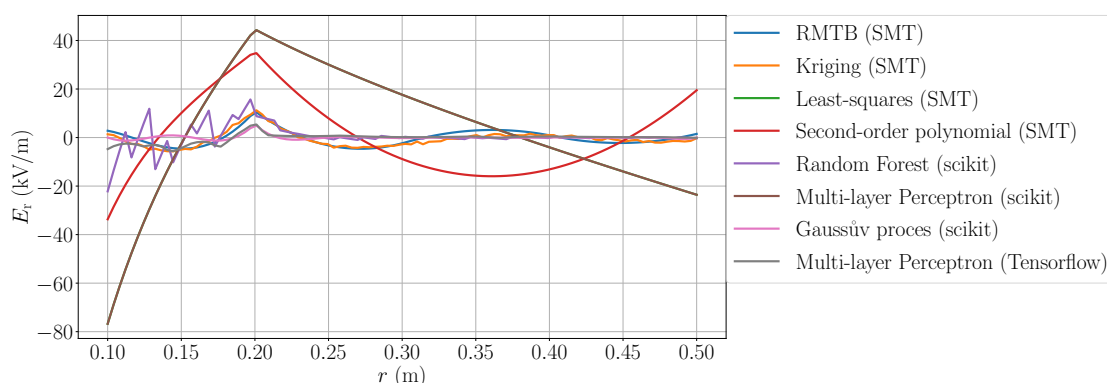
Tab. 1: Odchylka intenzity pole náhradního modelu od analytického řešení

Metoda	rovnoměrné	náhodné
	dělení (kV/m)	dělení (kV/m)
Regularized minimal-energy tensor-product (SMT)	3,21	7,38
Kriging (SMT)	3,28	10,85
Least-squares (SMT)	25,57	25,92
Second-order polynomial (SMT)	15,74	15,75
Random Forest (Scikit)	5,07	14,11
Multi-layer Perceptron (Scikit)	25,57	26,86
Gaussian Process (Scikit)	0,87	13,27
Multi-layer Perceptron (Tensorflow)	1,27	11,41

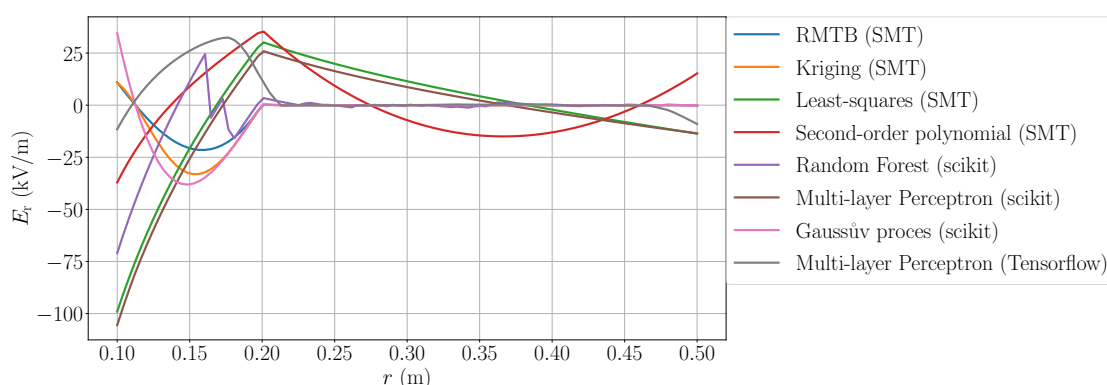
3. Optimalizace

Pro optimalizační proces v případě uvedeného ilustrativního příkladu byla minimalizována cílová funkce pro kvadratickou odchylku od nulové intenzity (viz obr. 12). Uvažujme E_0 rovno nule, pak lze cílovou funkci zapsat ve tvaru

$$OF = (E - E_0)^2.$$

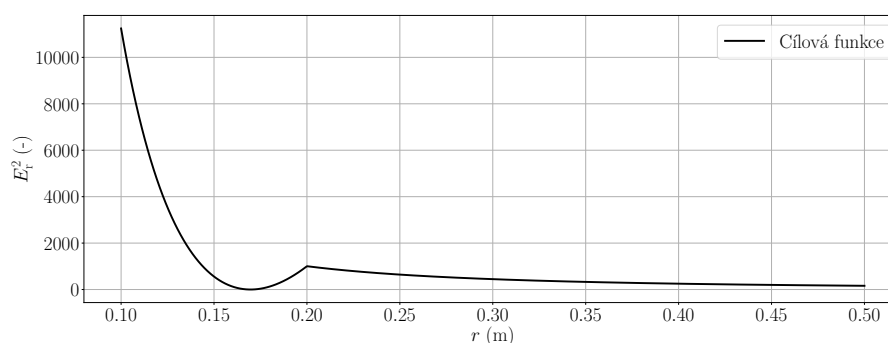


(a) Aproximace pro rovnoměrně navzorkovaný prostor - odchylka od původního řešení



(b) Aproximace pro náhodně navzorkovaný prostor - odchylka od původního řešení

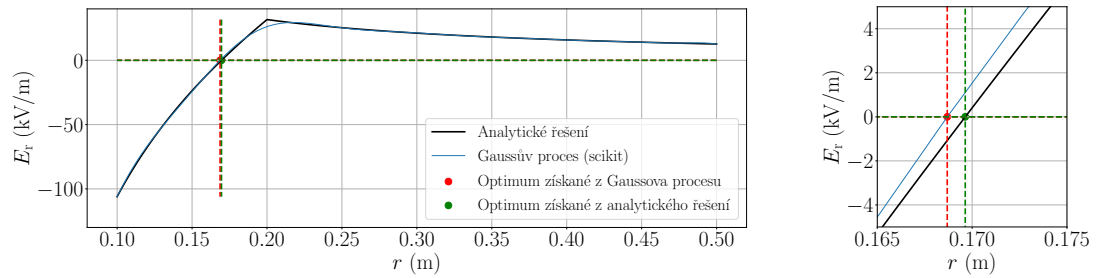
Obr. 11: Aproximace navzorkovaného prostoru - odchylka od původního řešení



Obr. 12: Cílová funkce

Řešení optimalizačního procesu proběhlo pomocí často používaného genetického algoritmu NSGA-II s populací o velikosti 10 jedinců při počtu 1000 generací.

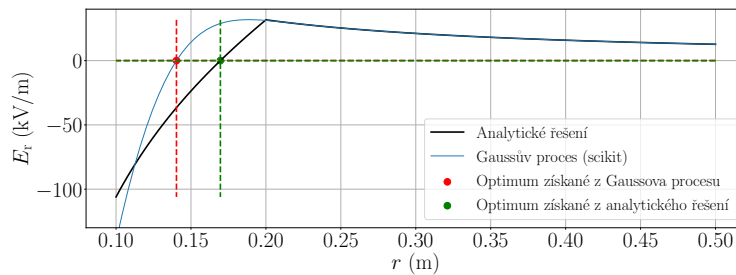
Průběh cílové funkce je zobrazen na obr. 12. Z grafu je patrné, že funkce nabývá dvou lokálních minim pro $r = 0,17$ m a $r = 0,5$ m. První z nich je zároveň globální minimum nulové intenzity. Deterministické optimalizační metody, jako například BOBYQA, vzhledem k zá-



Obr. 13: Optimalizace rovnoměrně navzorkovaného prostoru získané z analytického řešení a Gaussova procesu

vislosti na vhodné volbě počátečního bodu v tomto případě selhávají. Proto je nutné použít některou z globálních optimalizačních metod. Na obr. 13 a 14 je znázorněno řešení problému pomocí optimalizačního algoritmu NSGA-II.

Obrázek 13 znázorňuje hodnotu optima získaného z aproximace rovnoměrně navzorkovaného návrhového prostoru gaussovským procesem. Je zřejmá velmi dobrá shoda nalezení globálního optima pomocí náhradního modelu a přímého volání funkce, která vyčísluje analytické řešení. Oproti tomu řešení s nevhodně zvoleným vzorkováním na obr. 14 vykazuje chyby řádově desítky procent. U reálných problémů je správná volba vzorkování naprosto klíčová. V tomto případě byl pro jednoduchost použit statický návrh experimentu. U větších problémů bývá vhodnější použití adaptivních návrhů. V případě Gaussova procesu lze sledovat směrodatnou odchylku odhadu a na jejím základě řídit proložení návrhového prostoru.



Obr. 14: Optimalizace náhodně navzorkovaného prostoru získané z analytického řešení a Gaussova procesu

SOUČASNÝ STAV POZNÁNÍ

POKROČILÉ optimalizační metody společně s modelováním se stávají dnes již zavedeným standardem při návrhu elektrotechnických zařízení. Cílem takového návrhu je konstrukce zařízení s ohledem na vyšší výkon, nižší spotřebu, nižší výrobní náklady, apod.

3.1 VZORKOVÁNÍ NÁVRHOVÉHO PROSTORU

Realizace návrhu experimentu patří mezi jeden z nejdéle využívaných statistických nástrojů, viz článek Stephena Stiglera z roku 1974 [8]. Příklady plánovaných, kontrolovaných experimentů značně předcházejí první pokusy o formální analýzu náhodných dat, které sahají přinejmenším do Starého zákona, možná dále. Jeden takový příklad lze dle Stiglera najít již ve Starém zákoně přibližně 200 let př. n. l. v Knize Danielově, kde Daniel prokázal, že jeho strava je lepší než krále Nebuchadnezzara. Jako kontrolní skupina přitom posloužili služebníci samotného krále.

Myšlenky na velikost zkoumaného vzorku lze najít v práci Walzera [9] a v pracích Rabinowitche [10] a Sheynina [11], uvádějící shodně příklad z období kolem roku 150 n. l., kde řecký lékař Galén představuje debatu mezi empirikem a dogmatikem o relativních výhodách zkušeností a teorie v lékařském výzkumu. Předmětem debaty je otázka, jak se člověk, který se zavázal k experimentálnímu přístupu, může rozhodnout, že je k dispozici dostatek důkazů. V 11. století popsal mnoho moderních principů arabský lékař Avicenna ve svazku *Canon of Medicine*, jenž byl předním lékařským textem po téměř osm století. Avicenna zde uvedl sedm pravidel pro lékařské experimenty, která zdůrazňovala nutnost kontrol a replikace, nebezpečí příznivých účinků a nutnosti měnit jeden faktor po druhém a pozorování účinků pro více různých faktorů [12].

Další výraznější snahy lze zaznamenat až v 17. a 18. století. Za zmínku stojí již zmiňované Georgenovo pojednání uvedené v [8] z roku 1815. To se zabývá problémem, kdy reakce systému závisí na jedné nezávislé proměnné a je potřeba odhadnout její hodnotu v jednom bodě, a jakým způsobem tuto proměnnou vybrat za předpokladu náhodných chyb při odpovědích systému, který se snaží řešit dostatečně velkým množstvím pozorování. Představuje, co může být první explicitní aplikací principu nejmenších čtverců na obecný polynomiální regresní model. Georgenova práce je jednou z prvních, která se pokouší matematicky se vypořádat s problémem návrhu v regresním rámci, což ukazuje, že plánování experimentů bylo z matematického hlediska zvažováno již v roce 1815. Přesto je za průlom považována až práce sira R. A. Fishera [13].

Následně byly pro reálné (fyzikální) experimenty vyvinuty různé návrhy, obecně známé jako následující klasické vzory:

- Full and Half Factorial [14],
- Central Composite,
- Plackett and Burman (1946) a
- Box and Behnken (1960).

Zatímco klasický DoE se v zásadě zaměřuje na fyzikální experimenty, nástup počítačů vyvolal novou větev DoE. Jedná se o moderní DoE, kdy jsou reálné experimenty nahrazovány již zmíněnými počítačovými návrhy, jež umožňují mnohdy v praxi neproveditelné experimentování. Primárním cílem DoE je rozhodnutí o bodech, v nichž bude chování systému simulováno. Klasické metody jsou již poměrně dobře známé a jejich přímá aplikace naráží na různá úskalí kvůli zásadním rozdílům mezi fyzikálními a počítačovými experimenty. Většina fyzikálních experimentů má stochastický charakter kvůli řadě neznámých nebo špatně kontrolovatelných faktorů, které mají za následek generování náhodných chyb. Klasické metody návrhu experimentu jsou navrženy tak, aby řešily nevyhnutelnou náhodnost.

Metodami návrhu vzorkování prostoru se v současnosti zabývá mnoho autorů. Za zmínku stojí práce následujících autorů:

- Koehler a Owen (1996) - důkladné přezkoumání DoE v [15],
- Giunta a kol. (2003) - představení základních technik DoE v [5],
- Chen a kol. (2006) - přezkoumání počítačových návrhů ze statistického hlediska v [16],
- Garud (2017, 2018) - přezkoumání DoE, souhrn nejnovějších poznatků v oblasti [7, 17].

Velkou oblibu dnes získávají i tzv. semiempirické modely, které slučují počítačové a fyzikální experimenty, čímž se zabývá např.:

- Chen a kol. v práci [16] zaměřené na přezkoumání návrhů počítačových experimentů ze statistického hlediska,
- Blaue (2008) v [18],
- Levy a Steinberg (2010) v [19] obsahující krátký přehled počítačových experimentů (ale samotnému DoE je věnována jen malá část),
- Mosbach (2012, 2014, 2015, 2017) v [20–23].

V roce 2012 Pronzato a Müller v [24] zveřejnili přehled zaměřený na podrobný popis pokroku nad rámec prací Koehlera a Owena z roku 1996 a Chena a kol. z roku 2006. Ačkoli diskutovali o technikách vyplňování prostoru, nediskutovali o nadcházejících možnostech adaptivního vzorkování. Pro adaptivní vzorkování bylo vyvinuto značné množství technik v průběhu let, ale objevilo se i několik výzev, které doposud nebyly překonány, jako je například „*kletba dimenzionality*“, kdy výpočetní náročnost problému narůstá spolu se zvětšováním dimenze návrhového prostoru a požadovaný počet bodů návrhového prostoru roste spolu s rozměry exponenciálně. Ve svých publikacích se na problém rostoucí výpočetní náročnosti v souvislosti se zvětšováním dimenze návrhového prostoru zaměřují:

- Shan a Wang (2010) v [25],
- Cai (2016) v [26],
- Liu a kol. (2018) v [27].

Kupresanin a Johannesson upozorňují v [28], že o této problematice je dosud poměrně velmi málo známo a dostupná literatura poskytuje jen omezené množství informací. Proto se ve

své publikaci snažili porovnat různé přístupy pro modelování Gaussova procesu, přičemž však nebyli schopni jednoznačně identifikovat metodu, která by byla pro tuto problematiku nejvhodnější.

3.2 NÁHRADNÍ MODELÝ

Jedním z problémů je rychlost řešení dopředného problému. Jako perspektivní oblast pro urychlení výpočtu se jeví využití zjednodušených modelů. Výchozí koncept je znám a byl poprvé popsán v článku Johna Bandlera [29] z roku 1994, a pak dále rozpracován v dalších publikacích, jako je [30]. Základní myšlenka spočívá v nahrazení plného reálného modelu modelem zjednodušeným, který v jistých konkrétních směrech s dostatečnou přesností reprezentuje plný model společně s výrazně nižší výpočetní náročností, viz obr. 5 v předchozí kapitole.

Zmíněný přístup bývá často označován jako space mapping (SM) nebo také surrogate model. Tento model může být reprezentován například plným modelem řešeným na hrubé síti, neuronovou sítí [31] nebo dalšími metodami.

Metoda náhradního modelování je v současné době jedním ze slibných přístupů k náročnějším počítačovým simulacím. Ukazuje se, že může mít poměrně široké využití a celé spektrum možností, jak k ní přistupovat, například:

- Sanchez a kol. (2005) - lineární mapování prostoru při návrhu vysokofrekvenčních obvodů,
- Encica a kol. (2007) - tvarová optimalizace hlasového aktuátoru s využitím SM metod pro stanovení přibližného modelu využití linearizace v blízkém okolí pracovního bodu návrhových parametrů,
- Vivier a kol. (2011) - návrh rezonančního lineárního aktuátoru, který se typicky používá v membránových čerpadlech. Cílem je zde navrhnout aktuátor s maximální generovanou silou a s minimálním objemem a hmotností hlavních konstrukčních částí. Jako aproximaci využívá zjednodušený konečněprvkový model v poměru elementů 1:24.

Mezi další oblíbené techniky v aproximaci parametrů, zejména v elektrických strojích, patří náhradní obvody. Jsou vhodné pro stanovení základních elektrických vlastností, avšak nejsou dost dobře použitelné pro optimalizaci tvaru nebo mechanických parametrů. Pokročilou metodu ekvivalentního obvodu uvádí Amrhein a Krein (2009) v článku [32]. Autoři zde pro stanovení magnetické síly z Maxwellova tenzoru pnutí využívají náhradní odporový síťový model. Z uvedených článků je zřejmé, že se jedná o perspektivní směr výzkumu.

V roce 2012 se Gupta a Bhakta snažili vyvinout integrovaný náhradní model, který by zlepšil screening inhibitorů v hostitelských buňkách pro patogen *Mycobacterium tuberculosis* [33]. Při simulaci náhradních modelů jsou v literatuře často využívány genetické algoritmy, jako například

- práce Loschilov a kol. (2013) [34] uvádí možnost intenzivnějšího využívání náhradního modelu pro větší velikosti populací v souvislosti s optimalizací hyperparametrů pomocí samoadaptivní metody Covariance Matrix Adaptation Evolution Strategy (saACM-ES),

- práce Liu a kol. (2014) [35] je zaměřena na náhradní modely, které jsou podporovány evolučními algoritmy (SAEA).

Většina SAEA algoritmů je zaměřena jen na problémy o poměrně malém rozsahu, zatímco jejich efektivita pro řešení středně velkých problémů není zatím dost dobře prozkoumána. Zpravidla je zde třeba posoudit 20–50 rozhodovacích proměnných. Liu a kol. se snažili navrhnout a prozkoumat algoritmus náhradního modelu Gaussova procesu pro výpočetně náročné problémy právě středního rozsahu. Hlavní součástí je přitom náhradní model s vyhledávacím algoritmem pro větší problémy, v nichž je nákladné vytvořit kvalitní náhradní model a techniky redukce dimenzí kvůli *kletbě dimenzionality*. Snaží se použít nový rámec pro koordinaci náhradního modelování a evolučního vyhledávání pro vyhledávání zaměřené jen na velmi malou slibnou oblast.

Garud, Karimi a Kraft v roce 2017 představili nový algoritmus inteligentního vzorkování (SSA) pro řešení řady náhradních nelineárních programovacích problémů pro umístění bodů se zaměřením na větší přesnost procesu a snížení celkové zátěže výpočetních mechanismů [36]. Wistuba se ve své práci zaměřuje na klíčovou roli volby hyperparametrů a výběru algoritmů jako klíčových prvků pro strojové učení. Ve své publikaci [37] představil model, který přenáší znalost výkonu algoritmu na souborech dat k automatické akceleraci optimalizace [35] pro nový soubor dat ve dvou fázích. V první fázi je funkce, která mapuje konfigurace hyperparametrů, aproximována pro již dříve známé datové sady. Ve druhé fázi se aproximace kombinují za účelem seřazení konfigurace hyperparametrů pro novou sadu dat. Blackam a kol. [38] využívá náhradní modely pro gravitační vlny vyzařujícího binárního systému černé díry. Využití náhradních modelů v této oblasti by zde mělo umožnit studovat odvození obecné relativity a pozorovat gravitační vlny. Na svoji práci navázali v další publikaci [39] zaměřené na náhradní modely křivek numerické simulace relativity za účelem získání přesnější binární gravitační křivky černé díry.

Generování náhradního modelu má obecně problémy s velkým počtem nezávislých proměnných, což má za následek velký vzorkovací prostor. Straus a Skogestad se ve své publikaci [40] z roku 2018 zaměřují na aplikaci „samooptimalizujících“ se konceptů pro efektivnější generování náhradních modelů v ustáleném stavu. Pokud by měl být pro optimalizaci použit náhradní model, samooptimalizační proměnné by měly být schopné mapovat povrch odezvy, jež se blíží k optimální a tím snížit i náročnost modelu. Mapovaný povrch by se tak stal mnohem plošším, což by mohlo umožnit i jednodušší zobrazení nebo dokonce zanedbat vliv určitých proměnných.

Song a kol. vyvinuli náhradní model s multiobjektivním memetickým algoritmem (SMOMA) pro optimální čerpací strategie rozsáhlých problémů pobřežních podzemních vod, který integruje efektivní náhradní model založený na datech s vylepšeným nedominantním tříděním genetickým algoritmem II (NSGA-II), který využívá operátora místního vyhledávání k urychlení jeho konvergence v optimalizaci [41]. V jejich studii je model založený na stroji Kernel Extreme Learning Machine (KELM) během evolučního hledání adaptivně trénován, aby uspokojil požadovanou úroveň věrnosti náhradního signálu, takže brání hromadění chyb předpovědi a vede ke správnému konvergování na skutečnou Pareto-optimální frontu. Adaptivní náhradní model nejen zlepšuje přesnost predikce Pareto-optimálních řešení ve srovnání s jednorázovým náhradním modelem, ale také udržuje ekvivalentní kvalitu Pareto-optimálních řešení ve srovnání s řešeními NSGA-II ve spojení s původním simulačním modelem, zatímco

zachovává výhody náhradních modelů při snižování výpočetní zátěže až na 94 % úspory času.

V současné době se začínají objevovat příspěvky věnované pokročilým surrogate modelům využívajících metody strojového učení, které zahrnují pravděpodobnost odhadu cílové veličiny; příkladem jsou již zmiňované publikace zaměřené na gaussovský proces. Mezi další perspektivní techniky patří využití neuronových sítí, jež jsou založeny na principu fungování neuronů v lidské nervové soustavě. McCulloch a Pitts [42] pracovali na modelování funkcí mozkových neuronů, na tomto základě pak definovali matematický model neuronu, který se používá dodnes. V roce 1949 prezentoval Hebb v [43] zákon o modelování funkcí synaptických neuronů; jedná se o základní pravidlo pro učení neuronových sítí. V roce 1958 pak definoval Rosenblatt [44] *perceptron* - jednoduchý matematický model popisující funkci neuronů v lidském mozku. V roce 1980 pak Hecht a Nielsen navrhli vícevrstvý perceptron, který řešil problémy s lineární separací u původního perceptronu [45]. Rumelhart ve své práci [46, 47] značně zpopularizoval trénování vícevrstevných perceptronů a dokázal, že se jedná o univerzální aproximátory. Poprvé byly také neuronové sítě nasazeny v komerčním odvětví pro rozpoznávání ručně psaného PSČ. Velkou oblibu zaznamenávají konvoluční neuronové sítě a LSTM algoritmy pro problémy opakujícího se školení NN. Velkým průlomem byla možnost trénování neuronových sítí s velkým počtem vrstev spolu s vhodným inicializováním, viz [48]. Glorot ve své práci [49] navrhl schéma inicializace hmotnosti, běžně známé jako Xavierova inicializace s trénováním bez učitele, které je bráno jako standard pro hluboké učení. Ve své práci rovněž zdůraznil, že výběr aktivační funkce má velký vliv na relevanci výstupů, což vedlo ke zvýšenému zájmu zaměřenému na výzkum aktivačních funkcí a definici aktivační funkce ReLU. Rozmach hlubokého učení měl značný dopad na prognózy časových řad [50–53]. V současnosti se navrhuje nové modely a architektury speciálně pro prognostické úlohy, které využívají hluboké učení k přeplnění klasických prognostických modelů nebo se vyvíjejí zcela nové přístupy. Tyto nové modely poskytují významné zlepšení výkonu ve srovnání s restriktivní jednovrstvou architekturou vícevrstevného perceptronu (MLP – Multi Layer Perceptron), která byla dominantní (hlavně kvůli praktickým omezením) v předchozích desetiletích [54] a stávají se nejmodernějšími aplikacemi pro předpovídání časových řad [55].

3.3 POKROČILÉ OPTIMALIZAČNÍ TECHNIKY

Tradiční optimalizační algoritmy jako je například Nelder-Mead nebo BOBYQA jsou implementované v mnoha softwarových balících. Vzhledem k tomu, že jsou vhodné pouze pro hledání lokálního optima, jejich použití při tvarové optimalizaci nebo identifikaci je velmi omezené. Mohou ovšem najít uplatnění při urychlení nových a perspektivních algoritmů. Jedním z nich je Bayesovská optimalizační metoda [56] a [57]. Tato metoda využívá pravděpodobnostního přístupu k vytvoření surrogate modelu předem neznámých funkcí (hyperparametrů). Konkrétní implementace [56] využívá knihovnu NLOpt [58], která zahrnuje mimo jiné i tradiční optimalizační algoritmy pro hledání optimálních parametrů jader. Na tyto funkce dále aplikuje metody strojového učení pro výběr vhodných bodů pro zlepšení náhradního modelu. Tento přístup se využívá i v dalších aplikacích jako je například Sequential Kriging Optimization (SKO), Sequential Model-Based Optimization (SMBO) nebo Efficient Global Optimization (EGO). Vedle těchto algoritmů se zdají být perspektivní i přírodou inspirované metaheuristické evoluční algoritmy (EA) jako například algoritmus hnízda kuřky [59] nebo rojové algoritmy. Výhoda těchto algoritmů je jejich snadná implementace. Nepotřebují časově náročný výpočet gradientu a jsou ze své podstaty globální. Při imple-

mentaci EA nestačí pouze určit vhodná kódovací schémata a evoluční operátory, ale je také potřeba zvolit vhodné nastavení parametrů tak, aby se zajistila úspěšnost algoritmu, což má za následek velkou náročnost výpočtu vzhledem k potřebnému poměrně značnému počtu nutných vyčíslení [60]. Velkou pozornost vzbuzují v současnosti již zmíněné rojové algoritmy inspirované analogií ptačích hejn hledajících potravu. První simulace provedená v roce 1995 Kennedym a Eberhartem dala základ nové optimalizační metodě Particle Swarm Optimization (PSO) [61]. Na obdobném principu Dervis Karaboga [62] navrhl například algoritmus pro umělé včelí kolonie Artificial Bee Colony (ABC) založený na podobnosti s chováním včel při hledání potravy. Algoritmy založené na chování zvířat jsou v dnešní době hojně využívány a dočkaly se mnoha modifikací; jako příklad lze uvést algoritmus Chicken Swarm Optimization (CSO), jež napodobuje hierarchické pořadí v hejnech kuřat i ostatní drůbeže [63], Magnetotactic Bacteria Moment Migration Algorithm (MBMMA), který je založený na migraci magnetotaktických bakterií, viz [64], případně Power Spectrum Based Magnetotactic Bacteria Algorithm (PSMBA) dle [65].

3.4 OPTIMALIZAČNÍ TECHNIKY S VYUŽITÍM NÁHRADNÍCH MODELŮ

Optimalizace s využitím náhradních modelů (SBO) představuje typ optimalizačního algoritmu, který využívá náhradní modely k přiblížení nákladným funkcím cíle a omezení, což vede k přidávání a vyhodnocování nových vzorových bodů směrem k optimálnímu. Ukázalo se, že SBO je velmi efektivní pro konstrukční problémy, kde se často používá nákladná numerická analýza. S nástupem vysokorychlostních počítačů ve druhé polovině 20. století bylo možné využívat efektivněji možnosti numerického modelování. Schmit v roce 1960 v [66] položil základy moderní strukturální optimalizace a prezentoval první návrh syntézy matematické optimalizace a numerických metod. Tehdejší využívané metody založené na derivátech, jako je matematické programování nebo přístupy kritérií optimality, ve své práci v roce 1978 zhodnotil Venkaya v [67]. Sobieszcanski-Sobieski však tento přístup v roce 1984 zpochybnili v [68] kvůli

- příliš velkým nákladům a špatné spolehlivosti výpočtu derivací cílových a omezujících funkcí,
- vysokým výpočetním nárokům v důsledku opakování hodnocení nákladných analýz pomocí FEM.

Hsu ve své práci z roku 1994 [69] poukazuje na některé problematické aspekty v této oblasti, jako je

- problém praktické implementace,
- využití přímé vazby s procesem vyhodnocení pro efektivní použití optimalizační metody, např. výpočet derivátů současně s přímými výsledky,
- nutnost specifické formulace pro daný problém.

Jak se zmiňuje Blum v [70], implementace do komerčních řešičů byla značně problematická, případně nebyl k dispozici jejich zdrojový kód, takže od 60. let se začaly rozvíjet metody bez derivátů. Místo metod s využitím derivátů byly tyto metody navrženy tak, aby se spoléhaly přímo na vyhodnocení funkcí, ovšem za cenu většího množství analýz, viz [71].

Velké úsilí bylo věnováno výpočetně levnějším aproximačním modelům k poskytnutí řešení optimalizace s využitím méně výpočetních zdrojů [72]. Gorissen a kol. v [73] se zaměřují na reprodukci aproximace odezvy původní simulace, přičemž tyto aproximační modely byly

navrženy tak, aby nahradily nákladné simulace ve fázi optimalizace.

Navzdory rostoucí popularitě SBO se zřídka používá jako obecný optimalizační algoritmus, a to kvůli jeho nedostatečným konvergenčním vlastnostem, obtížím spojeným s takzvanou *kletbou dimenzionality* a neúplným funkcím obecného optimalizačního algoritmu. Během uplynulého desetiletí řada odborníků nepřetržitě usilovala o vývoj SBO směrem k efektivnímu globálnímu optimalizačnímu algoritmu, který dokáže vyřešit libovolné optimalizační problémy ve spojitém návrhovém prostoru. Han se ve své práci z roku 2016 snaží o vývoj optimalizace s využitím náhradních modelů směrem k optimalizačnímu algoritmu, který dokáže vyřešit libovolné optimalizační problémy se spojitým návrhovým prostorem.

HLAVNÍ CÍLE PRÁCE

V SOUČASNÉ době jsou již matematické modely dopředných úloh popsaných zpravidla soustavou parciálních diferenciálních rovnic ve velké většině známy, stejně jako metody jejich numerického řešení.

Pozornost bude proto věnována oblastem, jejichž výzkum není doposud ukončen, jelikož doposud publikované práce, zaměřené spíše na řešení určitých konkrétních příkladů, ještě nemají zobecňující charakter. Jedná se zejména o využití metod návrhu experimentů a náhradních modelů při výpočtech fyzikálních polí. Tyto metody najdou využití hlavně při hledání optimálního řešení problémů v elektrotechnice.

Hlavní cíle dizertační práce byly stanoveny na základě současného stavu poznání. Se stoupající výpočetní náročností počítačových simulací využívaných při návrhu zařízení nejenom v elektrotechnice se zvyšuje zájem o metody, který umožňují redukovat složitost modelu. Jedním ze slibných přístupů je aproximace složitého fyzikálního modelu podstatně jednodušším náhradním modelem. Tyto metody najdou uplatnění zejména v prvních návrzích zařízení nebo u expertních systémů. Hlavním cílem dizertační práce je formulace a následná implementace dostatečně přesných náhradních modelů, které mohou být využity při optimalizacích náročných elektrotechnických problémů.

Tento cíl může být pak shrnut do následujících bodů:

1. Výběr vhodných metod vzorkování prostoru parametrů.
2. Výběr vhodných náhradních modelů pro využití u výpočetně náročných počítačových simulací.
3. Výběr vhodných optimalizačních metod s využitím náhradního modelu, které mohou být použity pro pokročilé řešení problémů v elektrotechnice.
4. Ověření vhodnosti vybraných metod na problémech s analytickým řešením a na konkrétních praktických aplikacích s experimentální verifikací.

Část II

OPTIMALIZACE S VYUŽITÍM NÁHRADNÍCH MODELŮ

VZORKOVÁNÍ NÁVRHOVÉHO PROSTORU

Pro vzorkování prostoru nejčastěji využíváme metody návrhu experimentu - DoE (**D**esign **o**f **E**xperiments).

Aby byl celý proces návrhu experimentu uskutečnitelný v rozumném měřítku (zdroje, čas, výpočetní náročnost), jsou na jeho proveditelnost kladeny dva základní požadavky:

1. účinnost - minimalizace množství experimentů k prozkoumání návrhového prostoru poskytnutím co nejrelevantnějších parametrů a
2. přesnost - minimalizace nesrovnalostí predikce náhradního modelu a plné simulace.

Primárně lze metody experimentu dělit na

- **klasické**, které jsou zaměřené především na reálné fyzické experimenty,
- **moderní**, které pracují na principu klasických metod aplikovaných na počítačové simulace místo na fyzický problém.

Klasické metody se vyvíjely v průběhu času spolu s konkrétními požadavky na reálné experimenty a jsou v literatuře již poměrně dobře popsány. Vzhledem k uzpůsobení požadavků na metody dle konkrétního experimentu je jejich aplikace na počítačové experimenty značně problematická z následujících důvodů:

- existence zásadních rozdílů mezi počítačovými a reálnými experimenty,
- většina reálných experimentů je stochastického charakteru díky vlivu náhodných faktorů,
- díky náhodným faktorům vzniká velké množství náhodných chyb.

Tyto metody jsou navrženy tak, aby řešily nevyhnutelnou nutnost působení náhodného faktoru a předpokládají lineární či kvadratickou aproximaci pro odezvu systému. Pokud má systém odezvu $f(x)$, pak lze tuto odezvu v experimentu popsat jako

$$f_t(x) = f(x) + \epsilon,$$

kde $f_t(x)$ je reálná odezva systému a ϵ představuje náhodnou chybu. Hlavním cílem DoE je odvodit nejlepší možnou aproximaci i přes náhodnou chybu. Pokud existuje předpoklad pro lineární nebo kvadratickou odezvu systému, pak vrcholy na hranicích návrhového prostoru budou nejvhodnějšími ukázkovými body, viz obr. 17.

Moderní metody návrhu již ale neuvažují lineární nebo kvadratickou odezvu systému a počítají s deterministickými simulacemi. Dle Giunty [5] je primárním hlediskem moderních metod DoE vyplňování prostoru a nejvhodnější vzorové body jsou rozmístěny nejen na hranicích ale po celém návrhovém prostoru.

Tato práce se věnuje právě moderním metodám, k nimž lze přistupovat ze **statického** a **adaptivního** hlediska [74].

1.1 STATICKÉ METODY

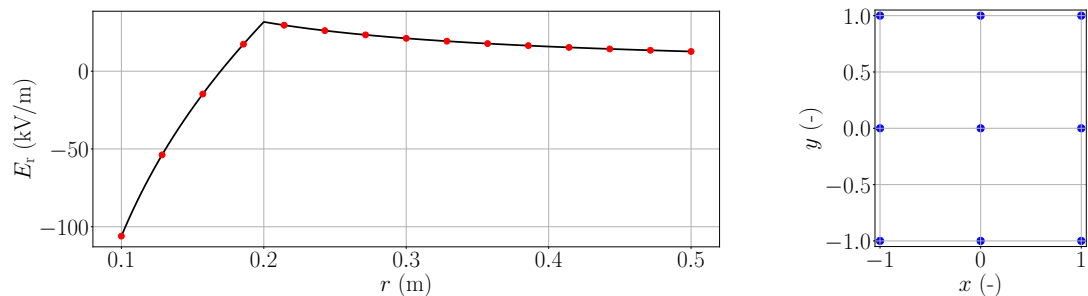
Statické metody DoE pracují pouze s vyplňováním vstupního prostoru a díky tomu je jejich implementace poměrně jednoduchá. V průběhu let byla navržena celá řada statických metod DoE, ale pro účely tohoto přehledu je diskuse omezena na ty nejběžnější - rovnoměrné, náhodné, faktoriální, Latin-hypercube sampling (LHS) a Haltonovy sekvence.

1.1.1 Rovnoměrné rozdělení

Rovnoměrné rozdělení je nejspíše nejjednodušší pravděpodobnostní rozdělení pro vyplnění návrhového prostoru:

- má konstantní hustotu pravděpodobnosti $P(X) = \text{konst.}$,
- rozložení vzorových bodů v návrhovém prostoru lze vyjádřit jako náhodnou veličinu, rozloženou rovnoměrně po celém experimentálním prostoru.

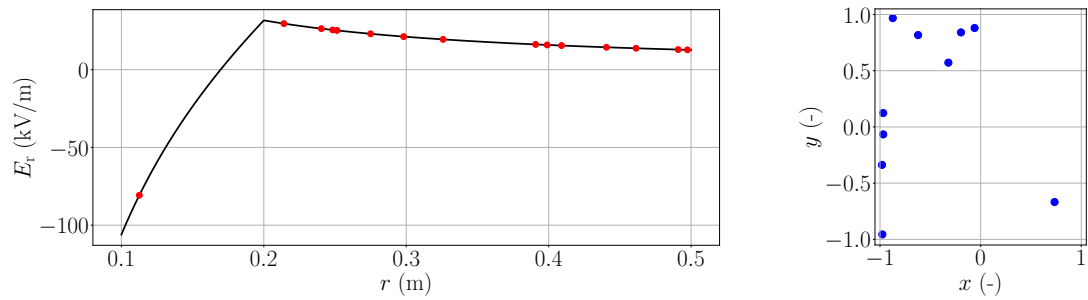
Příklad rovnoměrně navzorkovaného prostoru je na obr. 15 vlevo pro ilustrativní příklad 2 a vpravo pro obecný jednotkový prostor.



Obr. 15: Rovnoměrné rozdělení - aplikace na ilustrativní příklad 2 (vlevo) a obecné zobrazení (vpravo)

1.1.2 Náhodné rozdělení

Vyplnění návrhového prostoru je náhodné a lze jej čerpat například z homogenního Gaussova rozdělení. Toto rozdělení bylo použito u ilustrativního příkladu, viz kapitola 2, obr. 2.9(b). Z aplikace na ilustrativním příkladu je patrné, že náhodné rozdělení nezaručuje rovnoměrné pokrytí návrhového prostoru, což může vést k poměrně zásadním nepřesnostem. Náhodné rozdělení může mít uspokojivé výsledky v případě lineární nebo kvadratické odezvy systému, nebo v případě většího množství vzorků, což je méně efektivní varianta z hlediska výpočetní náročnosti systému. Náhodné rozdělení návrhového prostoru je uvedeno na obr. 16 vlevo pro ilustrativní příklad 2 a vpravo pro obecný jednotkový prostor.



Obr. 16: Náhodné rozdělení - aplikace na ilustrativní příklad 2 (vlevo) a obecné zobrazení (vpravo)

1.1.3 Faktoriální

Mnoho metod pro řešení experimentálních návrhů je založeno na principu OFAT (One Factor At a Time), kdy je brán v úvahu vliv pouze jednoho vstupního faktoru. Oproti tomu faktoriální metody berou v úvahu vliv minimálně dvou faktorů na vstupu na odezvu systému. Každý faktor může nabývat předem daných hodnot, které určují úroveň návrhu.

Díky tomu mají tyto metody několik výhod:

- poskytují více informací rychleji a za srovnatelné náklady,
- mohou zkoumat více faktorů při stejných nákladech,
- mohou poskytnout lepší informaci o interakci systému, pokud existuje odlišný účinek pro jeden faktor.

Faktoriální metody jsou využívány právě pro detekci takových interakcí systému a mohou odhadnout účinky faktoru na několika úrovních ostatních faktorů, což vede k závěrům platným pro celou řadu experimentálních podmínek. K sestavení faktoriálního návrhu je zapotřebí určit vstupní faktory a jejich úrovně. Potom množství členů N_f lze vyjádřit jako

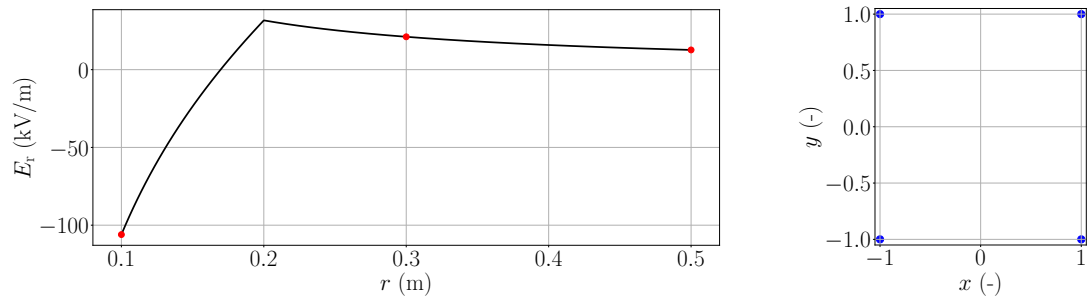
$$N_f = n^k,$$

kde n = počet úrovní a k = počet faktorů.

Pokud tedy existují 3 dvojúrovňové faktory, pak tzv. *plný faktoriální návrh* bude mít 2^3 členů. Dvojúrovňové faktoriální experimenty jsou velmi často využívány, především díky [75, 76]:

- jednoduchosti pro analýzu a vyjádření dat,
- možnosti využití grafických metod,
- malému množství potřebných vstupních dat (pro dvojúrovňový experiment se čtyřmi faktory postačí 16 návrhů experimentu, pro trojúrovňový je to již 81 návrhů),
- udržitelnosti v rozumném rozsahu i pro velké množství faktorů.

Faktoriální rozdělení návrhového prostoru je zobrazeno na obr. 17 vlevo pro ilustrativní příklad 2 a vpravo pro rozdělení obecného jednotkového prostor.

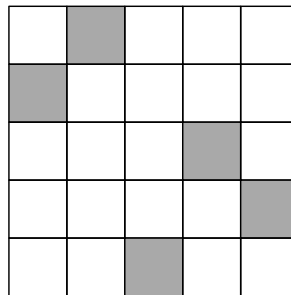


Obr. 17: Faktoriální návrh - aplikace na ilustrativní příklad 2 (vlevo) a obecné zobrazení (vpravo)

1.1.4 Latin-Hypercube Sampling (LHS)

Patří mezi statistické metody typu Monte Carlo, jež jsou však časově poměrně náročné a většinou je nutno k uspokojivému vyčíslení použít velké množství iterací. Proto se při jejich aplikacích vyžívá mnoho redukčních metod jako je metoda LHS, která:

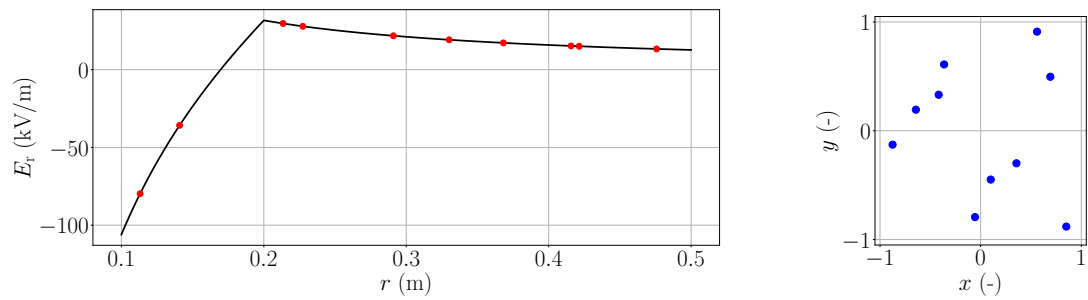
- je schopná snížit výpočetní náročnost,
- uspokojivě pokryje celý interval.



Obr. 18: Mřížka Latin Square

Princip metody vychází z čtvercové mřížky **Latin Square** obsahující v každém řádku a sloupci právě jeden prvek, jak je patrné z obr. 18. **Latin Hypercube** zobecňuje tento koncept na libovolný počet dimenzí, kde je každý vzorek jediným v každé nadrovině prostoru [77].

Principiálně se generuje N -násobek vzorků ke každé náhodné proměnné. Definiční obor distribuční funkce $F(x)$ odpovídá funkci hustoty pravděpodobnosti $f(x)$ a je rozdělen právě na N intervalů nebo vrstev nabývajících pravděpodobnosti $1/N$. Z každé vrstvy je pak vybrána jedna hodnota, která reprezentuje celý interval a v simulaci se použije právě jednou. Pomocí inverzní transformace distribuční funkce je získána reprezentativní hodnota náhodné proměnné. Existuje více způsobů výběru vzorků z jednotlivých intervalů na oboru distribuční funkce [78]:



Obr. 19: LHS - aplikace na ilustrativní příklad 2 (vlevo) a obecné zobrazení (vpravo)

1. **LHS random** - generuje N náhodných čísel z intervalu $\langle 0, 1 \rangle$ o rovnoměrném rozdělení. Tyto hodnoty se přiřadí k odpovídajícím intervalům a pomocí inverzní transformace se pak stanoví hodnoty vzorků $x_{i,k}$ pro proměnnou X_i

$$x_{i,k} = F_i^{-1} \left(\frac{n + (k + 1)}{N} \right),$$

kde k je k -tý vzorek i -té náhodné proměnné X_i , F_i^{-1} je inverzní distribuční funkce této proměnné a n značí náhodně vygenerované číslo z intervalu $\langle 0, 1 \rangle$.

2. **LHS median** - výběr hodnoty ze středu intervalu $1/N$ distribuční funkce spojený s přímým použitím inverzní distribuční funkce

$$x_{i,k} = F_i^{-1} \left(\frac{k - 0,5}{N} \right).$$

3. **LHS mean** - výběr reprezentativních vzorků $x_{i,k}$ jako střední hodnoty intervalu vymezeného pro funkci hustoty pravděpodobnosti dané proměnné X_i . Díky tomu se pak vzorky budou nacházet přesně v těžišti rozdělovacích ploch

$$x_{i,k} = N \int_{y_{i,k-1}}^{y_{i,k}} x f(x) dx.$$

Meze integrálu lze přitom určit ze vztahu

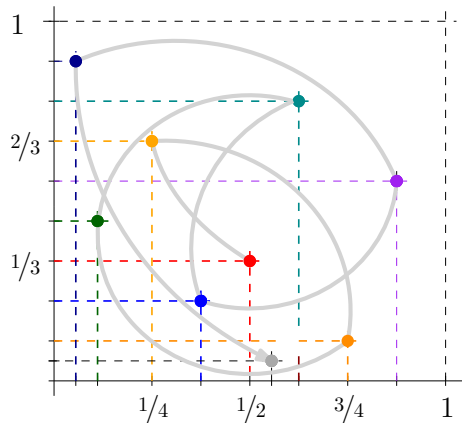
$$y_{i,k} = F_i^{-1} \left(\frac{k}{N} \right).$$

Použití LHS metody na dělení experimentálního prostoru je zobrazeno na obr. 19 pro ilustrativní příklad 2 vlevo a vpravo pro obecný prostor.

1.1.5 Haltonova sekvence

Metoda Haltonovy sekvence je často používána pro generování bodů v prostoru pro numerické metody, jako je například Monte Carlo. Ač se může zdát jako náhodná, funguje na deterministickém principu nesoudělitelných čísel (čísla a a b jsou nesoudělitelná, pokud jejich největším společným dělitelem je 1).

Jako příklad lze uvažovat Haltonovu sekvenci, kdy dělení jedné dimenze bude založeno na čísle 2 a druhé na čísle 3. Aby bylo možné vygenerovat posloupnost pro číslo 2, je potřeba



Posloupnost bodů pro vzorkování prostoru
pomocí Haltonovy sekvence:

1. $[1/2, 1/3]$, 2. $[1/4, 2/3]$, 3. $[3/4, 1/9]$,
4. $[1/8, 4/9]$, 5. $[5/8, 7/9]$, 6. $[3/8, 2/9]$,
7. $[7/8, 5/9]$, 8. $[1/16, 8/9]$, 9. $[9/16, 1/27]$.

Obr. 20: Princip Haltonovy sekvence

rozdělit interval $(0, 1)$ na polovinu, poté na čtvrtinu a na osminu. Pak tedy prvních 9 hodnot bude

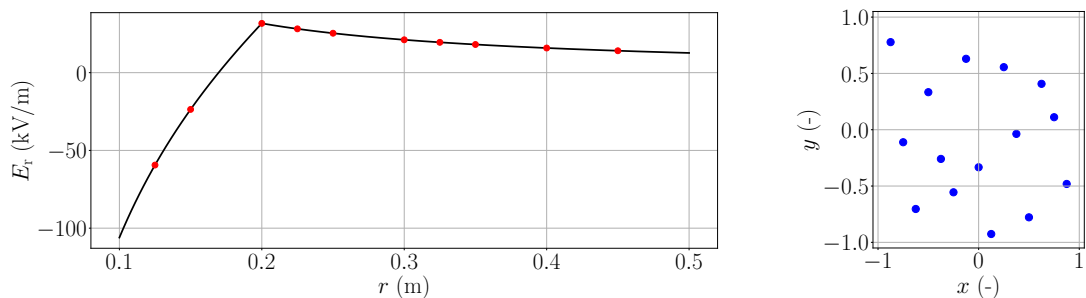
$$1/2, \quad 1/4, \quad 3/4, \quad 1/8, \quad 5/8, \quad 3/8, \quad 7/8, \quad 1/16, \quad 9/16, \quad \dots$$

Pro vygenerování sekvence pro 3 je potřeba rozdělit interval na třetiny, pak na devítiny a na dvacetisedminy, takže prvních 9 hodnot bude

$$1/3, \quad 2/3, \quad 1/9, \quad 4/9, \quad 7/9, \quad 2/9, \quad 5/9, \quad 8/9, \quad 1/27, \quad \dots$$

Při spárování je vygenerována posloupnost bodů v jednotkovém čtverci pro vzorkování, která je znázorněna na obr. 20.

Příklad navzorkovaného prostoru pomocí Haltonovy sekvence je patrný na obr. 21 vlevo pro ilustrativní příklad 2 a vpravo pro obecný prostor.



Obr. 21: Haltonova sekvence - aplikace na ilustrativní příklad 2 (vlevo) a obecné zobrazení (vpravo)

1.2 ADAPTIVNÍ METODY

Nevýhodou statického DoE je, že nebere v úvahu odezvu systému, pouze vyplní vstupní prostor. Může se však stát, že v určitých oblastech může být vzorkovaná funkce relativně plochá. Potom nám postačí méně vzorků než například u funkce nelineární, kde je zapotřebí

vzorků více ke kvalitnímu zachycení variací.

Adaptivní metody jsou dynamickou třídou DoE, jelikož berou v úvahu nejen kvalitu vstupních dat, ale i výstupních [79]. V jejich případě je tedy žádoucí propojení s náhradním modelem, na jehož základě můžeme určit důležitost výběru vzorků v daných oblastech. Výhodná je i přesnost náhradního modelu trénovaného na menším množství vzorků.

Statické metody návrhu experimentu generují všechny body v návrhovém prostoru najednou. Důsledkem toho může být podvzorkování nebo naopak převzorkování prostoru, a tím pádem i špatná aproximace systému. Výhoda adaptivních metod je tedy v poměrně nízkých výpočetních nákladech vzhledem k jejich využívání s náhradními modely a také v lepší aproximaci návrhového prostoru.

Efektivní přístup adaptivního sekvenčního vzorkování musí zahrnovat:

1) *Globální průzkum prostoru (global exploration)*

Jedná se o pokrytí návrhového prostoru pro špatně zastoupené nebo podvzorkované dílčí prostory; průzkum se zaměřuje na homogenní umístění vzorku pomocí vyplňování prostoru. Nejjednodušším případem adaptivního průzkumného vzorkování (ASED) je adaptivní 2D mřížka o velikosti 3×3 . Pokud těchto 9 elementů není dostatečných pro konstrukci přijatelné aproximace, lze mezi existující body mřížky přidat další body (např. ve středech zmíněných elementů). I když se jedná o nejjednodušší ASED, může být obvykle hrubě zrnitý a rychle se setkává s kletbou dimenzionality. Termín adaptivní naznačuje, že tyto techniky obsahují systémové informace pro umístění vzorku. Jde však pouze o postupné/iterativní techniky vyplňování prostoru. Jakoukoli techniku vyplňování prostoru lze tedy přizpůsobit. Některé techniky fungují například na principu vnořených LHD a MCS optimalizací mezipaměti a vzdáleností mezi lokalitami.

2) *Lokální těžbu dat (local exploitation)*

Umístování bodů do složitých nebo vysoce nelineárních oblastí představuje užitečný nástroj proti riziku místního shlukování návrhových bodů. Cílem těžby je kontrola kvality vzorku ve výstupním prostoru, k čemuž jsou sledovány zejména čtyři hlavní ukazatele:

- techniky křížové validace,
- lineární aproximace gradientu,
- míry zakřivení a
- umístění extrémů.

Obecně lze tedy prohledávání nových vzorků vyjádřit vztahem

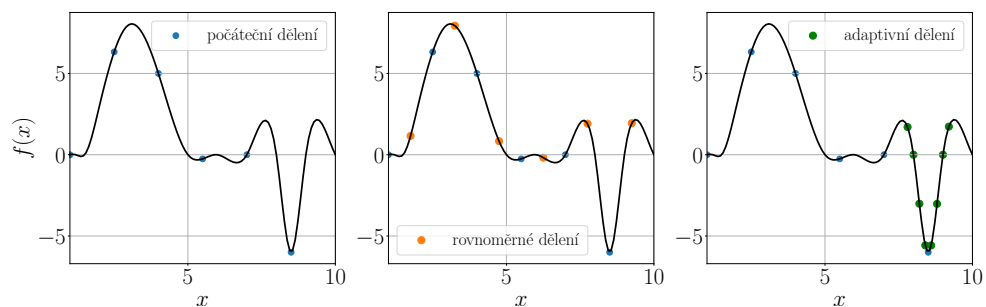
$$x_{\text{new}} = \arg \max_{x \in D} (S_{\text{local}}(x), S_{\text{global}}(x)), \quad (3)$$

kde $S_{\text{local}}(x)$ lze pojmout jako lokální využití dat a $S_{\text{global}}(x)$ jako globální průzkum prostoru.

Pro počáteční vzorek je v první řadě potřeba vyhodnotit metriky právě pro průzkum a těžbu dat. Jakmile jsou tyto metriky stanoveny, je velmi důležité nalézt rovnováhu mezi nimi, což je zobrazeno na obr. 22.

Nejprve je nutno počítat s extrémny. První graf zobrazuje prostor s počátečním vzorkováním. Na grafu uprostřed je zobrazen jen čistý průzkum prostoru realizovaný pomocí rovnoměrného dělení. Zatímco čistý průzkum může skončit nedostatečně navzorkovaným prostorem komplexních nebo nelineárních oblastí nebo naopak převzorkováním v jednoduchých lineárních regionech, tak čisté využití může vyústit např. v místní vzorkování a umístit tak body jen do zajímavých oblastí, zatímco ostatní oblasti se ponechají prázdné. To je velmi dobře patrné ze třetího grafu.

Techniky adaptivního DoE jsou většinou zaměřeny buď na využití dat, nebo se snaží najít rovnováhu mezi průzkumem a využitím.



Obr. 22: Rizika pro extrémní průzkum nebo extrémní využití dat [80]

Pro správný přístup adaptivního sekvenčního vzorkování je zapotřebí korektní dodržení několika kritérií:

1. Počáteční vzorkování

Počáteční vzorek lze vytvořit pomocí kritérií pro vyplňování prostoru, jelikož musí pokrýt celý vstupní prostor pro průzkum a dosud není známá odezva systému. Hlavním účelem je získat počáteční představu, které regiony jsou zajímavé pro další využití. Pro počáteční vzorek lze tedy použít kteroukoli ze statických strategií DoE. Poněvadž další vzorky nebudou založeny na definici počátečního vzorku, je také přijatelné použít faktoriální návrhy.

2. Rovnováha mezi průzkumem a těžbou dat

Po dokončení počátečního vzorku se vyhodnotí metriky průzkumu a vytěžených dat. Jakmile jsou tyto metriky stanoveny, je třeba najít mezi nimi rovnováhu. Jak bylo zmíněno výše, viz obr. 22, čistý průzkum odpovídá statickému DoE a čistá „těžba“ dat by se mohla zaměřit jen na silně nelineární regiony a mohla by opominout ostatní oblasti. Existuje zde riziko vynechaných oblastí, které nebyly objeveny v počáteční sadě.

Jednou z možností je normalizace těchto dvou požadavků a udržení konstantní váhy během odběru vzorků. Dále je možné sledovat dynamickou odezvu a na základě průběžně dosažených výsledků hodnotit vliv metrik. Lze například postupně omezovat průzkum a zvyšovat časové využití na základě předpokladu, že se zvýšeným množstvím vzorků klesá nejistota ohledně nových regionů a je možné se stále více zaměřovat na aktuální reakci. Je také možné

vyhodnotit, zda byl pokrok v posledních několika iteracích příznivý nebo ne, a na základě toho pak rovnováhu upravit. Pokud tedy nedávná těžba vedla ke snížení chyby, pak je rozumné v tomto pokračovat. Pokud se přínos zmenší, je vhodné se obrátit směrem k průzkumu a hledat nové oblasti.

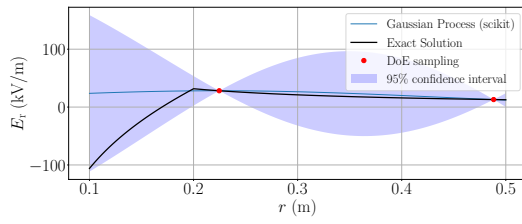
3. Výběr nových vzorků

Jakmile jsou metriky a jejich rovnováha známy, je třeba vybrat nové vzorky. Obecně se akceptují dva přístupy:

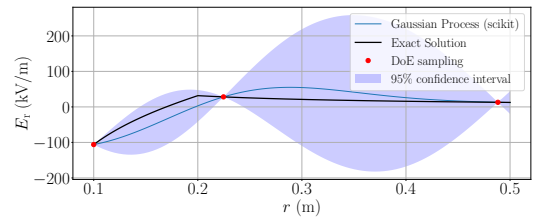
- **návrh kandidátských vzorků** a přiřazení ke každému z nich skóre tak, aby bylo možné vybrat ty, jejichž skóre je co nejvyšší nebo nejnižší,
- formulace vyhledávání jako **problému s postupnou optimalizací**.

Vzhledem k tomu, že jsou zvažována dvě protichůdná kritéria, je další postup založen na znalosti Paretovy fronty optimálních řešení, přičemž dalším krokem je výběr vzorku z přední části této fronty. Některé metody to řeší převodem problému na optimalizační úlohu s jednou cílovou funkcí, např. sčítáním nebo vynásobením obou cílů nebo použitím jednoho z kritérií jako omezení při optimalizaci druhého. Jiné určují logiku pro výběr vzorků přímo z přední části Paretovy fronty.

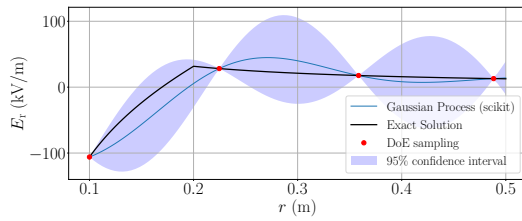
V praxi je také důležité zvážit, kolik simulací lze spustit současně, aby bylo vybráno vhodné množství nových vzorků a aby se plně využily dostupné výpočetní zdroje. Podobně lze výběr více vzorků najednou motivovat omezením množství náhradních tréninkových iterací. Nejjednodušší případ nastane, když je současně vyhodnocen pouze jeden vzorek, protože výběr optimální kombinované metriky je v takovém případě triviální. Komplikací u více vzorků je skutečnost, že kandidátské vzorky mohou být blízko vzorků původních a mají podobná skóre. Navíc formulace průběžné optimalizace nabízí nekonečné množství možných řešení. V takovém případě je třeba vzorkování provést vhodnou adaptivní metodou DoE. Na obrázku 23 je zobrazeno řešení ilustrativního příkladu pro adaptivní DoE.



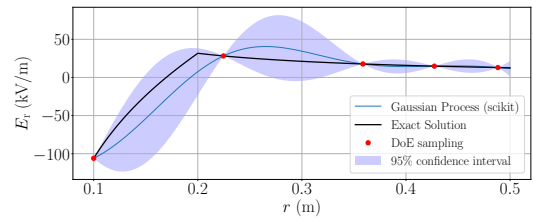
(a) Krok 1



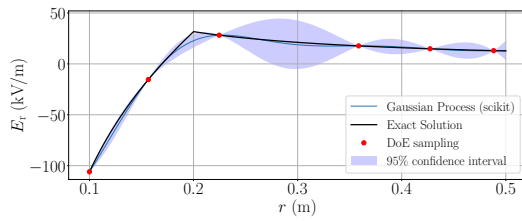
(b) Krok 2



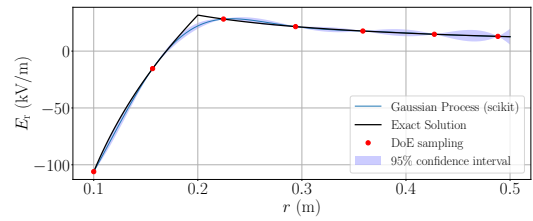
(c) Krok 3



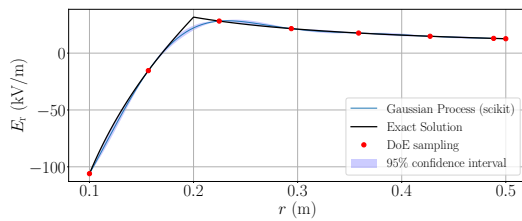
(d) Krok 4



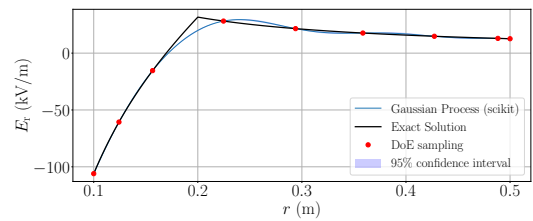
(e) Krok 5



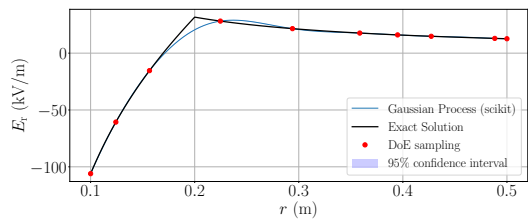
(f) Krok 6



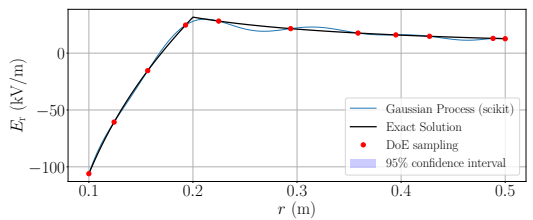
(g) Krok 7



(h) Krok 8



(i) Krok 9



(j) Krok 10

Obr. 23: Proces adaptivního vzorkování - aplikace na ilustrativní příklad 2

NÁHRADNÍ MODELŮ

Mezi nové perspektivní techniky pokročilé analýzy lze zařadit využití zjednodušených modelů nahrazujících s jistou (avšak přípustnou) chybou modely úplné. Tyto modely mohou být založeny na velkém množství algoritmů, které mají určité společné vlastnosti. Obecně lze vhodnou volbu parametrů těchto modelů nazvat strojovým učením.

Uvažujme soubor n vzorků dat obsahující vstupní parametry a výstupní veličiny. Tyto hodnoty mohou být získány například z měření nebo s využitím počítačového modelování opakovaným řešením problému. Každému vzorku parametrů odpovídá jedna nebo více výstupních veličin.

2.1 STROJOVÉ UČENÍ

Strojové učení je druh umělé inteligence, díky níž je umožněno počítači se učit. Lze je dále rozlišit podle přístupu k učení:

- Učení s učitelem (supervised learning) - vstupní data mají předem určený výstup.
- Učení bez učitele (unsupervised learning) - vstupní data nemají předem určený výstup.

Pro tuto práci jsou relevantnější přístupy pro učení s učitelem, proto bude následující kapitola zaměřena na tuto problematiku.

Problémy strojového učení s učitelem (supervised learning) lze obecně rozdělit do dvou kategorií:

- **Klasifikace** - vzorky dat spadají do dvou nebo více diskrétních kategorií a na základě již určených trénovacích dat je třeba určit neroztříděné, výstupem je volba mezi třídami, např. ano/ne, červená/modrá/zelená.
- **Regrese** - výstup tvoří jedna nebo více spojitých proměnných a cílem je najít jejich náhradní spojitou reprezentaci. Principem regresního přístupu je superpozice vlivu všech funkčních proměnných za účelem získání předvídatelné hodnoty.

V případě náhrady výstupních veličin počítačových modelů se jedná o regresi. Tato data lze získat vhodným prohledáním prostoru parametrů modelu. Obecnou praxí je rozdělení vzorku získaných dat do dvou množin. První z nich představuje *tréninková data*, na kterých se náhradní model učí, a *testovací data* sloužící k ověření vlastností již naučeného modelu.

2.2 OVĚŘENÍ NA TESTOVACÍCH DATECH

Ověření na testovacích datech lze provést implementací velkého množství metod. Obecně lze říci, že tyto metody měří vzdálenost mezi modelem a testovacími daty. Jako příklad lze uvést metody jako

- **Variační míra** je poměr, s jakým matematický model počítá s rozptylem daného souboru dat. Pokud \hat{y} je odhadovaný cílový výstup, y je správný cílový výstup a σ^2 je rozptyl, pak lze vztah pro určení variační míry vyjádřit v následujícím tvaru

$$V(y, \hat{y}) = 1 - \frac{\sigma^2\{y - \hat{y}\}}{\sigma^2\{y\}}. \quad (4)$$

Nejlepší možná dosažená hodnota $V(y, \hat{y})$ je 1. Čím je tato hodnota nižší, tím horší je i výsledek.

- **Koeficient determinace R^2** je poměr rozptylu odhadovaného závislého cílového výstupu \hat{y} od nezávislého výstupu y . Poskytuje měřítko, jak úspěšně bude model předvídat budoucí hodnoty. Stejně jako v předchozím případě je nejlepší možná hodnota 1. Konstantní model, který vždy předpovídá očekávanou hodnotu y , a to bez ohledu na vstupní vlastnosti, by získal skóre 0. Tyto hodnoty lze získat ze vztahu

$$R^2(\hat{y}) = 1 - \frac{\sum_{i=0}^{n_{\text{samples}}-1} (y_i - \hat{y}_i)^2}{\sum_{i=0}^{n_{\text{samples}}-1} (y_i - \bar{y})^2}, \quad (5)$$

kde

$$\bar{y} = \frac{1}{n_{\text{samples}}} \sum_{i=0}^{n_{\text{samples}}-1} y_i$$

je střední hodnota vzorku dat a n_{samples} je počet vzorků.

- Jednou z často používaných metod je také **průměrná absolutní chyba** (Mean Absolute Error - MAE), která vyjadřuje míru rozdílu mezi dvěma spojitými proměnnými. Pokud \hat{y}_i je predikovaná hodnota i -tého vzorku a y_i je odpovídající skutečná hodnota, potom je odhadnuté skóre R^2 definováno jako

$$\text{MAE}(\hat{y}) = \frac{1}{n_{\text{samples}}} \sum_{i=0}^{n_{\text{samples}}-1} |y_i - \hat{y}_i|. \quad (6)$$

- **Střední kvadratická chyba** (Mean Squared Error - MSE) vyjadřuje přesnost odhadů pomocí průměrné střední hodnoty druhých mocnin pro rozdíl odhadu a měření (reality).

$$\text{MSE}(\hat{y}) = \frac{1}{n_{\text{samples}}} \sum_{i=0}^{n_{\text{samples}}-1} (y_i - \hat{y}_i)^2. \quad (7)$$

2.3 METODY REGRESE A KLASIFIKACE

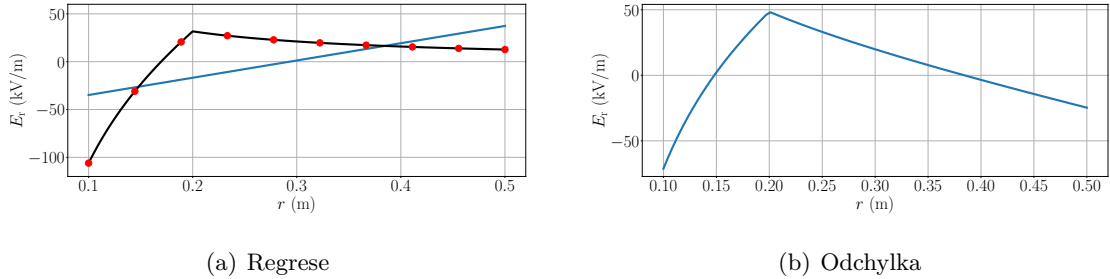
Jak již bylo zmíněno, v případě náhrady výstupních veličin u numerických modelů hovoříme o proložení výstupních dat pomocí regresní funkce. Pro náhradu proměnných na výstupu daného problému lze využít následující metody.

2.3.1 Polynomiální regrese

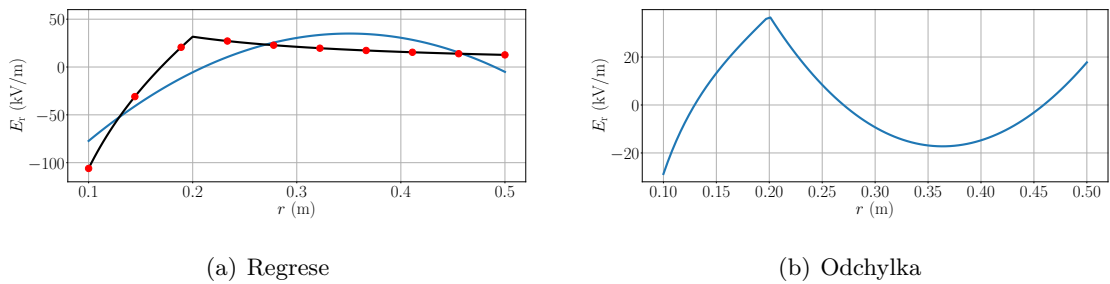
Nejjednodušším případem regresní aproximace je lineární regrese. V případě lineární regrese hovoříme o proložení dat přímkou. Tato data jsou reprezentována body, přičemž lze předpokládat, že jejich souřadnice na ose x je známá, zatímco hodnota těchto dat a tedy souřadnice bodů na ose y je známá s určitou náhodnou chybou. Pokud tedy přímkou popíšeme rovnicí

$$y = a_1 + a_2x$$

pak koeficienty a_1 a a_2 jsou optimální koeficienty pro aproximační přímku, viz obr. 24. Na obrázku 25 je příklad regrese druhého řádu. Obě proložení jsou realizována na ilustrativním příkladu 2 uvedeném v první části.



Obr. 24: Lineární regrese s rovnoměrným dělením



Obr. 25: Regrese polynomem druhého řádu s rovnoměrným dělením

2.3.2 Gaussovské procesy

Gaussovské procesy jsou jednou z obecných metod. Tato metoda je velmi často využívána právě v souvislosti se strojovým učení. Lze ji velmi dobře uplatnit pro regresi pravděpodobnostních problémů. Výhodou je, že předpoklad lze interpolovat pravděpodobnostním Gaussovým rozložením

$$f(\mathbf{x}) = \mathcal{GP}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')).$$

kde $m(\mathbf{x})$ je funkcí střední hodnoty a $k(\mathbf{x}, \mathbf{x}')$ je funkcí kovariance (jádru), kterou lze definovat mnoha způsoby, například jako exponenciální funkci (Squared Exponential Function)

$$k(\mathbf{x}, \mathbf{x}') = \exp\left(-\frac{1}{2}\|\mathbf{x} - \mathbf{x}'\|^2\right)$$

nebo konkrétněji

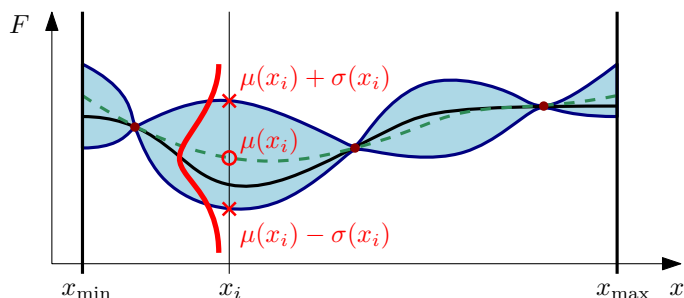
$$k(\mathbf{x}, \mathbf{x}') = \exp\left(-\frac{1}{2\theta^2}\|\mathbf{x} - \mathbf{x}'\|^2\right),$$

kde θ vyjadřuje hyperparametr, kterým lze naladit potřebné chování procesu.

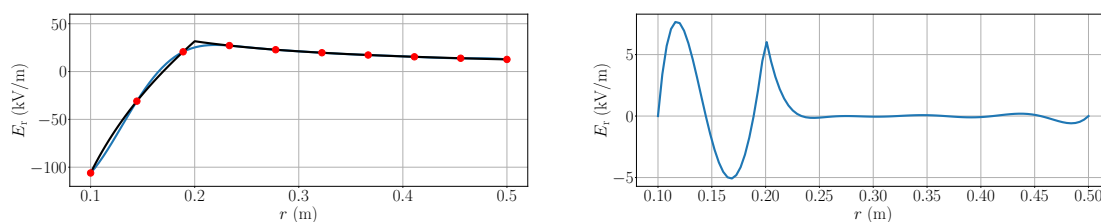
Následně je možné určit empirické intervaly spolehlivosti, přičemž princip je velmi dobře patrný z obr. 26. Lze se pak snadněji rozhodnout, které předpoklady je nutné upravit (online

přizpůsobení nebo adaptivní přizpůsobení, viz kapitola 1.2). V prostorech s vysokými dimenzemi, kde počet parametrů přesahuje i několik desítek, ale ztrácí efektivitu. K formulaci předpokladů využívají kompletní informaci o vzorcích.

Na obrázku 27 je znázorněna aproximace návrhového prostoru ilustrativního příkladu z předchozí části pomocí Gaussova procesu.



Obr. 26: Toleranční pásmo vymezené gaussovským procesem



Obr. 27: Gaussův proces - aproximace rovnoměrně navzorkovaného prostoru s počtem vzorových bodů $N = 15$ (vlevo) a odchylka aproximace (vpravo)

Jednou z nejznámějších metod založené na aproximaci gaussovským procesem jsou **metody krigování**. Kriging bývá často označována akronymem BLUE (Best Linear Unbiased Estimator), což vyjadřuje nejlepší lineární nezkreslený odhad, což vystihuje obecnou formuli pro výchozí podmínky krigování, kdy je odhadovaná hodnota vypočtena jako lineární kombinace vstupních hodnot

$$\hat{Z}(x_0) = \sum_{i=1}^N \lambda_i Z(x_i), \quad (8)$$

kde $Z(x_i)$ je naměřená hodnota pro i -tou variantu z N naměřených hodnot, λ_i je váhový koeficient, pro který platí $\sum_{i=1}^N \lambda_i = 1$ a koeficient λ_0 je koeficient pro predikované hodnoty.

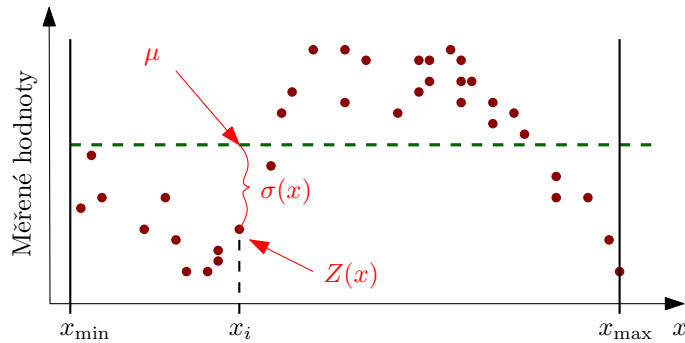
V současnosti existuje poměrně značné množství algoritmů krigingu. K charakteristice několika uvedených typů předpokládejme obecný model v následujícím tvaru

$$Z(x_i) = \mu(x_i) + \sigma(x_i), \quad (9)$$

kde $Z(x_i)$ popisuje proměnnou v bodě x_i , $\mu(x_i)$ je deterministická hodnota trendu a $\sigma(x_i)$ je autokorelovaná náhodná proměnná. Vzhledem k tomu, že většinou není známá hodnota trendu $\mu(x_i)$, která je odhadována, pak je touto chybou zatížena i náhodná proměnná. Potom tedy platí, že průměrná chyba bude rovna nule a potom autokorelace hodnot $\epsilon(x_i) + \sigma(x_i + h)$

nebude závislá na aktuální pozici, ale pouze na vzdálenosti h , viz obr. 28. Hodnota trendu může být konstantní nebo může mít daný trend, může být předem známá nebo může být odhadována. V této souvislosti potom lze rozlišit základní metody krigingu uvedené v [81]:

- základní kriging,
- jednoduchý kriging,
- univerzální kriging,
- pravděpodobnostní kriging, atd.



Obr. 28: Základní princip krigování

2.3.3 Hluboké neuronové sítě

Neuronové sítě jsou založeny na poznacích o neuronech a nervových sítích. Uvažujme *formální neuron*, který je popsán n vstupy, viz obr. 29. K aktivaci neuronu potřebujeme aktivační funkci $y(\varphi)$, kterou lze zapsat ve tvaru

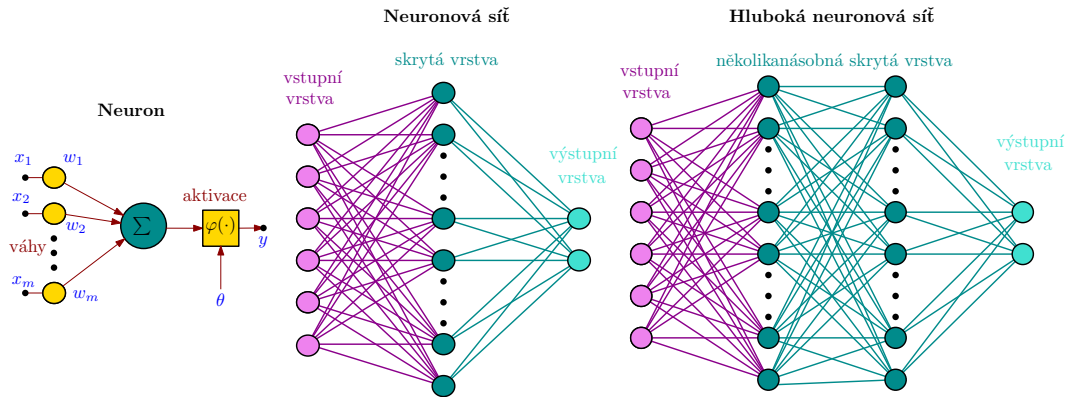
$$y(\varphi) = \sum_{i=1}^N (w_i x_i - \theta),$$

kde φ představuje sumarizační funkci přičemž $\mathbf{x} = (x_1, \dots, x_n)$ je vektor popisující jednotlivé vstupy neuronu, $\mathbf{w} = (w_1, \dots, w_n)$ je vektor přiřazující jednotlivým vstupům neuronu váhové koeficienty a θ značí prahovou hodnotu této funkce. Jako vstupní veličinu lze využít buď vnější informaci nebo výstup jiného neuronu. Po ováhování jednotlivých vstupních informací a jejich součtu je výsledek porovnáván se stanovenou prahovou hodnotou. K tomu, aby byl neuron aktivován, je často využívána prahová hodnota **aktivační funkce** $\varphi(\cdot)$. Nejčastěji používané aktivační funkce jsou sigmoidní funkce, hyperbolická tangenciální sigmoidní funkce, ReLU (Rectified Linear Unit) a skoková funkce, v tomto pořadí jsou zobrazeny na obr. 30. Pokud platí

$$\sum_{i=1}^N (w_i x_i - \theta) \geq 0,$$

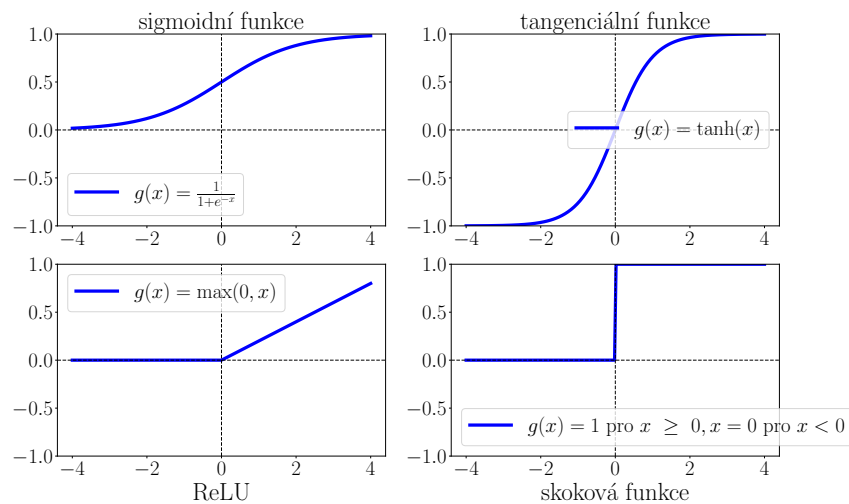
dojde k aktivaci neuronu a neuron vygeneruje signál. V opačném případě zůstane neuron pasivní.

Tyto neurony jsou vzájemně spojeny tak, aby výstup jednoho neuronu byl vstupem dalšího (nebo i více) neuronu, čímž je vytvořena neuronová síť. Princip neuronových sítí je velmi dobře patrný na obr. 29. V levé části je znázorněn základní neuron, jehož spojení s ostatními



Obr. 29: Struktura neuronové sítě

neurony vytvoří vstup dalších neuronů. V pravé části obrázku je pak znázorněno spojení těchto neuronů (znázorněných malými kruhy) do složitějších struktur, které jsou následně označovány jako „*neuronové sítě*“.

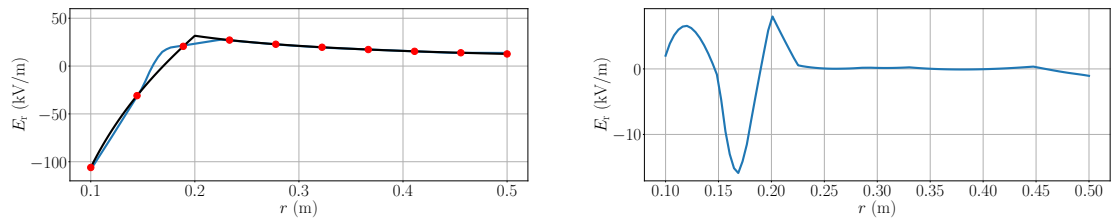


Obr. 30: Nejčastěji používané aktivační funkce

Multi-Layer Perceptron (MLP) jsou třídou neuronových sítí. Jedná se o řízený algoritmus, který se naučí funkci pomocí trénování na datové množině. Vzhledem k množině vlastností a cíle se může nelineární aproximátor naučit funkci pro klasifikaci nebo regresi. Mezi vstupní a výstupní vrstvou je alespoň jedna nebo více nelineárních vrstev, které bývají označovány jako skryté vrstvy. MLP využívá pro proces učení s učitelem backpropagation algoritmu, který při učení algoritmu pracuje s informací o gradientu ztrátové funkce pro jeden vstup a jeho výstup. Díky tomu je možné během procesu trénování výrazně omezit ztráty. Algoritmus zpětného šíření funguje tak, že počítá gradient ztrátové funkce s ohledem na každou váhu pomocí pravidla řetězce, počítá gradient jednu vrstvu po druhé a iteruje zpětně od poslední vrstvy, což zabrání nadbytečným výpočtům.

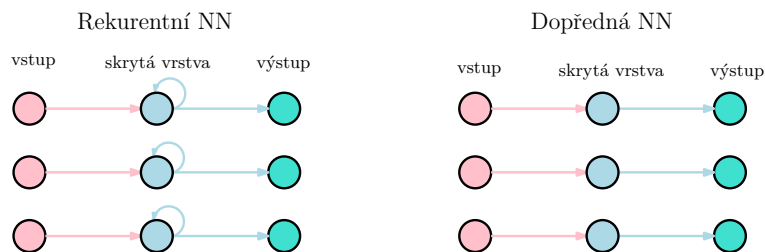
Na obr. 31 je znázorněna aproximace rovnoměrně navzorkovaného prostoru spolu s odchylkou

pomocí regresního algoritmu MLP.



Obr. 31: MLP - aproximace rovnoměrně navzorkovaného prostoru s počtem vzorových bodů $N = 15$ (vlevo) a odchylka aproximace (vpravo)

Rekurentní neuronová síť (RNN) využívá sekvenční data nebo data získaná z časových řad. Velmi často jsou RNN využívány pro jazykové překlady, rozpoznávání řeči nebo u popisu obrázků. Na jejich principu funguje velké množství běžně dostupných aplikací (Siri, Google Translate). Stejně jako v předchozím případě se jedná o algoritmus, který k učení využívá sadu trénovacích dat. Vyznačují se svou „pamětí“, protože přijímají informace z předchozích vstupů, aby ovlivnily aktuální vstup a výstup, což je rozdíl oproti tradičním dopředným neuronovým sítím, které předpokládají, že vstupy a výstupy jsou na sobě nezávislé. Oproti tomu výstup rekurentní neuronové sítě je závislý na předchozích prvcích a jeho princip je patrný z obr. 32.

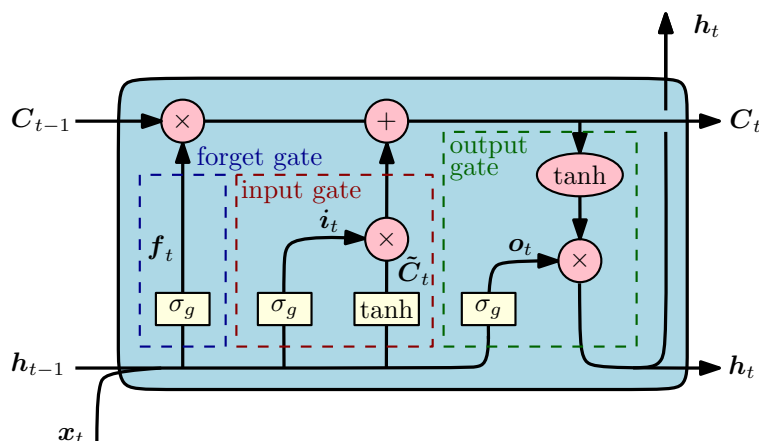


Obr. 32: Princip rekurentní a klasické neuronové sítě

Rekurentní neuronové sítě využívají k určení gradientů algoritmus zpětné propagace v čase (Backpropagation Through Time – BPTT), který se mírně liší od tradičního algoritmu, protože je specifický pro sekvenční data. Principy BPTT jsou stejné jako u tradičního algoritmu, kde se model trénuje počítáním chyb z výstupní vrstvy do vstupní vrstvy. Na základě těchto výpočtů je potom možné vhodně upravit parametry daného modelu. Oproti tradičnímu přístupu se zde sčítají chyby v každém časovém kroku, zatímco dopředné sítě nemusí sčítat chyby, protože nesdílejí parametry napříč každou vrstvou. BPTT přístupy mohou při výpočtech narazit na dva problémy známé jako explodující a mizející gradienty.

Tyto problémy jsou definovány velikostí gradientu, který vyjadřuje sklon ztrátové funkce podél chybové křivky. Pokud je gradient příliš malý, pak se stále zmenšuje a aktualizuje parametry (např. hmotnost), až se stanou absolutně nevýznamnými a algoritmus se pak již neučí. Explodující gradienty vznikají v případě příliš velkého přechodu (dlouhodobá předpověď) a vedou k nestabilitě modelu. Jedním z řešení je snížení počtu vrstev neuronové sítě, čímž se eliminuje určitá složitost modelu RNN.

Jednou z variant je algoritmus *Long Short-Term Memory (LSTM)*, který nabízí řešení problému mizejícího gradientu u klasických RNN. Tento algoritmus byl představen na úloze dlouhodobých závislostí v [82]. Pokud předchozí vstup ovlivňující výstup nenastal v nedávné minulosti, pak RNN není schopna přesně určit současný stav. Lze uvést příklad předpovědi založené na slovech zvýrazněných v textu: „Celý týden se neučil. Nemůže jít na zkoušku.“ Z kontextu je patrné, že student nemůže na zkoušku kvůli nedostatečné přípravě. Pokud ovšem tyto věty budou v textu od sebe příliš vzdáleny, bude pro RNN problém dát je do souvislosti. LSTM toto řeší pomocí vstupní brány (input gate), výstupní brány (output gate) a brány zapomenutí (forget gate). Tyto brány řídí tok informací, které jsou potřebné k predikci výstupu v síti.



Obr. 33: Principiální schéma LSTM algoritmu

Princip algoritmu je velmi dobře patrný na obrázku 33. V prvním kroku LSTM rozhodne, jakou informaci z buňky vyloučit pomocí sigmoidní vrstvy σ_g (forget gate). Pomocí stavového vektoru \mathbf{h}_{t-1} a vstupního vektoru do jednotky LSTM \mathbf{x}_t přiřadí číslo mezi 0 (*zahodit*) a 1 (*uchovat*) pro každé číslo ve stavu buňky \mathbf{C}_{t-1} . Aktivační funkci $\mathbf{f}(t)$ lze zapsat jako

$$\mathbf{f}(t) = \sigma_g(\mathbf{W}_f \mathbf{x}_t + \mathbf{U}_f \mathbf{h}_{t-1} + \mathbf{b}_f).$$

Dalším krokem je rozhodnutí, jaké nové informace uložit. Nejprve sigmoidní vrstva σ_g rozhodne, které hodnoty budou aktualizovány. Aktivační vektor této vrstvy \mathbf{i}_t lze zapsat jako

$$\mathbf{i}_t = \sigma_g(\mathbf{W}_i \mathbf{x}_t + \mathbf{U}_i \mathbf{h}_{t-1} + \mathbf{b}_i).$$

Poté vrstva \tanh vytvoří vektor nových požadovaných hodnot $\tilde{\mathbf{C}}_t$, který lze vyjádřit jako

$$\tilde{\mathbf{C}}_t = \tanh(\mathbf{W}_c \mathbf{x}_t + \mathbf{U}_c \mathbf{h}_{t-1} + \mathbf{b}_c).$$

Následně se z těchto hodnot vytvoří aktualizace stavu, a poté je nutno aktualizovat stav buňky \mathbf{C}_{t-1} na nový stav buňky \mathbf{C}_t

$$\mathbf{C}_t = \mathbf{f}_t \circ \mathbf{C}_{t-1} + \mathbf{i}_t \circ \tilde{\mathbf{C}}_t,$$

kde \circ značí Hadamardův součin. Výstup bude vycházet ze se stavu buňky, kdy se nejprve aktivuje sigmoidní vrstva σ_g , která rozhodne, které členy lze vyloučit. Poté vrstva \tanh

opět vytvoří nové hodnoty a po vynásobení výstupem sigmoidního hradla jsou stanoveny požadované hodnoty. Aktivační výstupní funkci \mathbf{o}_t lze zapsat ve tvaru

$$\mathbf{o}_t = \sigma_g(\mathbf{W}_o \mathbf{x}_t + \mathbf{U}_o \mathbf{h}_{t-1} + \mathbf{b}_o).$$

Vektor skrytého stavu \mathbf{h}_t lze zapsat ve tvaru

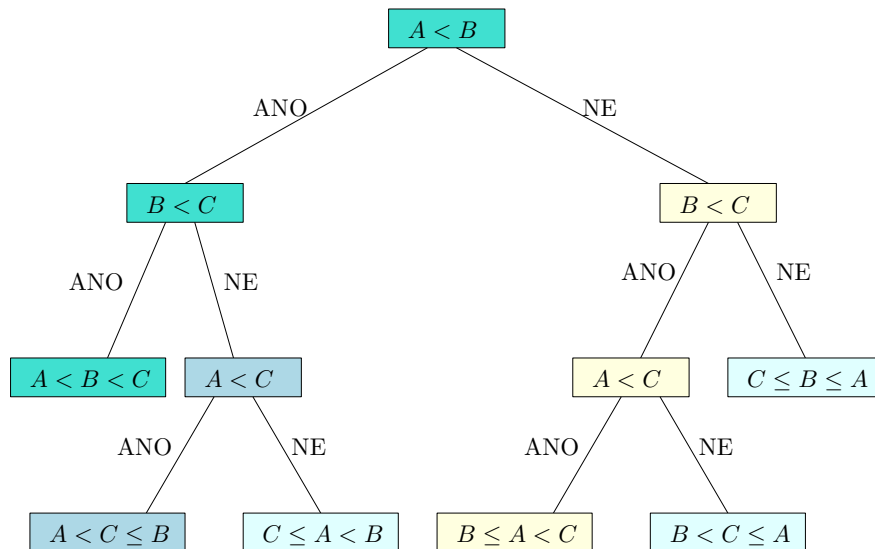
$$\mathbf{h}_t = \mathbf{o}_t \circ \tanh(\mathbf{C}_t),$$

kde \mathbf{W} a \mathbf{U} jsou váhové matice a \mathbf{b} je vektor předpětí (bias). Symboly σ_g označují sigmoidní aktivační funkci a \tanh je hyperbolická tangenciální aktivační funkce.

Gated Recurrent Units (GRU) je varianta algoritmu, která je podobná LSTM, protože také řeší problém s krátkodobou pamětí. Oproti LSTM má místo tří bran dvě — resetovací a aktualizací. Podobně jako v rámci LSTM, tyto brány řídí, jaké informace a v jakém množství se mají uchovávat. Dalším poměrně často používaným algoritmem je zjednodušený algoritmus *SimpleRNN*.

2.3.4 Rozhodovací stromy

Rozhodovací stromy (Decision Trees) jsou neparametrickou regresní metodou. Jedná se o metodu klasifikační, kdy lze hodnotu cílové funkce předpovědět na základě učení jednoduchých rozhodovacích pravidel. Lze si představit, že každý vnitřní uzel stromu je podmínkou a každý list stromu je následné rozhodnutí, viz obr. 34.

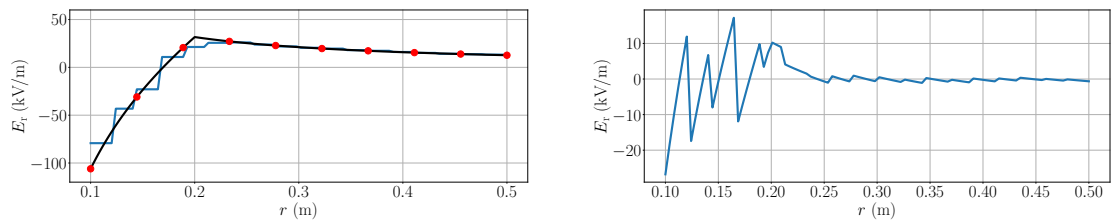


Obr. 34: Princip klasifikační metody rozhodovacích stromů

Výhodou použití rozhodovacích stromů je snadná pochopitelnost a interpretace výsledků, přičemž je zapotřebí jen velmi malé množství dat. Klíčové pro tvorbu rozhodovacích stromů je správná volba charakteristického atributu, podle kterého se budou data vyhodnocovat.

Jiné techniky totiž často vyžadují datovou normalizaci, kdy je zapotřebí vytvoření fiktivní

proměnné. Rozhodovací stromy se vyznačují dobrou proveditelností, i když příslušné předpoklady jsou poněkud ovlivněny skutečným modelem, z něhož byla data generována. Tato metoda je ale ne vždy stabilní, i malé odchylky v datech by mohly mít za následek generaci stromu zcela odlišného. Tento problém se dá zmírnit použitím rozhodovacích stromů uvnitř souboru, to znamená, že je zapotřebí pro kořenový uzel rozhodovacího stromu vybrat takový atribut, díky kterému půjdou od sebe jednotlivé objekty co možná nejvíce odlišit. Praktické algoritmy pro rozhodovací strom jsou založeny na heuristických algoritmech, kde jsou lokálně optimální rozhodnutí prováděna v každém uzlu. Tyto algoritmy ale nemohou zaručit návratnost globálně optimálního rozhodovacího stromu. Aproximace rovnoměrně navzorkovaného prostoru pomocí klasifikační metody rozhodovacích stromů je znázorněno na obr. 35.



Obr. 35: Rozhodovací stromy - aproximace rovnoměrně navzorkovaného prostoru s počtem vzorových bodů $N = 15$ (vlevo) a odchylka aproximace (vpravo)

Shrnutí

Strojové učení je velmi silným nástrojem pro vytváření aplikací, které využívají ke spolupráci umělé inteligence. Kromě výše uvedených metod existují tisíce nejrůznějších algoritmů, které strojové učení využívá. I přes obrovské množství metod je jejich hlavní myšlenka stejná - na základě dat známých z reálného systému se jej pokusit popsat pomocí správného matematického vztahu tak, aby jej bylo možné pochopit v jeho úplnosti s následnou schopností predikce budoucnosti.

Ačkoliv se dnes metody strojového učení používají vcelku běžně, výběr vhodné metody pro konkrétní problém není zcela jednoduchou věcí. Ve většině případů se však lze spokojit s predikcí s určitou chybou, která vždy bude ovlivněna volbou metody s ohledem na řešený problém. V tabulce 2 je znázorněno porovnání úspěšnosti predikce výše uvedených metod regresního a klasifikačního přístupu.

Tab. 2: Odchylka intenzity pole náhradního modelu od analytického řešení pro počet vzorových bodů $N = 10, 15, 20$

Metoda ($N = 10$)	rovnoměrné dělení (kV/m)	náhodné dělení (kV/m)	LHS (kV/m)	Haltonova sekvence (kV/m)
Kriging (SMT)	2,19	5,25	3,41	6,89
Rozhodovací stromy (Scikit)	5,65	15,80	11,61	12,43
MLP (Scikit)	3,86	3,46	25,27	25,70
Gaussův proces (Scikit)	2,40	174,50	4,67	73,16
Metoda ($N = 15$)	rovnoměrné dělení (kV/m)	náhodné dělení (kV/m)	LHS (kV/m)	Haltonova sekvence (kV/m)
Kriging (SMT)	0,72	10,85	1,9	5,58
Rozhodovací stromy (Scikit)	4,90	14,11	8,81	9,60
MLP (Scikit)	2,42	26,86	25,20	10,41
Gaussův proces (Scikit)	0,87	13,27	9,57	53,24
Metoda ($N = 20$)	rovnoměrné dělení (kV/m)	náhodné dělení (kV/m)	LHS (kV/m)	Haltonova sekvence (kV/m)
Kriging (SMT)	3,21	88,06	3,18	3,32
Rozhodovací stromy (Scikit)	3,81	35,03	7,45	7,54
MLP (Scikit)	2,84	39,77	2,87	2,77
Gaussův proces (Scikit)	0,74	279,51	9,17	1,66

OPTIMALIZAČNÍ METODY

Optimalizační metody se v současnosti využívají pro řešení problémů v širokém spektru vědních disciplín. Oprávněně je jim proto věnována stále větší pozornost. Tato kapitola bude zaměřena na představení nejčastěji využívaných optimalizačních metod.

3.1 DEFINICE OPTIMALIZAČNÍ ÚLOHY

Obecný optimalizační problém lze popsat jako minimalizaci cílové funkce $f(x)$ pomocí změny parametrů nad množinou X

$$\min f(x) \quad x \in X.$$

Lze předpokládat minimalizační a maximalizační problém zároveň? Pokud uvažujeme

$$\max f(x) \quad x \in X,$$

pak lze předpokládat, že tento problém je ekvivalentní pro

$$\min -f(x) \quad x \in X.$$

V praxi jsou většinou výsledky odvozeny pouze pro minimalizační problémy. Pokud je potřeba řešit maximalizační problém, převedeme jej na optimalizační problém pomocí minimalizace záporné hodnoty funkce $f(x)$.

Ukažme principiální řešení optimalizačního příkladu na konkrétní funkci, hledejme tedy minimum funkce $f(x, y)$, kterou lze vyjádřit ve tvaru:

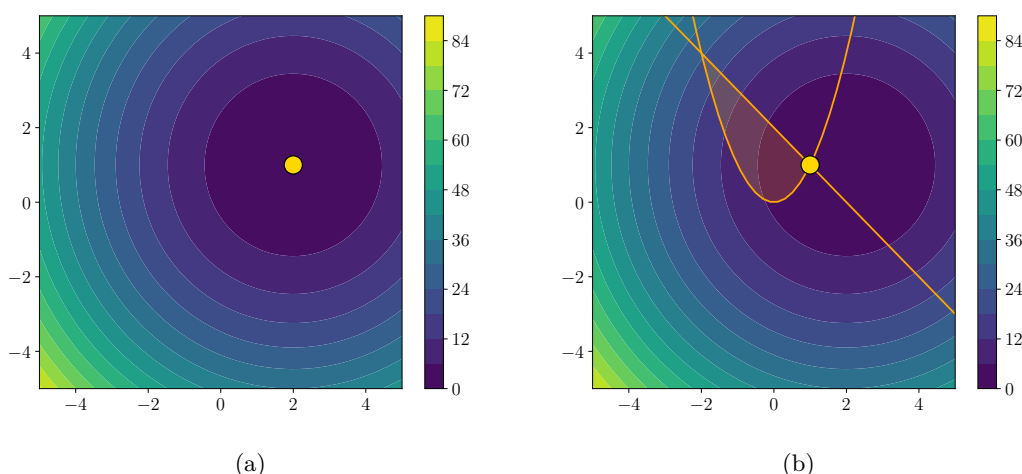
$$f(x, y) = (x - 2)^2 + (y - 1)^2 \quad x \in (-5; 5), y \in (-5; 5)$$

Tato funkce nemá na celém definičním oboru žádné omezení a hledaným optimem je bod $[2; 1]$, viz. obr. 36 a), ve kterém je dosaženo minima

$$f(2, 1) = 0.$$

Optimalizace s vazbou

Pokud máme pro řešení optimalizačního problému předem daná omezující optimalizační kritéria, pak hovoříme o optimalizačním problému s vazbou. V praxi mohou být tato omezení vztažena například na provozní teplotu materiálu, předem dané požadavky na rozměry, apod.



Obr. 36: Optimum pro funkci a) bez omezení a za b) s omezením

Optimalizační proces s vazbou (omezeními) lze zapsat ve tvaru

$$\begin{array}{ll} \min & f(x, y) & x \in X \\ \text{za podmíněk} & g(x) = 0 & \text{omezení typu rovnost} \\ & h(x) \leq 0 & \text{omezení typu nerovnost,} \end{array}$$

kde funkce $g(x)$ představuje omezení vstupních parametrů typu rovnost a funkce $h(x)$ typu nerovnost. Těmito podmínkami lze omezit oblast možných řešení úlohy.

Omezme tedy funkci $f(x, y)$:

$$\begin{array}{ll} \min & (x - 2)^2 + (y - 1)^2 & x \in (-5, 5) \\ & & y \in (-5, 5) \\ \text{za podmíněk} & & x + y \leq 2 \\ & & x^2 - y \leq 0 \end{array}$$

viz obr. 36 b), kde nyní funkce nabývá optima v bodě $[1; 1]$, cílová funkce v optimálním bodě nabývá hodnoty:

$$f(1, 1) = (1 - 2)^2 + (1 - 1)^2 = 1.$$

Metoda Lagrangeových multiplikátorů

V případě omezení typu rovnost lze úlohu řešit jako multikriteriální optimalizační problém převedením do tvaru

$$\mathcal{L}(x, \lambda) = f(x) - \lambda g(x).$$

Metodu Lagrangeových multiplikátorů si lze ukázat na příkladu pro určení rozměrů válcové nádoby o objemu 0,5 litru tak, aby vnitřní povrch nádoby byl co nejmenší. Povrch a objem nádoby určíme ze vztahů

$$P = \pi R^2 + 2\pi Rv,$$

$$V = \pi R^2 v,$$

kde R je poloměr a v je výška nádoby. Povrch P uvažujeme jako cílovou funkci $f(R, v)$ a pro objem V jako funkce $g(R, v)$ musí platit rovnost

$$g(R, v) - 0,5 = 0.$$

Potom lze formulovat Lagrangeovu funkci $\mathcal{L}(R, v, \lambda)$ jako novou cílovou funkci

$$\mathcal{L}(R, v, \lambda) = \pi R^2 + 2\pi Rv + \lambda(\pi R^2 v - 0,5).$$

Řešení metody s vazbou

Provedeme parciální derivace podle R, v, λ a jejich porovnáním s nulou určíme kritické body:

$$\frac{\partial \mathcal{L}}{\partial R} = 2\pi R + 2\pi v + \lambda 2\pi Rv = 0,$$

$$\frac{\partial \mathcal{L}}{\partial v} = 2\pi R + \lambda \pi R^2 = 0,$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = \pi R^2 v - 0,5 = 0.$$

Obdrželi jsme soustavu tří rovnic o třech neznámých, kterou lze vcelku jednoduše vyřešit. Úpravou druhé rovnice obdržíme rovnici

$$\frac{\partial \mathcal{L}}{\partial v} = \pi R(2 + \lambda R) = 0,$$

jejímž řešením jsou hodnoty $R_1 = 0$ a $R_2 = -\frac{2}{\lambda}$, kde ovšem první řešení nedává z interpretačního hlediska žádný smysl.

Dosazení vztahu pro koeficient R_2 do první a třetí rovnice vede na soustavu dvou rovnic:

$$\frac{\partial \mathcal{L}}{\partial R} = 2\pi \left(-\frac{2}{\lambda}\right) + 2\pi v + \lambda 2\pi \left(-\frac{2}{\lambda}\right) v = 0,$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = \pi \left(-\frac{2}{\lambda}\right)^2 v - 0,5 = 0.$$

Z první rovnice si lze nyní vyjádřit vztah pro výšku řešeného válce v

$$-2 + \lambda v - 2\lambda v = 0$$

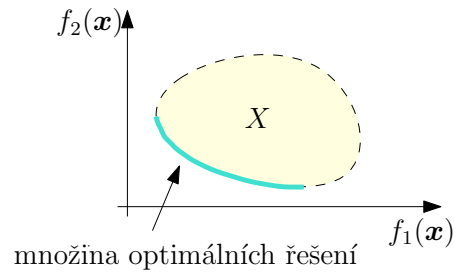
$$v = -\frac{2}{\lambda} = R_2,$$

jehož dosazením do třetí rovnice obdržíme řešení pro výšku v

$$\pi v^2 v - 0,5 = 0$$

$$v = \sqrt[3]{\frac{0,5}{\pi}}.$$

Řešení soustavy rovnic pro poloměr je $R_2 = v$, kde $v = \sqrt[3]{\frac{0,5}{\pi}}$.



Obr. 37: Příklad Paretoovy fronty

3.2 KLASIFIKACE OPTIMALIZAČNÍCH ÚLOH

Uvažujme nyní optimalizační problém s k proměnnými

$$\min f_k(x) \quad x \in X, k \in \{1, \dots, m\}.$$

Pokud pro $f_k(x)$ platí

- $m = 1$, pak hovoříme o **monokriteriální** optimalizační úloze, tato úloha při hledání optima zohledňuje pouze jedno kritérium,
- $m > 1$, pak hovoříme o **multikriteriální** optimalizační úloze, pro tuto úlohu existuje více kritérií.

Na rozdíl od monokriteriálního problému, multikriteriální problém může mít nespočetné množství optimálních řešení, která zpravidla vytvoří tzv. **Paretovu frontu**, z níž se vybere takové řešení, kde jsou hodnoty všech cílových funkcí přijatelné z hlediska rozhodovacích faktorů:

- pro dvě kriteriální funkce $m = 2$ je řešením křivka v rovině (Paretova fronta je zobrazena na obrázku 37 modrou barvou),
- pro tři cílové funkce $m = 3$ je plocha ve 3D prostoru atd.

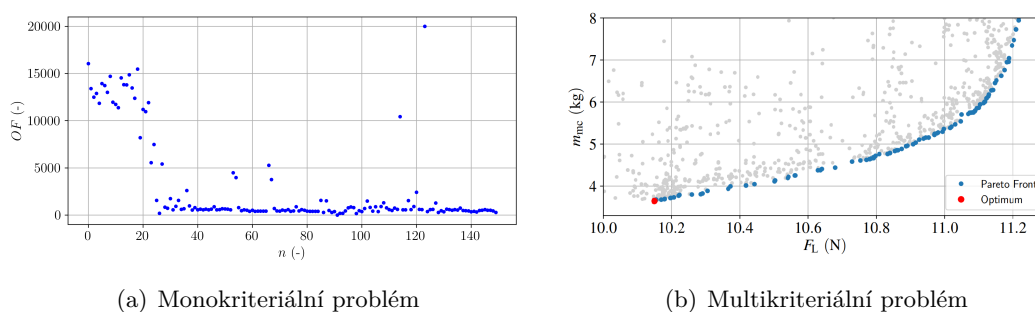
Na obr. 38 jsou znázorněny příklady možné interpretace průběhu řešení optimalizačního problému. V jeho části a) je znázorněno řešení pro **monokriteriální optimalizační problém**, kde

- cílová funkce je minimalizována,
- průběh je zobrazen pro závislost cílové funkce na počtu iterací.

Řešení pro **multikriteriální optimalizaci** je patrné z části b), kde

- maximalizujeme první cílovou funkci, minimalizujeme druhou cílovou funkci a
- optimální výsledky jsou zobrazeny na Pareto frontě.

Obecně existuje mnoho způsobů klasifikace optimalizačních úloh. V technické praxi lze optimalizační problémy rozlišovat dle [83] například na základě vstupních parametrů pro optimalizaci. Jedním z přístupů je optimalizace zdrojových nebo materiálových parametrů, druhou možností je optimalizace geometrie zařízení a jako zvláštní skupinu lze uvažovat řešení inverzních úloh s využitím optimalizačních algoritmů.



Obr. 38: Ilustrativní příklady možné prezentace řešení pro jednotlivé optimalizační problémy

Optimalizace zdrojových veličin

Jako optimalizaci zdrojových veličin lze uvažovat takový optimalizační problém, kdy jsou známé výsledky z experimentu, ovšem nejsou přesně známy parametry experimentálního zařízení (geometrie, materiálové charakteristiky, apod.). Tyto veličiny jsou pak předmětem optimalizace.

Optimalizace geometrie

Optimalizaci tvaru geometrie lze rozlišovat z pohledu přístupu ke změnám geometrie, viz obr. 39. K tomu lze využít tři optimalizační principy změn, uvedené v [83]:

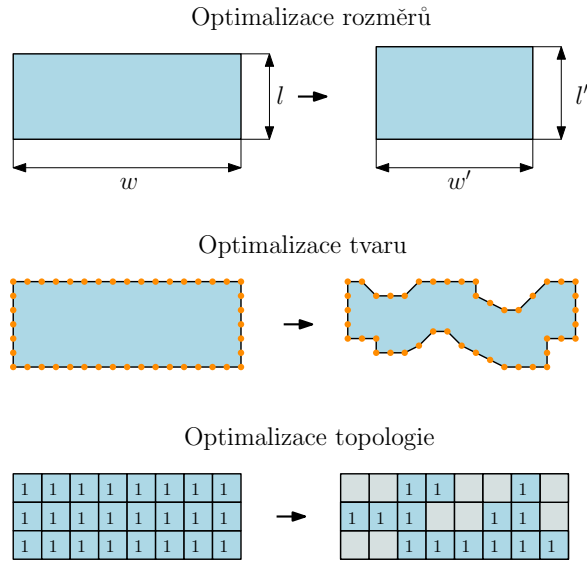
1. **rozměrů** - parametry optimalizace jsou rozměry zařízení, je nejjednodušší z uvedených přístupů. Velkou výhodou je nízká výpočetní náročnost, na druhou stranu jsme limitováni omezeným množstvím řešení, kdy jsme omezeni na předem daný tvar zařízení.
2. **tvaru** - parametry optimalizace jsou pozice bodů, které jsou rozloženy na hranicích geometrie popisující zařízení. Tyto body jsou určeny pomocí zadané volnosti a během procesu optimalizace dochází ke změnám souřadnic těchto bodů. Tvarová optimalizace je sice výpočetně náročnější vzhledem ke stoupající dimenzi úlohy, než je úloha předchozí, ale poskytuje větší variabilitu množiny optimálních řešení.
3. **topologie** - optimalizace rozložení materiálu na základě rozdělení řešené oblasti na konečný počet podoblastí, které během procesu optimálně vyplní danou optimalizovanou oblast. Jedná se o výpočetně nejnáročnější přístup z uvedených, ovšem zaručuje velkou variabilitu optimálních řešení, jelikož lze pracovat i s rozložením uvnitř geometrie.

Inverzní úloha

Pokud je předem známé řešení problému a je potřeba určit vhodnou kombinaci vstupních faktorů, která poskytne právě takové řešení, hovoříme o inverzní úloze, kterou lze zapsat jako

$$\mathbf{y} = M(\mathbf{x}), \quad (10)$$

kde \mathbf{y} představuje vektor dat předem známého řešení, \mathbf{x} je vektor vstupních parametrů a M je operátor zavedený pro řešený matematický model. Cílem inverzní úlohy je určení vektoru \mathbf{x} nebo operátoru M tak, aby výše uvedená rovnice platila.



Obr. 39: Možnosti optimalizace geometrie zařízení [83]

Optimalizační problém lze potom formulovat jako minimum cílové funkce v obecném tvaru, kdy

$$\min f(x) = |y - Mx|. \quad (11)$$

Tento přístup je využit při řešení problémů v ilustrativních příkladech v následujících kapitolách, které budou věnovány vyhledávání

- maximální dovolené provozní teploty zařízení a
- rovnoměrného rozložení hodnoty proudů na výstupu zařízení.

3.3 METODY OPTIMALIZACE

V současné době je již k dispozici velké množství numerických iteračních postupů vedoucích většinou ke spolehlivému a dostatečně přesnému řešení z hlediska rozhodovacích faktorů. K ukončení zmíněného iteračního procesu dojde po splnění některého konvergenčního kritéria, například když se změny vstupních parametrů od iterace k iteraci stanou nevýznamnými.

Optimalizační metody lze dělit z mnoha hledisek. Pro řešení numerických složitých problémů se dnes využívá velké množství numerických iteračních postupů vedoucích většinou ke spolehlivému a dostatečně přesnému řešení z hlediska rozhodovacích faktorů.

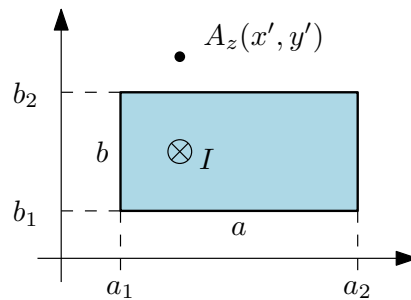
Vzhledem k rozsahu práce není možné analyzovat do této doby desítky navržených a funkčních optimalizačních metod. V dalším textu proto budou popsány metody, které jsou implementovány v komerčním programu COMSOL Multiphysics a vlastní aplikaci Artap a které byly použity pro řešení optimalizačních problémů, na něž je tato dizertační práce zaměřena. Řešení trojrozměrné úlohy bylo prováděno převážně v aplikaci COMSOL Multiphysics, ostatní úlohy jsou řešeny v kombinaci Artapu s tímto programem, dále je využito prostředí Agros Suite. V dalším textu tedy budou ve stručnosti popsány pouze techniky, které se v těchto balíčcích nacházejí.

Numerické metody je možné využít pro řešení optimalizačních úloh, kde cílovou funkci nelze vyjádřit pomocí uzavřené formule. Rozdělení optimalizačních metod lze dělit například na:

1. **Gradientní metody a metody s aproximací gradientu** - informace o změně funkce $f(x)$ pro nalezení optimálního směru vyhledávání.
2. **Stochastické metody** (evoluční metody, metody inspirované přírodou) - oproti předšlým metodám se jedná o nahodilý proces vyhledávání.
3. **Ostatní**

Různé metody vedou k různým výsledkům. Některé metody mohou zkonvergovat k lokálnímu minimu, na druhou stranu je jejich vyhodnocení poměrně rychlé. Jedná se zejména o metody založené na výpočtu gradientu. Stochastické algoritmy jsou velmi pomalé a výsledky optimalizace se mohou lišit pro jednotlivé běhy. Uvedené metody budou ukázány na ilustrativním příkladu řešení soustavy vodičů.

Uvažujme nekonečně dlouhý vodič obdélníkového průřezu protékaný harmonickým proudem I (viz obr. 40).



Obr. 40: Principiální schéma uspořádání (jeden pásový vodič)

Matematický model a simulace úlohy

Rozložení magnetického vektorového potenciálu v obecném bodě x' a y' v okolí takového vodiče lze zapsat ve tvaru

$$A_z(x', y') = \frac{\mu_0 I}{4\pi ab} \int_{b_1}^{b_2} \int_{a_1}^{a_2} \ln \left[(x - x')^2 + (y - y')^2 \right] dx dy.$$

Magnetickou indukci lze vyjádřit jako

$$\mathbf{B}(x', y') = \text{rot } \mathbf{A}(x', y').$$

Tento integrál lze řešit numericky nebo rovněž snadno analyticky. Analytické řešení bylo provedeno s využitím programu Wolfram Mathematica.

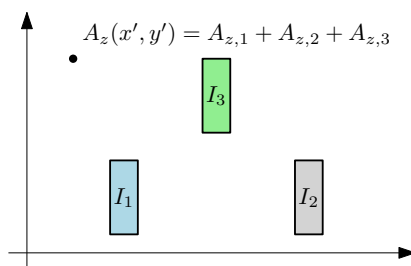
$$A_z(x', y') = \frac{\mu_0 I}{4\pi(a_2 - a_1)(b_2 - b_1)} \cdot (b_1 - y)^2 \arctg \frac{a_1 - x}{b_1 - y} - (b_1 - y)^2 \arctg \frac{a_2 - x}{b_1 - y} - (b_2 - y)^2 \arctg \frac{a_1 - x}{b_2 - y} + (b_2 - y)^2 \arctg \frac{a_2 - x}{b_2 - y}$$

$$\begin{aligned}
& (a_1 - x) \left[-3b_1 + y + (a_1 - x) \operatorname{arctg} \frac{b_1 - y}{a_1 - x} + (b_1 - y) \ln \left[(a_1 - x)^2 + (b_1 - y)^2 \right] \right] - \\
& (a_2 - x) \left[-3b_1 + y + (a_2 - x) \operatorname{arctg} \frac{b_1 - y}{a_2 - x} + (b_1 - y) \ln \left[(a_2 - x)^2 + (b_1 - y)^2 \right] \right] - \\
& (a_1 - x) \left[-3b_2 + y + (a_1 - x) \operatorname{arctg} \frac{b_2 - y}{a_1 - x} + (b_2 - y) \ln \left[(a_1 - x)^2 + (b_2 - y)^2 \right] \right] + \\
& (a_2 - x) \left[-3b_2 + y + (a_2 - x) \operatorname{arctg} \frac{b_2 - y}{a_2 - x} + (b_2 - y) \ln \left[(a_2 - x)^2 + (b_2 - y)^2 \right] \right]. \quad (12)
\end{aligned}$$

Nyní uvažujme 3 takové vodiče (obr. 41) protékané proudy s fázovým posuvem 120° . Rozložení magnetického vektorového potenciálu v obecném bodě x' a y' lze zapsat jako superpozici příspěvků od všech tří vodičů

$$A_z(x', y') = A_{z,1}(x', y') + A_{z,2}(x', y') + A_{z,3}(x', y').$$

Uvažujme proud o amplitudě $I_0 = 1000$ A a frekvenci 50 Hz v čase 0,06367 s. Rozměry vodičů uvažujme dle tabulky 3.



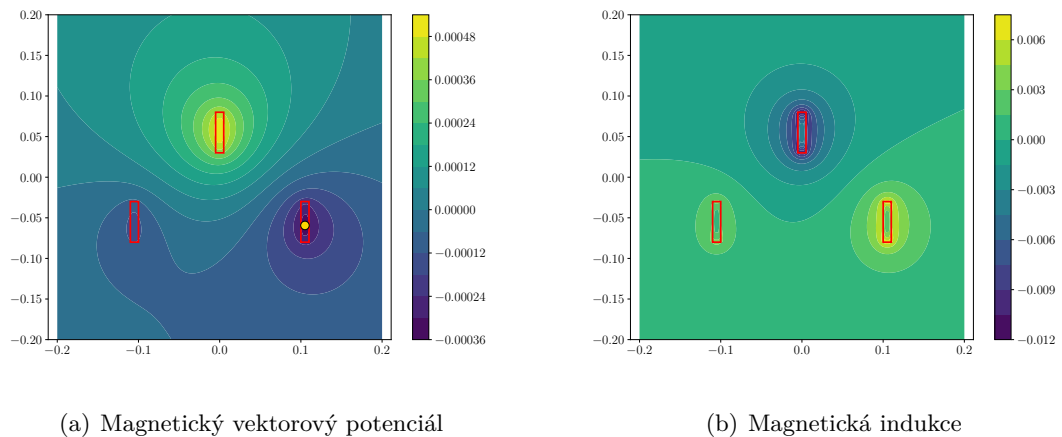
Obr. 41: Principiální schéma uspořádání (tři pásově vodiče)

Tab. 3: Proudů v jednotlivých vodičích

	a_1 (m)	a_2 (m)	b_1 (m)	b_2 (m)	i (A)
Vodič 1	-0,110	-0,100	-0,080	-0,030	$I_0 \cdot \sin(\omega t)$
Vodič 2	0,100	0,110	-0,080	-0,030	$I_0 \cdot \sin(\omega t + 120^\circ)$
Vodič 3	-0,005	0,005	0,030	0,080	$I_0 \cdot \sin(\omega t + 240^\circ)$

Na obr. 42 je znázorněno řešení soustavy vodičů pro magnetický vektorový potenciál s vyznačeným optimem v bodě $[0, 1; -0, 08]$ a rozložení magnetické indukce. Cílem optimalizačního procesu bude nalezení minimální hodnoty magnetického vektorového potenciálu. Potom cílovou funkci zapíšeme ve tvaru

$$\min A_z(x, y), \quad x \in (-0, 2; 0, 2), \quad y \in (-0, 2; 0, 2).$$



Obr. 42: Rozložení vektorového potenciálu s vyznačeným optimem vlevo a rozložení modulu magnetické indukce vpravo

3.3.1 Metody využívající informace o změně funkce

U jednorozměrných (a případně vícerozměrných) problémů bez omezení prozkoumají vyhledávací metody prostor parametrů iterativně úpravou směru a rozsahu vyhledávání v každé iteraci, aby našly nižší a nižší hodnoty pro cílovou funkci.

Existuje několik iteračních vyhledávacích metod, obvykle nazývaných *line search methods*, které jsou navrženy tak, aby řešily jednorozměrné unimodální neomezené optimalizační problémy [84]. Jelikož gradientní metody používají informace o změně funkce $f(x)$ pro nalezení optimálního směru vyhledávání, jsou tím pádem účinné při hledání lokálního optima pro hladké jednorázové problémy s kontinuální doménou. Problém může nastat v případě nespojitě funkce, kde nelze zajistit získání optimální informace, nebo v případě, že cílová funkce $f(x)$ má multimodální charakter; v tomto případě hrozí, že iterační proces se ukončí v některém z případných lokálních extrémů.

Použitím optimální gradientní metody je zaručena konvergence k lokálnímu minimu z jakékoli výchozí hodnoty. V praxi existují metody vyhledávání směru, které používají různá kritéria pro volbu vhodné velikosti kroku, čímž se zabrání příliš dlouhým nebo naopak krátkým krokům ke zvýšení efektivity tohoto procesu.

Gradientní metody poklesu, tzv. (gradient descent methods) lze rozdělit do dvou skupin, přičemž každá z vyjmenovaných skupin má své výhody a nevýhody a vhodný výběr algoritmu záleží na konkrétní úloze [85]:

1. **Metody s výpočtem gradientu** (gradient-based methods) - Metoda gradientního sestupu, SNOPT,
2. **Metody založené na aproximaci gradientu** (gradient-free methods) - BOBYQA, Nelder-Mead.

Optimalizace pomocí metod založených na aproximaci gradientu nepoužívá klasické derivace pro nalezení optimálních řešení, jelikož řešení nelze popsat pomocí spojité analytické funkce. Jestliže si extremalizační úlohu převedeme na nalezení minima cílové funkce (jedná-li se

o maximum, stačí přidat záporné znaménko) a písmenem A označíme množinu dovolených vstupních parametrů, hledáme v každé iteraci takové $x_k \in A$, které splňuje podmínku

$$f(x_k) \leq f(x_{k-1}) \quad \text{pro} \quad x_k, x_{k-1} \in A$$

V tomto případě $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$ a podobný vztah platí i pro \mathbf{x}_0 , jež označuje počáteční bod iteračního procesu.

V případě metod založených na aproximaci gradientu jsou tedy místo derivací k řešení použity pouze hodnoty funkce $f(x)$. Pomocí těchto metod lze většinou dosáhnout požadovaného optima, ovšem v některých případech je řešení problému mnohem složitější a metody založené na principu porovnání konvexní optimalizace selhávají. Pro tyto účely je potřeba efektivně nalézt vhodné hodnoty parametrů, které mohou být téměř optimální (suboptimální).

Metoda sestupného gradientu

Pokud je výpočet založen na principu sestupného směru gradientu, odpovídající iterační vyhledávání lze nazvat **metodou sestupného gradientu** (známou jako nejprudší sestup nebo **Cauchyho metoda**). Směr záporného gradientu, při němž se cílová funkce snižuje co možná nejrychleji, je nejpřirozenější možností řešení, protože je to směr nejstrmějšího klesání hodnoty dané funkce. Vzhledem k tomu, že algoritmus kromě hodnoty funkce využívá i její gradient, pak funkce $f(x)$ musí být diferencovatelnou funkcí.

V příslušném algoritmu se využívá nastavení směru vyhledávání a optimální velikosti kroku. Matematicky lze tuto myšlenku zapsat v explicitním tvaru pomocí výrazu

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \alpha^k \cdot \text{grad } f(\mathbf{x}^k),$$

kde $f(\mathbf{x})$ je cílová funkce, α je délka kroku a k označuje pořadové číslo iterace. Vektor \mathbf{x} odpovídá jednotlivým osám v souřadnicovém systému.

```

1 # Základní algoritmus Gradient Descent
2 x_opt = []
3 y_opt = []
4 for i in range(steps):
5     x_opt.append(xt)
6     y_opt.append(yt)
7
8     [step_x, step_y] = gradient_descent_step(xt, yt, alpha)
9
10    xt = xt + step_x
11    yt = yt + step_y
12
13 def gradient_descent_step(xt, yt, alpha):
14    [gx, gy] = fce_gradient(xt, yt)
15
16    step_x = - alpha * gx
17    step_y = - alpha * gy
18    return [step_x, step_y]
```

Algoritmus 3.1: Základní algoritmus metody sestupného gradientu

Adaptivní gradientní sestup

Algoritmus s pevným krokem často není schopen najít hledané minimum. Jedna z možností vylepšení algoritmu je automatická volba kroku s využitím Armijovy podmínky ve tvaru

$$f(\mathbf{x}^k - \alpha^k \cdot \text{grad } f(\mathbf{x}^k)) \leq f(\mathbf{x}^k) - c \cdot \|\text{grad } f(\mathbf{x}^k)\|^2,$$

kde $c \in (0, 1)$ je malá konstanta. Často se pokládá $c = 10^{-4}$.

Vzhledem k tomu, že na levé straně nerovnice se nachází hodnota funkce v dalším iteračním kroku \mathbf{x}^{k+1} , podmínka předpokládá, že funkce bude vždy klesající. Tím se zabrání oscilacím. V praxi se často používá iterační způsob odhadu kroku α . Algoritmus lze nastartovat vhodnou volbou počáteční hodnoty kroku α_0 , která se v každém kroku půlí dokud je Armijova podmínka splněna.

```

1 # Adaptivní algoritmus Gradient Descent
2 x_opt = []
3 y_opt = []
4 for i in range(steps):
5     x_opt.append(xt)
6     y_opt.append(yt)
7
8     [step_x, step_y] = gradient_descent_adaptive_step(xt, yt, alpha, c)
9
10    xt = xt + step_x
11    yt = yt + step_y
12
13 def gradient_descent_adaptive_step(xt, yt, alpha, c):
14     [gx, gy] = fce_gradient(xt, yt)
15
16     alpha_t = alpha # Armijova podmínka
17     while fce(xt - alpha_t*gx, yt - alpha_t*gy) > fce_a_total(xt, yt) - c*alpha_t*(gx
18     **2 + gy**2):
19         alpha_t = alpha_t / 2.0
20         # zastavovací kritérium
21         if alpha_t < 1e-6:
22             break
23
24     step_x = - alpha_t * gx
25     step_y = - alpha_t * gy
26     return [step_x, step_y]

```

Algoritmus 3.2: Adaptivní algoritmus metody sestupného gradientu

Metoda sestupného gradientu s vazbou

Nejjednodušší metoda gradientního sestupu s vazbou jsou projektované gradienty

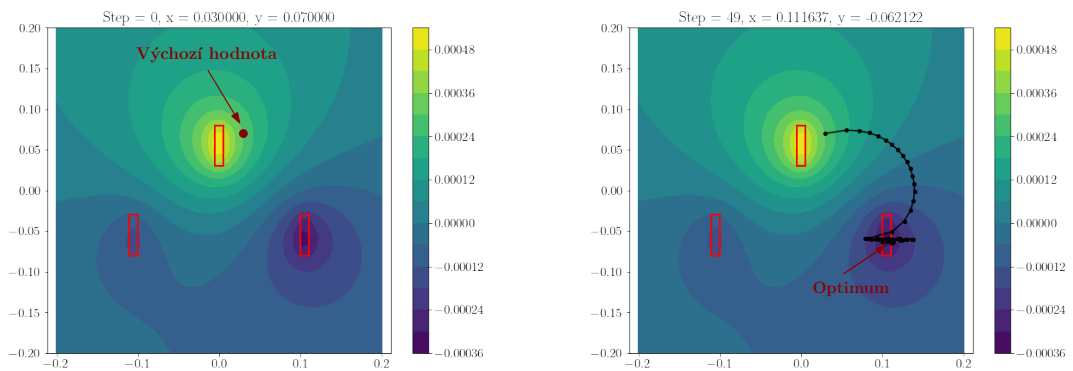
$$\begin{aligned} \mathbf{p}^{k+1} &= \mathbf{x}^k - \alpha^k \cdot \text{grad } f(\mathbf{x}^k), \\ \mathbf{x}^{k+1} &= P(\mathbf{p}^{k+1}). \end{aligned}$$

Nejprve se vypočte nový bod pomocí standardní metody gradientního sestupu a následně je zobrazen pomocí projekční funkce $P(x)$ do prostoru, který je vymezen vazbou. Omezení této metody je nutnost jednoduchého výpočtu transformace a proto je zpravidla omezena pouze na pravouhlé oblasti. Obecnější, ale výrazně pracnější na implementaci, je například již zmíněná metoda Lagrangeových multiplikátorů.

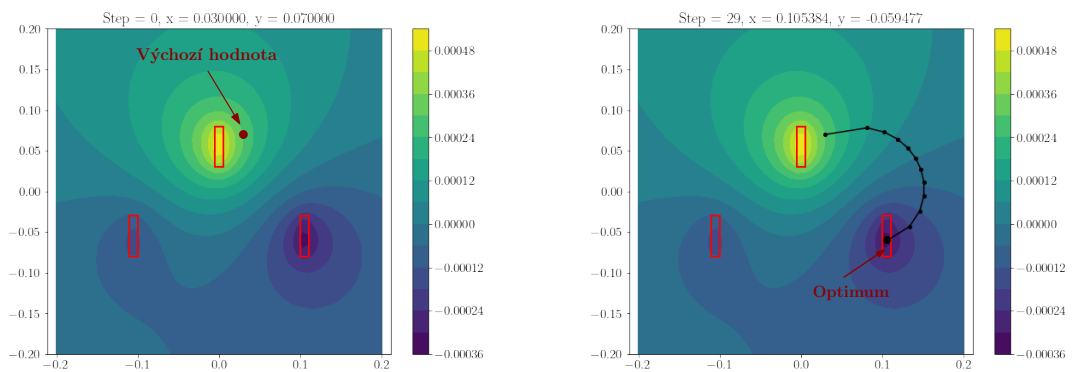
Příklady řešení problému pomocí metod sestupného gradientu jsou předvedeny na ilustrativním příkladu v kapitole 3.3. Výsledná řešení jsou patrná na obr. 43–46. Jak je z grafů patrné, pro úspěšné řešení optimalizačního problému metodou sestupného gradientu je důležitá

- vhodná volba kroku,
- vhodná volba výchozích hodnot.

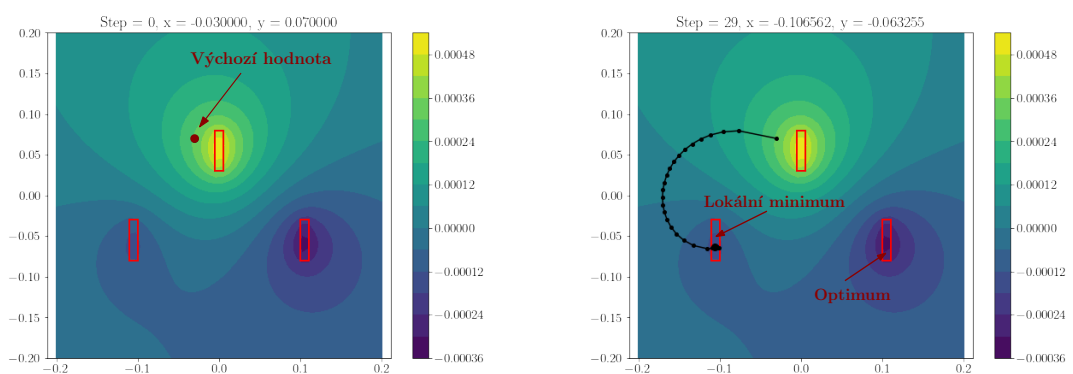
Pokud je zvolený krok optimalizačního algoritmu příliš malý, pravděpodobně algoritmus nakonec dosáhne přesného řešení, ale za cenu neúměrné časové náročnosti. V případě naopak



Obr. 43: Gradientní sestup – nevhodně zvolený krok

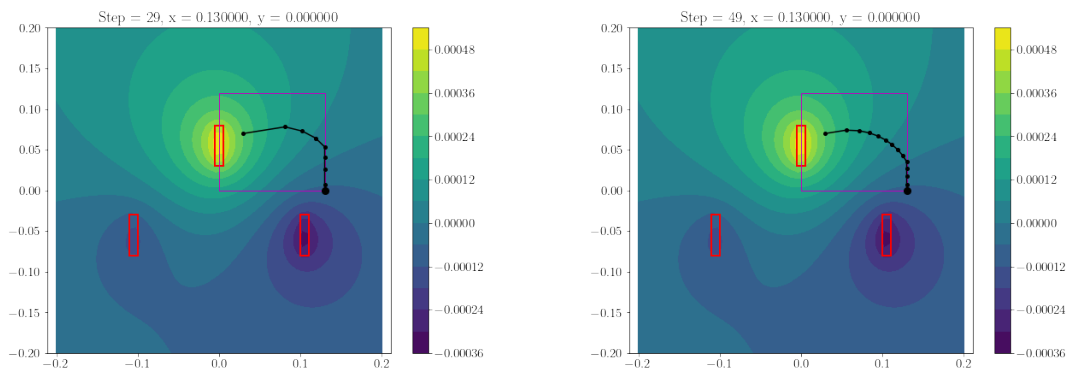


Obr. 44: Gradientní sestup – adaptivní krok



Obr. 45: Gradientní sestup – nevhodně zvolená počáteční poloha

většího kroku hrozí riziko nepřesného řešení. Tento problém je možné poměrně efektivně vyřešit volbou adaptivního kroku, viz obr. 44. Dalším úskalím gradientních metod je při



Obr. 46: Gradientní sestup s vazbou s adaptivním krokem (vlevo) a s velkým krokem (vpravo)

nevhodně zvolených výchozích hodnotách právě hrozba uvíznutí v lokálních minimech, viz obr. 43. V horní části obrázku je uveden počet kroků a pozice minima.

Sparse Nonlinear OPTimizer - SNOPT

Jedná se o univerzální systém pro rozsáhlou, nelineárně omezenou optimalizaci, která využívá sekvenční algoritmus kvadratického programování [86]. Metoda optimalizace vychází z flexibilní datové struktury, která zahrnuje celý problém optimalizace v rámci jedné proměnné. Hlavní funkce optimalizace vybírá správný řešitelský algoritmus, který závisí na vlastnostech cílové funkce a jejích omezeních. Tento optimalizační režim aplikace může skutečně řešit klasické čisté problémy s optimalizací; jeho skutečná síla je schopnost vypočítat přesné přechody, zejména když cílová funkce a omezení závisí na řešení modelu.

Nelder–Mead

Nelder–Mead (Downhill Simplex Method) [87] je heuristický algoritmus pro multikriteriální neomezené optimalizační problémy. Jedná se o jeden z nejběžněji používaných algoritmů a je základem pro mnoho optimalizačních technik v různých programovacích jazycích. K vyhodnocení minimální hodnoty cílové funkce $f(x)$ není potřeba informace o první nebo druhé derivaci, pracuje na principu aplikace operací trojúhelníkových simplexů $n + 1$ bodů v množině daných řešení (trojúhelníkový simplex ve dvou rozměrech, čtyřúhelníkový simplex ve třech rozměrech, atd.). Obsahuje funkce, které umožňují simplexu přizpůsobit se místnímu prostředí cílové funkce, tzn. při každé iteraci se simplex pohybuje směrem k minimu díky tomu, že provádí operace

1. zrcadlení (reflection)
2. protažení (expansion)
3. kontrakce (contraction)
4. zmenšení (shrink)

Kritérium pro ukončení výpočtu vychází ze standardní odchylky hodnoty funkce nad simplexem. Předpokládejme trojúhelníkový simplex ve 2D prostoru s vrcholy $x_1 = [0, 0]$, $x_2 = [2, 3]$, $x_3 = [4, 0]$. Pro daný trojúhelník můžeme určit těžiště:

$$\bar{x} = \frac{1}{3} \sum_{i=1}^3 x_i = [2, 1].$$

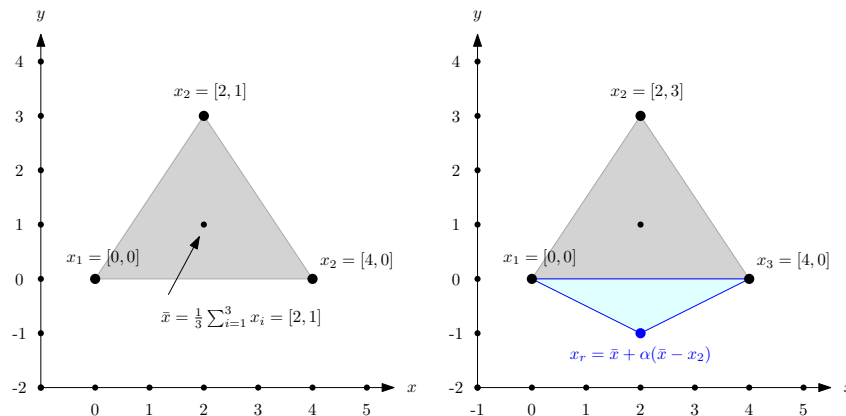
Algoritmus lze rozdělit do několika kroků

1. **Zrcadlení (Reflection):** Jako výchozí bod v rámci algoritmu uvažujeme x_2 . Zrcadlením tohoto bodu přes těžiště simplexu \bar{x} (obr. 47) vytvoříme nový bod pomocí vztahu

$$x_r = \bar{x} + \alpha(\bar{x} - x_i),$$

kde x_i je bod, který zrcadlíme, v tomto případě

$$x_r = \bar{x} + \alpha(\bar{x} - x_2) = [2, 1] + \alpha([2, 1] - [2, 3]) = [2, -1].$$



Obr. 47: Nelder–Mead – určení těžiště simplexu (vlevo) a operace zrcadlení (vpravo)

2. **Protážení (Expansion):** Nově vzniklý bod x_r (obr. 48) nyní protáhneme dále od těžiště \bar{x} pomocí vztahu

$$x_e = \bar{x} + \beta(x_r - \bar{x}).$$

V našem případě

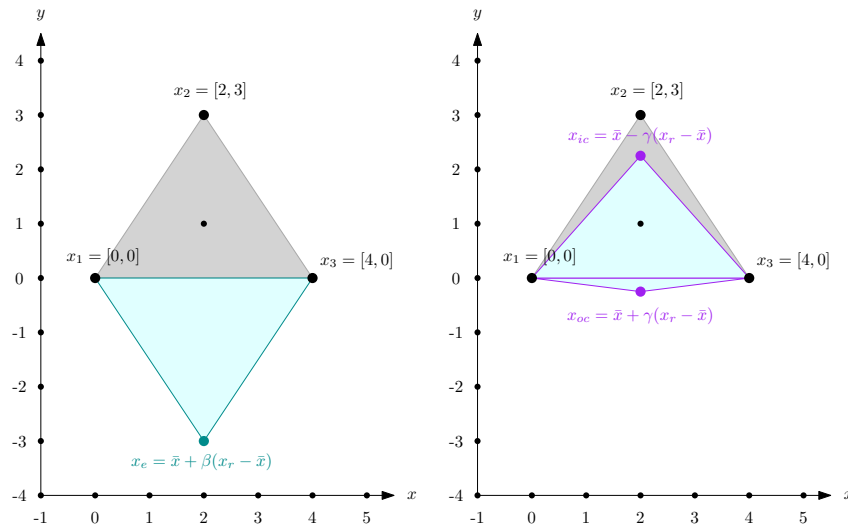
$$x_e = \bar{x} + \beta(x_r - \bar{x}) = [2, 1] + \beta([2, -1] - [2, 1]) = [2, -3].$$

3. **Kontrakce (Contraction):** Existují dva typy kontrakcí: **vnější** a **vnitřní** (obr. 48). Tento krok je opakem předchozího v tom smyslu, že namísto protahování se zrcadlený bod x_r přitahuje zpět směrem k těžišti \bar{x} . Pro vnější kontrakci tímto způsobem získáme bod x_{oc} pomocí vztahu

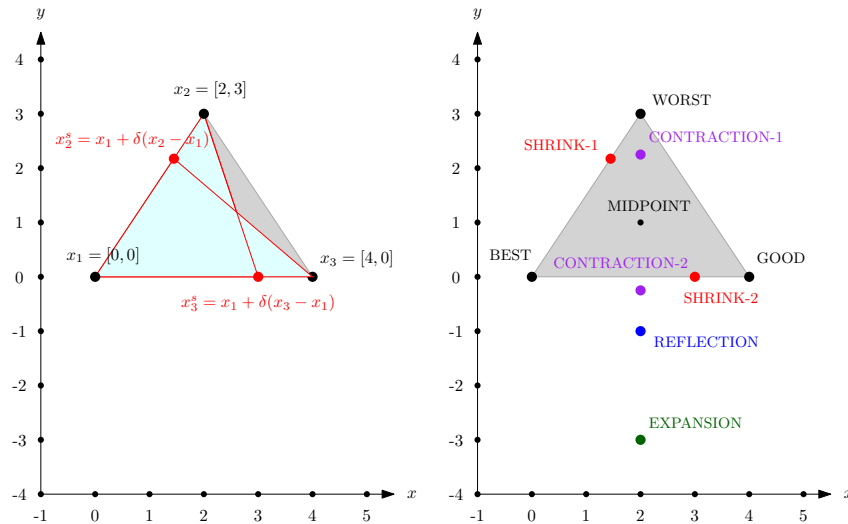
$$x_{oc} = \bar{x} + \gamma(x_r - \bar{x}).$$

Pro vnitřní kontrakci tímto způsobem získáme bod x_{ic} pomocí vztahu

$$x_{ic} = \bar{x} - \gamma(x_r - \bar{x}).$$



Obr. 48: Nelder–Mead – operace protažení (vlevo) a kontrakce (vpravo)

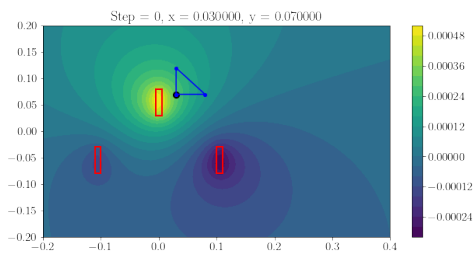


Obr. 49: Nelder–Mead – operace zmenšení (vlevo) a zobrazení základních bodů všech operací (vpravo)

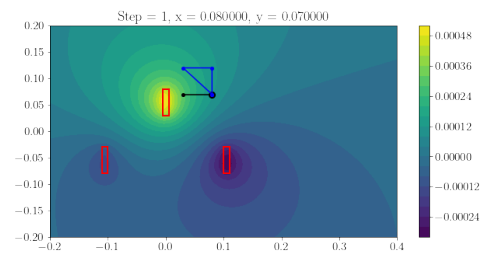
- Zmenšení (Shrink):** V posledním kroku zvolíme jeden referenční bod, ke kterému všechny ostatní body přitáhneme (obr. 49). Není vhodné v algoritmu volit stejný bod, pro který byly provedeny předchozí kroky, uvažujme tedy jako referenční bod x_1 . Potom všechny ostatní body zmenšíme pomocí vztahu

$$x_i^s = x_1 + \delta(x_i - x_1).$$

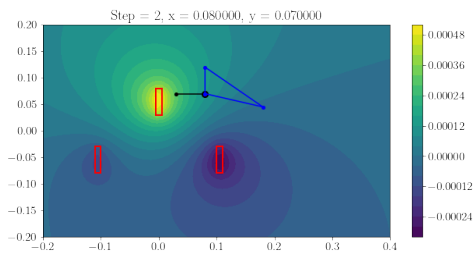
Na obr. 50 je zobrazeno řešení ilustrativního příkladu pomocí metody Nelder–Mead spolu s gradientem řešení zobrazených pomocí jednotlivých simplexů.



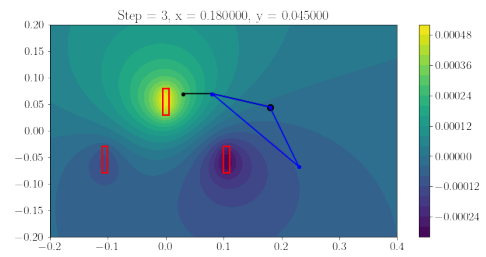
(a) Krok 1



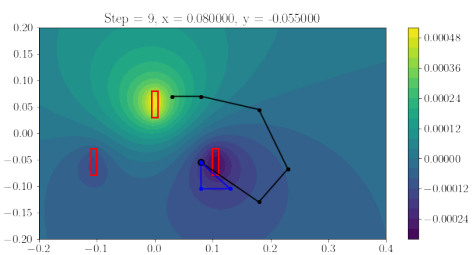
(b) Krok 2



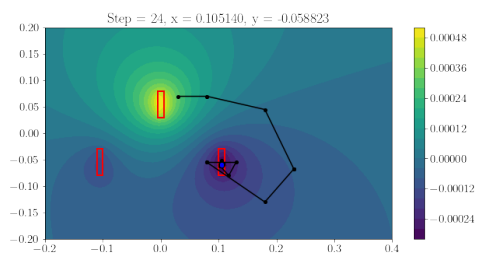
(c) Krok 3



(d) Krok 4



(e) Krok 10



(f) Krok 25

Obr. 50: Princip metody Nelder–Mead

```

1 # Vrcholy simplexu: WORST, GOOD, BEST + MIDPOINT
2 # Body zakladnich operaci: REFLECTION, EXPANSION, CONTRACTION, SHRINK
3
4 compute REFLECTION
5 if f(REFLECTION) < f(GOOD):
6     if f(BEST) < f(REFLECTION):
7         replace WORST -> REFLECTION
8     else:
9         compute EXPANSION
10
11     if f(EXPANSION) < f(BEST):
12         replace WORST -> EXPANSION
13     else:
14         replace WORST -> REFLECTION
15 else:
16     if f(REFLECTION) < f(WORST):
17         replace WORST -> REFLECTION
18
19 compute CONTRACTION
20 if f(CONTRACTION) < f(WORST):
21     replace WORST -> CONTRACTION
22 else:
23     compute SHRINK
24

```

```

25  replace WORST -> SHRINK
26  replace WORST -> MIDPOINT

```

Algoritmus 3.3: Základní algoritmus metody Nelder–Mead

Bound Optimization BY Quadratic Approximation - BOBYQA

je optimalizační algoritmus, který je založen na stejném principu jako algoritmus předchozí. Řeší problém metodou důvěryhodné oblasti (trust region), na níž se vytváří interpolační kvadratické modely. V každé iteraci se spočte jeden bod, obvykle vyřešením subproblémů důvěryhodné oblasti, který podléhá vázaným omezením, nebo alternativně výběrem bodu, který má nahradit interpolační bod za účelem podpoření lineární nezávislosti v interpolačních podmínkách [88].

3.3.2 *Stochastické metody*

Stochastické optimalizační metody obvykle potřebují více výpočtů během celého procesu optimalizace. Oproti metodám závislým na výpočtu gradientu se jedná o náhodný proces, nehrozí tedy možnost uvíznutí výpočtu řešení v lokálním minimu. Klasickým příkladem takového algoritmu je **Monte Carlo**, který je založen na náhodných volbách návrhových proměnných. Šanci na nalezení globálního optima lze velmi zlepšit velmi hustým statistickým vzorkováním [89].

Mezi další stochastické metody patří **evoluční algoritmy** a z nich vycházející **genetické algoritmy**. Základ každého evolučního algoritmu lze popsat tak, že každé možné řešení problému je reprezentováno jedincem v populaci. Populace se skládá z množiny jedinců, a to odpovídá množině různých řešení. Postupem času se populace vyvíjí. Špatní jedinci zanikají, objevují se lepší jedinci, kteří je v dalších populacích nahrazují. Pokud se tento postup bude opakovat dostatečně dlouho, populace bude tvořena pouze dobrými jedinci, kteří reprezentují správné řešení zadaného problému.

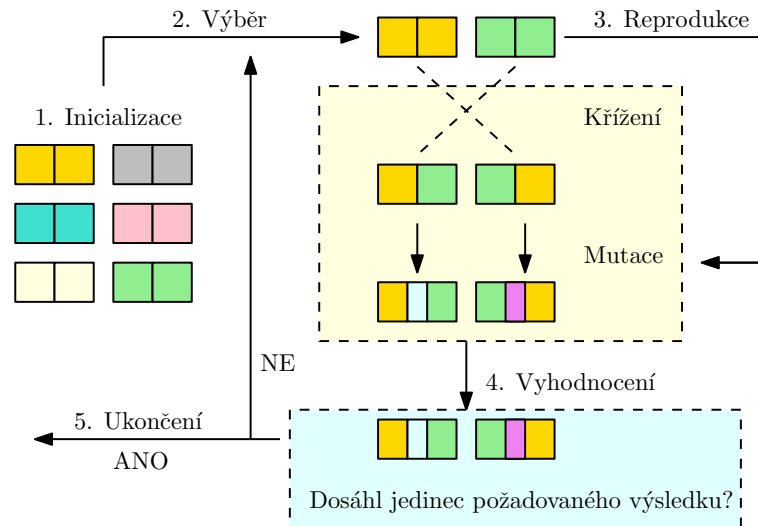
Aby byla evoluce funkční, je nezbytné provést následující kroky:

- Vytvořit ze dvou nebo více existujících řešení nové zprůměrované, této operaci se říká **křížení**.
- Vytvořené řešení umět náhodně pozměnit, této operaci se říká **mutace**.
- Je nezbytné pro vybraného jedince vybrat vhodného jedince ke křížení, této operaci se říká **selekce**.

Tato jednoduchá pravidla tvoří základ každého genetického algoritmu, viz obr. 51.

Proces genetické optimalizace pak bude fungovat následovně, podle např. [90] takto:

1. **Inicializace:** Vytvoření počáteční populace, všichni jedinci jsou (obvykle) náhodně vygenerováni tak, aby poskytovali celou škálu možných řešení ve vyhledávacím prostoru. Příležitostně mohou být řešení nasazena v oblastech, kde je pravděpodobné, že budou nalezena optimální řešení. Velikost populace závisí na povaze problému, ale obvykle obsahuje několik desítek až stovek potenciálních řešení kódovaných do jedinců.
2. **Výběr:** Pro každou po sobě následující generaci se poprvé vybírá určitý podíl stávající populace, aby se vytvořila nová generace. Jak již bylo zmíněno dříve, výběrový proces



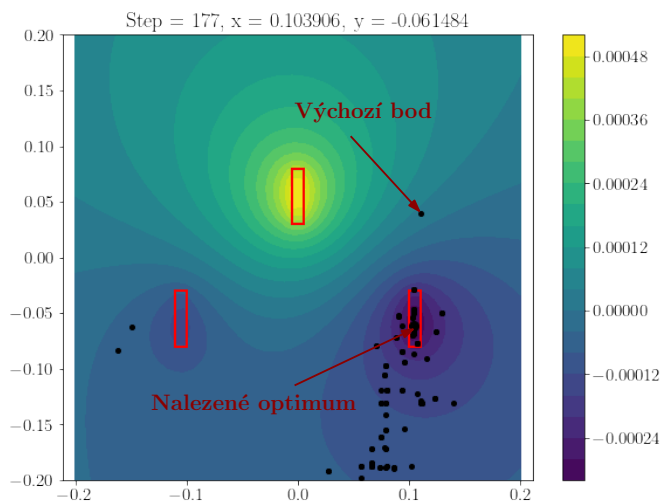
Obr. 51: Vývojový diagram genetického algoritmu

je náhodný nebo upřednostňuje jedince s vyšší vhodností. Některé metody výběru hodnotí vhodnost každého řešení a přednostně vybírají nejlepší řešení.

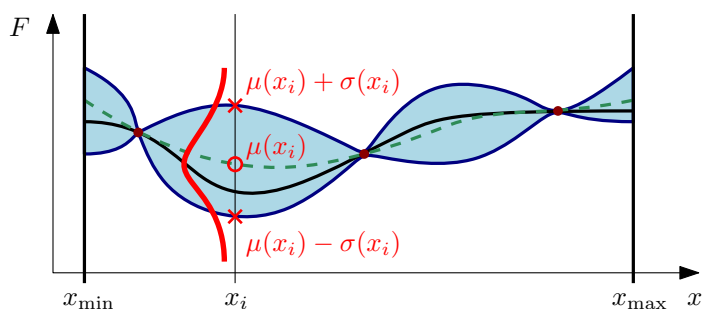
3. **Reprodukce:** Pro každou následnou generaci je druhým krokem generování jedinců příští generace z jedinců vybraných genetickými operátory, jako je křížení a mutace. Tyto genetické operátory nakonec vyústí v populaci jedinců další generace, která se liší od počáteční generace, ale kde jedinci obvykle sdílejí mnoho vlastností svých rodičů.
4. **Vyhodnocení:** Vlastnosti potomků jsou nejprve dekodovány a poté vyhodnoceny pomocí vhodné funkce a nahradí nejméně vhodné jedince v populaci tak, aby její velikost zůstala nezměněna.
5. **Ukončení:** V průběhu vyhodnocovacího kroku, pokud některý z jedinců dosáhne požadovaného výsledku nebo je dosaženo maximálního počtu iterací, je proces ukončen. V opačném případě se kroky 2 až 4 opakují za účelem vytvoření nové generace.

NSGA-II

Non-dominated Sorting Genetic Algorithm-II je druh genetického algoritmu. Je založen na principu Multi-Objective Evolutionary Algorithm (MOEA). NSGA-II využívá rychlé převzetí třídění a operátor, který vytváří pár pro množení tím, že kombinuje rodiče a populace potomků a vybere nejlepší sadu řešení (s ohledem na vhodnost a šíření). Algoritmus využívá náhodné rozložení výchozí populace, polynomiální mutaci a SBX (Simulated Binary Crossover) křížení. Je schopen najít velmi dobré řešení, které je blízko Pareto-optimální fronty. Nejvíce časově náročný úkol v multikriteriální optimalizaci je proces identifikace Pareto fronty ze sady řešení v multikriteriálním prostoru [89]. Řešení ilustrativního problému pomocí genetického algoritmu NSGA-II je znázorněno na obr. 52.



Obr. 52: Ukázka algoritmu NSGA-II



Obr. 53: Toleranční pásmo vymezené gaussovským procesem

3.3.3 Ostatní

Bayesovská optimalizace

Bayesovská optimalizace patří k technikám, které umožňují nalézt extrém reálné cílové funkce $f(x)$ i v případě, že informace o této funkci je neúplná a její vyčíslení je nákladné. Princip je velmi dobře patrný na obr. 53. Cílová funkce může být definována pomocí **gaussovského procesu** ve formě

$$f(\mathbf{x}) = \mathcal{GP}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')),$$

kde $m(\mathbf{x})$ je funkcí střední hodnoty a $k(\mathbf{x}, \mathbf{x}')$ je funkcí kovariance (jádro), kterou lze definovat mnoha způsoby, například jako exponenciální funkci (Squared Exponential Function)

$$k(\mathbf{x}, \mathbf{x}') = \exp\left(-\frac{1}{2}\|\mathbf{x} - \mathbf{x}'\|^2\right)$$

nebo konkrétněji

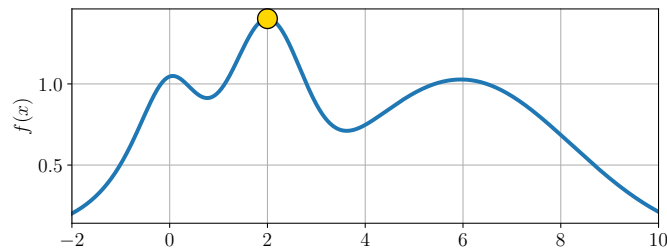
$$k(\mathbf{x}, \mathbf{x}') = \exp\left(-\frac{1}{2\theta^2}\|\mathbf{x} - \mathbf{x}'\|^2\right),$$

kde θ vyjadřuje hyperparametr, kterým lze naladit potřebné chování procesu.

Uvažujme cílovou funkci (obr. 54) na intervalu $x \in (-2, 10)$ ve tvaru

$$f(x) = e^{-(x-2)^2} + e^{-\frac{(x-6)^2}{10}} + \frac{1}{x^2 + 1},$$

která nabývá maximum 1,401 v bodě $x = 2$.



Obr. 54: Ilustrativní cílová funkce

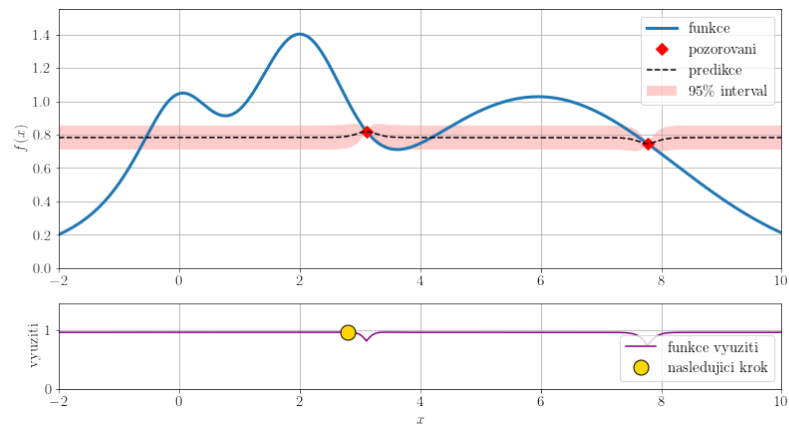
Zjednodušená verze této metody je nastíněna na obr. 53. Předpokládejme, že je nutno nalézt minimum cílové funkce, která není předem známá. Největší rozptyl vůči tomuto minimu je v bodě x_i . V tomto bodě se vypočítá hodnota cílové funkce (příklad na obr. 54) a následně se modifikuje tvar střední hodnoty. Stejným způsobem se poté postupuje v dalších bodech, dokud není nalezeno požadované optimum, viz obr. 55. Řešení ilustrativního příkladu pomocí metody je patrné na obr. 56.

Shrnutí

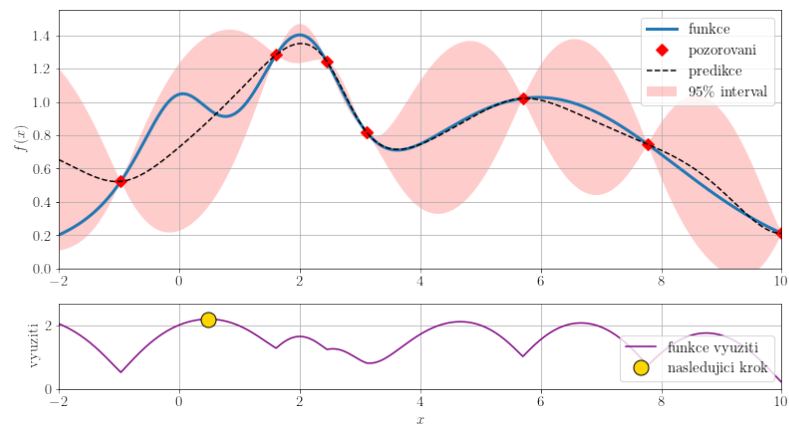
Různé metody vedou k různým výsledkům, jak je patrné v tabulce 4. Některé metody konvergují k lokálnímu minimu (většina gradientních metod), ale jejich vyhodnocení bývá rychlé. Pokud se gradientní metody dostanou do lokálního minima, bývá pro ně zpravidla obtížné dostat se ven. V případě, že lokální minimum metoda obejde, je velmi pravděpodobné, že dokonverguje k minimu globálnímu. Tento problém je možné odstranit pomocí opakování optimalizačního procesu a změnou výchozích hodnot pro první krok algoritmu. Oproti těmto metodám jsou stochastické algoritmy velmi pomalé a výsledky optimalizace se mohou značně lišit pro jednotlivé běhy, ovšem nehrozí jejich uvíznutí v lokálních minimech. Je velmi složité zvolit vhodnou optimální metodu pro konkrétní problém.

Tab. 4: Dosažené minimum magnetického vektorového potenciálu pomocí různých algoritmů

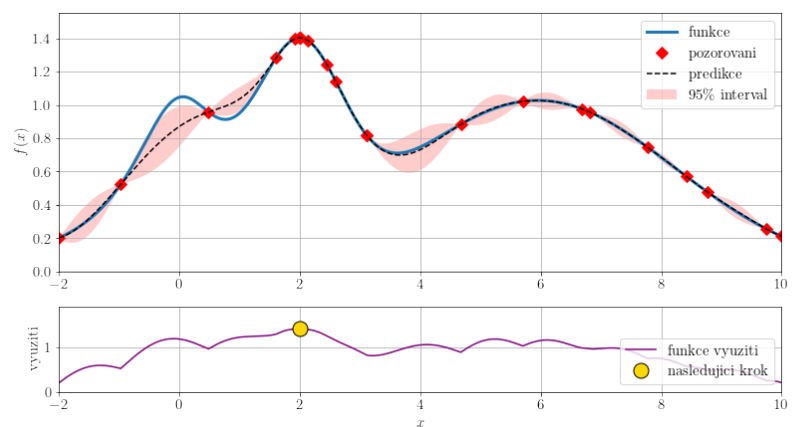
Metoda	Typ	x (m)	y (m)	A (Wb · m ⁻¹)	počet kroků
Gradientní sestup (příliš velký krok)	lokální	0,111637	-0,062122	-0,000295	50
Gradientní sestup (malý krok)	lokální	0,105385	-0,059476	-0,000321	200
Gradientní sestup (adaptivní krok)	lokální	0,105357	-0,059468	-0,000321	20
Nelder-Mead (NLOpt)	lokální	0,104883	-0,060826	-0,000320	42
BOBYQA (NLOpt)	lokální	0,105447	-0,061431	-0,000320	25
Bayesovská optimalizace	globální	0,105334	-0,059555	-0,000321	40
NSGA-II	globální	0,105577	-0,064151	-0,000318	210



(a) Krok 2

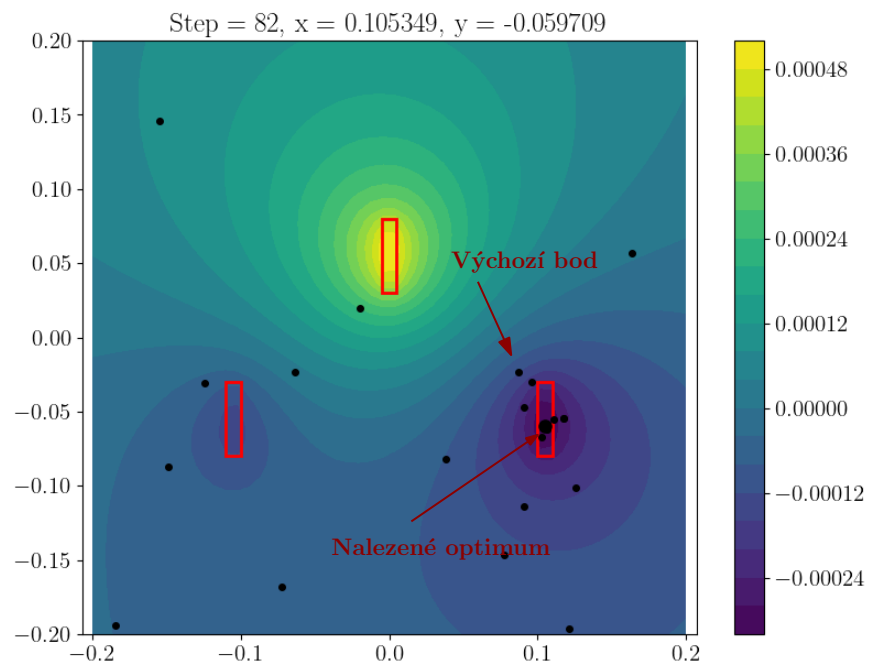


(b) Krok 8



(c) Krok 23

Obr. 55: Bayesovská optimalizace



Obr. 56: Ukázka Bayesovské optimalizace

Část III

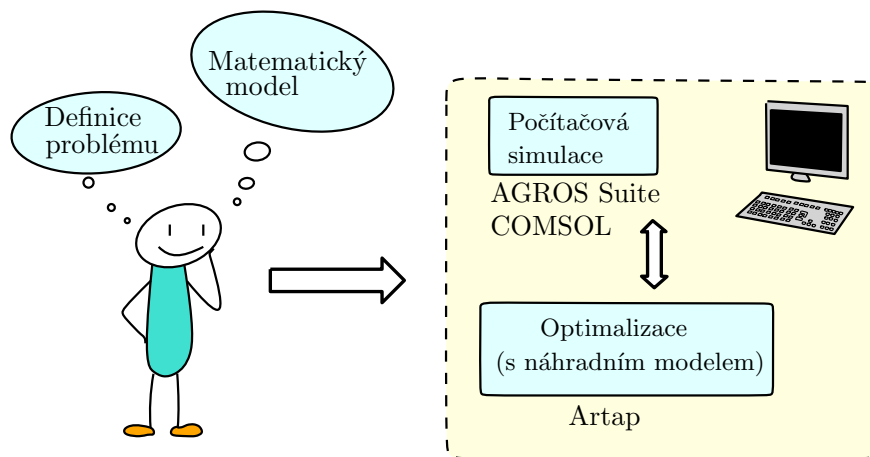
SOFTWAREVÁ IMPLEMENTACE

SOFTWAREVÁ IMPLEMENTACE

Metody popsané v předchozí kapitole byly implementovány v softwarových programech, které jsou vyvíjeny na katedře elektrotechniky a počítačového modelování. Tyto programy využívají vlastních implementací popsaných algoritmů s využitím volně dostupných knihoven třetích stran nebo komerčních výpočetních programů.

Na obr. 57 je zobrazen princip využití těchto výpočetních programů. V první fázi je potřebné definovat samotný problém a popsat jeho podstatu odpovídajícím matematickým modelem. V dalším kroku je na základě matematického modelu sestavena odpovídající počítačová simulace v libovolném softwaru. Na katedře elektrotechniky a počítačového modelování FEL ZČU v Plzni je pro řešení multifyzikálních problémů dlouhodobě vyvíjen vlastní software - Agros Suite. Z komerčních softwarů je dále využíván program COMSOL Multiphysics.

V dalším kroku je možné parametry takto definovaného modelu optimalizovat. Pokud je dané řešení provedené pomocí dostupných softwarů příliš složité, pak je možné provést optimalizaci pro parametry vhodně zvoleného náhradního modelu. K tomu lze využít open-source framework Artap, který je vyvíjený rovněž na katedře elektrotechniky a počítačového modelování a poskytuje možnosti využití mnoha optimalizačních technik a nástrojů a také algoritmů pro strojové učení.



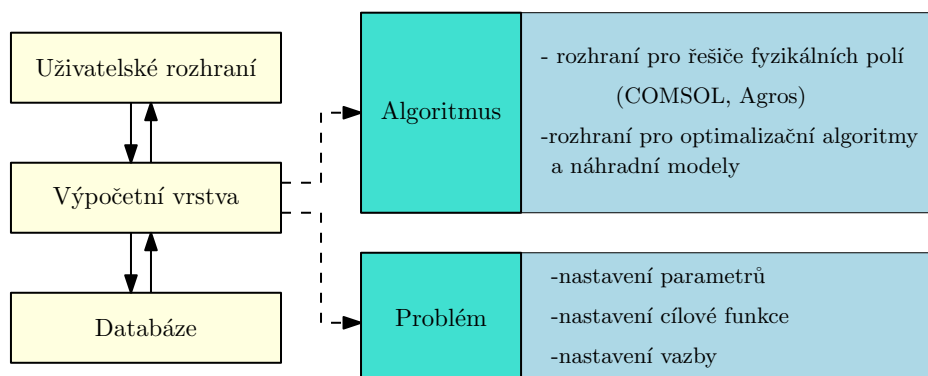
Obr. 57: Využívané softwary pro tvorbu počítačových simulací a řešení optimalizační problémy

ARTAP

Projekt Artap je open-source framework s rozsáhlou možností využití nástrojů optimalizace napsaný v jazyce Python. Poskytuje možnosti využití poměrně širokého množství optimalizačních algoritmů (BOBYQA, NSGA-II, Bayes Optimization, atd.), integrovaných i externích řešičů parciálních diferenciálních rovnic nebo nástrojů pro strojové učení.

K definici problému řešeného pomocí Artap serveru potřebujeme:

- řešiče fyzikálních polí: Agros Suite, COMSOL Multiphysics, FEMM, ANSYS, CST,
- optimalizační algoritmy: vlastní, SciPy, BayesOpt, NLOpt,
- distribuované výpočty: HTCCondor,
- zjednodušené (surrogate) modely: TensorFlow, Keras, Scikit, SMT.



Obr. 58: Třívrstvá struktura Artapu a hlavní třídy aplikační vrstvy

Artap má třívrstvou strukturu, viz obr. 58. Hlavní částí je aplikační vrstva, které je rozdělena do dvou hlavních tříd pro specifikaci:

- třída **Algoritmus** - specifikace postupu při řešení optimalizace, volba náhradního modelu, případně další paralelní aplikace nebo odkazy na externí řešiče,
- třída **Problém** - nastavení omezení a limitů samotného optimalizačního procesu, definice cílové funkce, volba parametrů.

KNIHOVNY TŘETÍCH STRAN

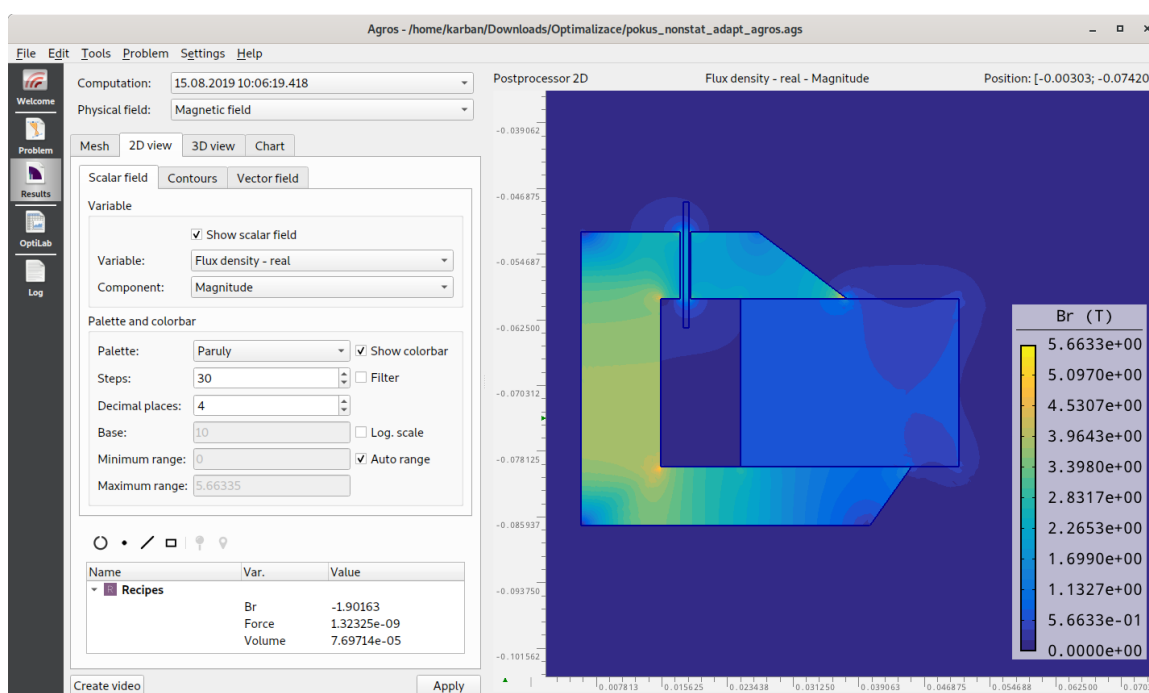
Artap je obecný robustní framework pro širokou škálu optimalizačních metod s jednotným rozhraním, který vhodně využívá další knihovny třetích stran. Nejdůležitější z nich jsou:

- **SciPy** je ekosystém otevřeného softwaru pro matematiku, vědu a výzkum založený na Pythonu. Zejména se jedná následující o základní balíčky:

- NumPy – základní balíček pro numerické výpočty, definuje typy číselných polí a matic a základní operace s nimi.
- SciPy – sada numerických algoritmů a nástrojů pro zpracování signálu, optimalizaci, statistiky, atd.
- Pandas – sada datových struktur pro analýzu dat
- Matplotlib – balíček pro tvorbu základních 2D a 3D. vykreslovacích struktur.
- **SMT: Surrogate Modeling Toolbox** je open source sada nástrojů pro tvorbu náhradních modelů a vzorkování prostoru v jazyce Python.
- **TensorFlow** je další z platforem pro strojové učení. Obsahuje komplexní a flexibilní ekosystém nástrojů, knihoven a komunitních zdrojů, který umožňuje práci s nejmodernějšími technologiemi strojového učení.
- Knihovna **scikit-learn** je toolbox pro strojové učení zahrnující polynomickou aproximaci prvního a druhého řádu, dále umělé neuronové sítě, nabízí možnost využití i vícerozměrných adaptivních regresních splinů a další množství nástrojů pro regresi a klasifikaci.
- **NLOpt** je otevřená knihovna pro nelineární optimalizaci, která poskytuje společné rozhraní pro řadu různých bezplatných optimalizačních rutin dostupných člancích i původní implementace dalších algoritmů. Jedná se zejména o následující algoritmy:
 - Nelder-Mead, BOBYQA,
 - DIRECT a DIRECT-L,
 - Control Random Search (CSR),
 - Multi-Level-Single-Linkage (MLSL).
- **BayesOpt: A Bayesian optimization library** je efektivní implementace Bayesovské optimalizační metody. Místo přímé optimalizace cílové funkce vytváří pravděpodobnostní model, který poté využije na základě jádra pro sestavení nové lehce optimalizovatelné funkce.

AGROS SUITE

Agros Suite je software pro řešení multifyzikálních problémů založený na metodě konečných prvků. Jedná se o nástupce softwaru Agros2D, který byl vyvíjen na katedře elektrotechniky a počítačového modelování od roku 2009. Původní verze využívá implementaci knihovny Hermes [94], které jsou určeny pro řešení parciálních diferenciálních rovnic. Jejich numerický výpočet je založen na metodě konečných prvků vyššího řádu, která je plně adaptivní. Novější druhá varianta využívá knihovnu deal.II [92]. Knihovna umožňuje tvorbu vlastní implementace konečných prvků. Obě knihovny jsou distribuovány pod licencí GNU General Public License v2.1. Ukázka hlavního okna programu Agros je znázorněna na obr. 59.



Obr. 59: Hlavní okno – postprocessor

Hlavním cílem projektu Agros je vytvořit univerzální nástroj pro numerické řešení širokého spektra parciálních diferenciálních rovnic druhého řádu popisujících různá fyzikální pole. Kromě primárního elektromagnetického pole lze řešit rozložení teplotního pole, mechaniku, nestlačitelné proudění, akustické pole, apod., viz tabulka 5. V průběhu více než 10 let svého vývoje se jeho vnitřní struktura podstatně změnila.

Fyzikální pole	Ustálený stav	Harmonický ust. stav	Přechodný děj
Elektrické pole	x		
El. proudové pole	x	x	
Magnetické pole	x	x	x
Vysokofrekvenční elmag. pole		x	x
Teplotní pole	x		x
Strukturální mechanika	x		
Nestlačitelné proudění	x		x
Akustické pole		x	x
Obecná parc. dif. rovnice	x	x	x

Tab. 5: Podporovaná fyzikální pole

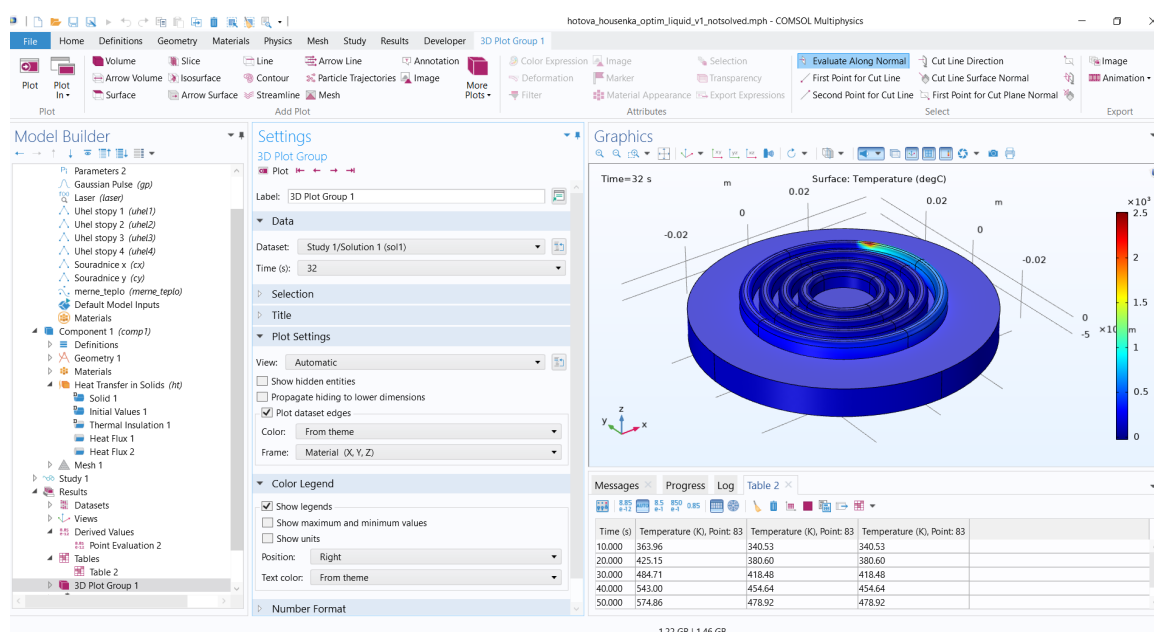
Klíčové vlastnosti Agros Suite

Agros Suite s Hermes a deal.II backendy obsahuje řadu výkonných funkcí, které ještě nejsou implementovány do dostupných komerčních kódů. Nejdůležitější z nich jsou:

- ***h*, *p* a *hp* adaptivita** – adaptivní techniky se s výhodou používají, pokud místní chyba řešení na jakémkoli místě překročí předepsanou hodnotu. Přesné řešení (ve většině případů) ale obecně není známo. Proto je nahrazeno referenčním řešením získaným na dostatečně husté síti. Toto řešení pak poskytuje informaci o místech, kde by měla být aplikována adaptivita [93, 94].
- **Multimesh** – v případě řešení spojených problémů několika fyzikálních polí lze každé pole vypočítat na nezávislé síti, která se může značně lišit od ostatních sítí používaných pro zbývající pole. Každou síť lze tedy přizpůsobit vlastnostem zpracovávaného pole.
- **Visící uzly** – Hermes obsahuje sofistikované postupy vyššího řádu pro podporu visících uzlů libovolné úrovně. deal.II umožňuje použít volné uzly prvního řádu. Jejich zpracování obvykle vyžaduje značný nárůst stupňů volnosti. Implementované algoritmy umožňují upřesnění pouze nejbližšího prvku kolem uzlu, zatímco ostatní části nevyžadují žádné další zjemnění.
- **Různé tvary prvků** – Hermes je schopen kombinovat trojúhelníkové a čtyřúhelníkové prvky. Libovolná strana může být popsána splinem. To představuje účinný nástroj pro vytváření sítí charakterizovaných zakřivenými rozhraními nebo hranicemi. Knihovna deal.II umožňuje používat pouze čtyřúhelníkové sítě s křivočarými prvky.
- **Časová adaptivita** – při řešení různých nestacionárních problémů je jedním z důležitých úkolů co nejvíce snížit počet časových kroků, aby se zkrátil čas výpočtu. Používání metod založených na implicitním nebo explicitním Eulerově algoritmu není efektivní. Strategie implementovaná v knihovně Hermes a deal.II je založena na diskretizaci parciálních diferenciálních rovnic v prostoru (Rotheho metoda), která poskytuje systém časově závislých obyčejných diferenciálních rovnic, který je následně vyřešen pomocí metody BDF2.

COMSOL MULTIPHYSICS

COMSOL Multiphysics je určen k řešení inženýrských úloh pomocí metody konečných prvků. Software lze s úspěchem používat v situacích, kdy je nutné do modelu zahrnout více fyzikálních polí – jedná se pak o komplexní, tzv. multifyzikální úlohy. Součástí programu COMSOL Multiphysics a jeho rozšiřujících nadstaveb jsou předdefinované typové úlohy, které ulehčují uživateli práci při definování modelů. Jako příklad těchto úloh lze uvést elektromagnetismus, proudění tekutin, přestup tepla, zatížení konstrukcí. Program obsahuje potřebné funkce k vytvoření a analýze modelu – počínaje od definování geometrie, zadání okrajových podmínek, vytvoření sítě, nastavení řešičů až po vizualizaci výsledků. Nastavení modelu je velice rychlé, díky množství přednastavených fyzikálních rozhraní (mechanika tekutin, přestup tepla, elektrostatika, mechanika pevných těles, atd.). Je možné definovat geometrii, materiálové vlastnosti a okrajové podmínky pomocí funkcí nebo závisle proměnných - tyto definice mohou být proměnné v čase nebo prostoru. Hlavní okno programu je na obr. 60.



Obr. 60: Hlavní okno – postprocessor

COMSOL Multiphysics je otevřeným rozhraním pro uživatele, kteří si chtějí vytvořit vlastní simulaci definováním matematických rovnic. Matematické rovnice mohou být ve formě PDR nebo ODR. Definice vlastních rovnic probíhá přímo v grafickém rozhraní programu a není nutné vytvářet jakékoliv přídavné kódy nebo skriptové soubory.

Pro složitější úlohy, jako je třeba tvorba uživatelských aplikací, pokročilých optimalizačních pro-

blémy, pokročilé metody zpracování dat, vizualizace, apod., je možné propojit program COMSOL se simulačním nástrojem Matlab.

COMSOL Multiphysics je dostupný na všech obvyklých operačních systémech a podporuje výpočty na vícejádrových počítačích i na clusterech.

Část IV

TESTOVACÍ PŘÍKLADY S ANALYTICKÝM ŘEŠENÍM

ÚVOD

Optimální návrh elektromagnetických struktur pomocí globálních optimalizačních technik je oblastí výzkumu s širokým spektrem aplikací. V současnosti se mnoho významných mezinárodních konferencí věnuje právě tématům problematiky optimalizace návrhu těchto zařízení (COMPUMAG, CEFC, ISEF, a další). COMPUMAG Team (Testing Electromagnetic Analysis Method) prezentuje benchmarkové příklady pro návrh elektromagnetických zařízení založené na metodě konečných prvků s využitím optimalizačních vazeb pro zajištění optimálního návrhu. Díky nim je poukázáno na optimalizační metody, které jsou pro tento druh optimalizace nejvhodnější.

Oproti modelům založeným na analýze pomocí metody konečných prvků, analytické modely jsou velmi často využívány v prvních fázích počítačového modelování, kdy je potřeba ověřit řešení získané pomocí numerických metod. V následující kapitole jsou uvedeny tři testovací optimalizační problémy formulované pro analytické modely. Analytické modely jsou ověřeny simulací v programu COMSOL Multiphysics.

Řešení numerických modelů pomocí metody konečných prvků jsou v současnosti jedním z nejvyužívanějších přístupů pro návrhy nejen elektromagnetických zařízení. V prvních fázích je ovšem nutné správnost předpokladů mnohdy ověřit na zjednodušeném řešení, které poskytují modely analytické. Ty jsou velmi důležitým nástrojem hned v počátcích návrhu, kdy je potřeba učinit rozhodnutí o struktuře zařízení a také o použitých materiálech. Pro tuto fázi je jejich použití mnohem výhodnější než konečněprvkový přístup – rychlejší výpočet i pro více parametrů a více fyzikálních jevů. Na druhou stranu vzhledem k uvažovaným zjednodušením bývají méně přesné a k jejich aplikaci je potřebný silný teoretický základ.

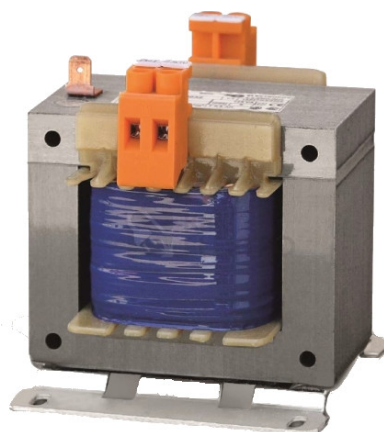
V této části práce jsou testovány vybrané optimalizační techniky a techniky náhradních modelů, které byly popsány v části (ii), na třech příkladech se známým analytickým řešením. Vzhledem k rychlosti řešení problémů popsanych pomocí analytických vztahů je možné otestovat optimalizační techniky s využitím náhradních modelů, které byly podrobněji popsány v kapitole ii, než budou implementovány na složitější počítačové simulace.

První testovací příklad je zaměřen na návrh oddělovacího transformátoru a druhý příklad popisuje návrh stejnosměrného kolového motoru použitého pro pohánění solárního vozidla. Samotná optimalizace u těchto dvou příkladů je vždy rozdělena na dvě části. Nejprve se ověří řešení pro jednokriteriální optimalizaci s požadavkem co nejmenší hmotnosti zařízení a ve druhém kroku je realizována vícekriteriální optimalizace s ohledem na co nejmenší hmotnost a zároveň maximální účinnost. Tyto dva příklady jsou benchmarkové příklady se známým analytickým řešením prezentované v [95] a v [96] a byly publikovány na konferencích. Třetí příklad ukazuje použití rekurentních neuronových sítí na predikci řešení obyčejné diferenciální rovnice popisující sériový RL obvod napájený ze zdroje napětí.

ANALYTICKÝ MODEL ODDĚLOVACÍHO TRANSFORMÁTORU

V této kapitole je představen první testovací příklad zaměřený na optimalizaci návrhu bezpečnostního oddělovacího transformátoru. Jedná se o příklad typu benchmark, který byl publikován na konferenci COMPUMAG v roce 2007, viz [95]. Příspěvek byl testován jako jednokriteriální i vícekriteriální optimalizační problém, který byl opět založen na řešení analytického modelu. Optimalizace byla provedena s využitím algoritmů ISRES, Nelder–Mead, Pattern Search a NSGA-II.

Na obr. 61 je pro ilustraci zobrazen jednofázový transformátor od českého výrobce ELEKTROKOV ZNOJMO se standardními plechy tvaru EI se svařeným jádrem, s převodem napětí 230 V/24 V, frekvencí 50/60 Hz, krytím IP00, svorkami IP20 a třídou izolace B, který principiálně zcela odpovídá řešenému příkladu.

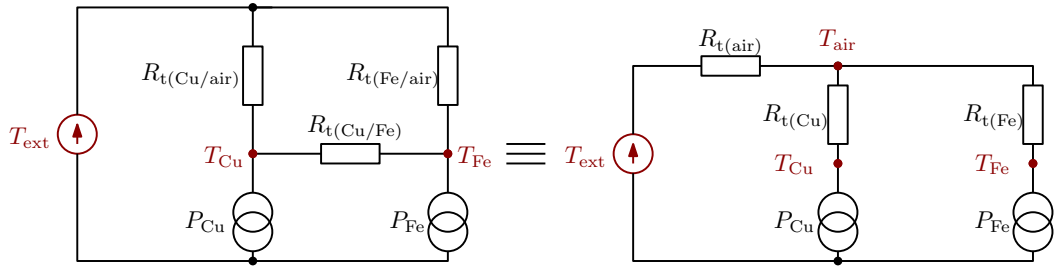


Obr. 61: Transformátor JBC 230 V/24 V 100 VA [97]

1.1 ANALYTICKÝ MODEL

Základní fyzikální procesy uvnitř transformátoru lze popsat pomocí vztahů pro elektromagnetické pole doplněné o teplotní pole. Rovnice jsou podrobně popsány v [95]. V této disertační práci je uvedeno pouze několik základních vztahů bez detailního odvození. Při tvorbě analytického modelu je uvažováno několik zjednodušujících předpokladů, jako rovnoměrné rozložení magnetické indukce v železném jádře či zanedbání vlivu výměny tepla mezi sekundární cívkou a magnetickým obvodem a s výjimkou centrálního jádra není uvažována výměna tepla mezi primární cívkou a magnetickým obvodem. Pro všechny povrchy je uvažován stejný koeficient konvektivního přestupu tepla a teploty v měděných a železných částech jsou stejné.

Zjednodušené schéma pro odvození tepelného modelu je patrné z obr. 62.



Obr. 62: Schéma tepelného modelu

Na obr. 62 je uvedeno zjednodušené schéma tepelného modelu transformátoru, ze kterého lze odvodit teplotu měděných cívek

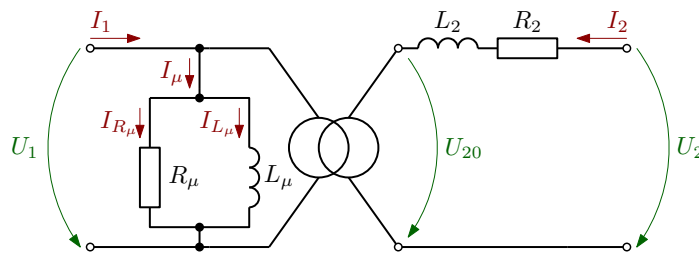
$$T_{\text{Cu}} = T_{\text{ext}} + \frac{R_{t(\text{Fe/air})} \cdot P_{\text{Cu}} + R_{t(\text{Fe/air})} \cdot P_{\text{Fe}} + R_{t(\text{Cu/Fe})} \cdot P_{\text{Cu}}}{R_{t(\text{Fe/air})} + R_{t(\text{Cu/air})} + R_{t(\text{Cu/Fe})}} \cdot R_{t(\text{Cu/air})} \quad (13)$$

a teplotu železa magnetického obvodu

$$T_{\text{Fe}} = T_{\text{ext}} + \frac{R_{t(\text{Cu/air})} \cdot P_{\text{Cu}} + R_{t(\text{Cu/air})} \cdot P_{\text{Fe}} + R_{t(\text{Cu/Fe})} \cdot P_{\text{Fe}}}{R_{t(\text{Fe/air})} + R_{t(\text{Cu/air})} + R_{t(\text{Cu/Fe})}} \cdot R_{t(\text{Fe/air})}, \quad (14)$$

kde T_{Cu} , T_{Fe} , T_{ext} jsou teploty mědi, železa a vnějšího prostředí, $R_{t(\text{Cu/air})}$ je tepelný odpor mezi mědí a vzduchem, $R_{t(\text{Fe/air})}$ je tepelný odpor mezi železem a vzduchem, $R_{t(\text{Cu/Fe})}$ je tepelný odpor mezi mědí a železem a konečně P_{Cu} a P_{Fe} jsou příslušné Jouleovy ztráty a ztráty v železe.

Na obr. 63 je znázorněno zjednodušené obvodové schéma. Předpokladem spojeným s tímto modelem je rovnoměrné rozložení indukce v magnetickém obvodu a absence poklesu napětí v důsledku magnetizačního proudu, parametry I_1 , I_2 , I_μ značí primární, sekundární a magnetizační proudy, U_1 , U_2 , U_{20} jsou napětí na primární a sekundární straně a při stavu naprázdno, R_μ , R_2 představují magnetizační a sekundární rezistanci a L_μ , L_2 značí magnetizační a sekundární indukčnost.



Obr. 63: Zjednodušené obvodové schéma řešeného elektromagnetického obvodu

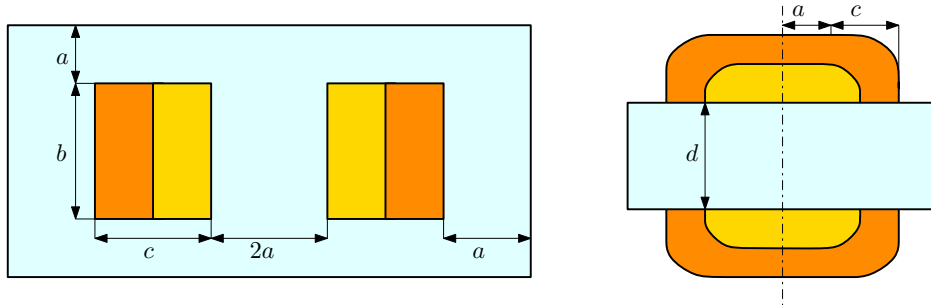
Maximální hodnotu magnetické indukce lze získat ze vztahu

$$B_m = \frac{1}{4\pi} U_1 \frac{\sqrt{2}}{n_1 a d f}, \quad (15)$$

kde n_1 je počet závitů na primárním vinutí, a a d jsou rozměry magnetického obvodu, viz obr. 64, a f je kmitočet.

1.2 ILUSTRATIVNÍ PŘÍKLAD

Uvažujme nyní transformátor s rozměry magnetického obvodu a , b , c a s tloušťkou d , viz obr. 64. Průřezy vodičů primárního a sekundárního vinutí jsou vyjádřeny parametry S_1 , S_2 . Počet závitů primární cívky je n_1 .



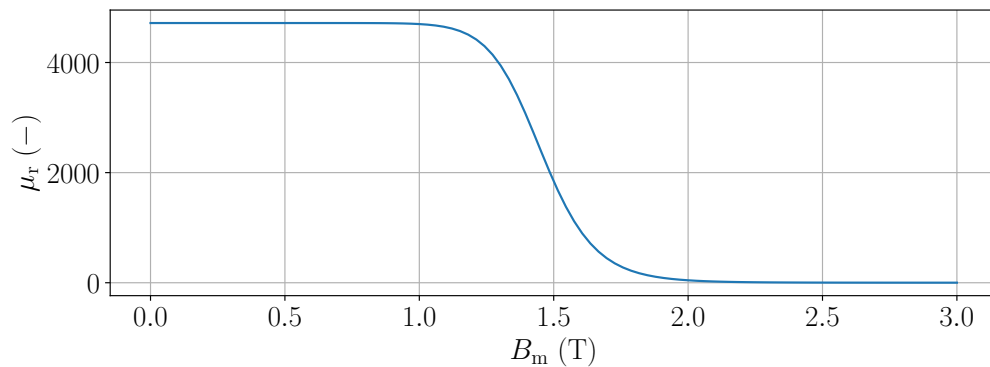
Obr. 64: Rozměry řešeného transformátoru

Dále uvažujme hodnoty napětí primárního vinutí $U_1 = 230 \text{ V}$ a sekundárního $U_2 = 24 \text{ V}$, frekvenci $f = 50 \text{ Hz}$, činitel plnění sekundární cívky $f_2 = 0,8$ a proud sekundárním vinutím $I_2 = 8 \text{ A}$. Počáteční teplota je $T_{\text{ext}} = 40^\circ\text{C}$, koeficient přestupu tepla konvekcí $h = 10 \text{ W} \cdot \text{m}^{-2}\text{K}^{-1}$, tepelná vodivost izolace $\lambda = 0,15 \text{ W} \cdot \text{m}^{-1}\text{K}^{-1}$, tloušťka izolace je $l = 1 \text{ mm}$. Měrná hmotnost železa je $\rho_{\text{Fe}} = 7800 \text{ kg} \cdot \text{m}^{-3}$ s elektrickou vodivostí $\sigma_{\text{Fe}} = 9,93 \cdot 10^6 \text{ S} \cdot \text{m}^{-1}$, měrná hmotnost mědi je $\rho_{\text{Cu}} = 8800 \text{ kg} \cdot \text{m}^{-3}$ s elektrickou vodivostí $\sigma_{\text{Cu}} = 56 \cdot 10^6 \text{ S} \cdot \text{m}^{-1}$.

Relativní permeabilitu transformátorových plechů lze vyjádřit pomocí následujícího vztahu

$$\mu(B) = \frac{1}{\chi + \frac{(1-\chi)B^{2\alpha}}{B^{2\alpha} + \beta}}, \quad (16)$$

kde $\chi = 2,12 \cdot 10^{-4}$, $\alpha = 7,358$ a $\beta = 1,18 \cdot 10^6$ jsou konstanty uvažovaného materiálu. Magnetizační charakteristika je zobrazena na obr. 65.



Obr. 65: Magnetizační charakteristika transformátorových plechů

1.3 OPTIMALIZACE

Řešený problém obsahuje celkem 7 návrhových proměnných – rozměry magnetického obvodu a , b , c a jeho tloušťku d , viz obr. 64, průřezy vodičů primárního a sekundárního vinutí S_1 , S_2 a počet závitů primární cívky je n_1 . Tyto parametry budou předmětem optimalizace a jejich spodní a horní meze a počáteční hodnoty jsou uvedeny v tabulce 6.

Tab. 6: Vstupní parametry pro optimalizaci

Parametr	a (mm)	b (mm)	c (mm)	d (mm)	n_1 (–)	S_1 (mm ²)	S_2 (mm ²)
výchozí hodnota	10	20	30	60	800	5	5
spodní mez	3	14	6	10	200	0,15	0,15
horní mez	30	95	40	80	1200	19	19

1.3.1 Jednokriteriální optimalizace

Cílem jednokriteriální optimalizace je minimalizovat celkovou hmotnost m_{total} železných a měděných materiálů. Optimalizační problém je popsán jako sdružená úloha magnetického a teplotního pole. Na tento problém je použito 7 omezujících vazeb – pro teplotu mědi a železa, hodnotu účinnosti, hodnotu magnetizačního proudu, úbytku napětí a činitele plnění obou cívek. Omezme tedy obecnou cílovou funkci m_{total} :

$$\begin{array}{ll} \min m_{\text{total}}, & \text{za podmínek} \\ & T_{\text{Cu}} < 120^\circ \text{ C} \quad T_{\text{Fe}} < 100^\circ \text{ C} \\ & \frac{I_{10}}{I_1} \leq 0,1 \quad \frac{\Delta U_2}{U_2} \leq 0,1 \\ & f_1, f_2 \leq 1 \quad \eta \geq 80\% \end{array}$$

Tab. 7: Hodnoty limitujících požadavků pro optimum úlohy (benchmark)

	m_{total} (kg)	T_{Cu} (°C)	T_{Fe} (°C)	I_{10}/I_1 (–)	$\Delta U_2/U_2$ (–)	η (–)	f_1 (–)	f_2 (–)
Hodnota	2,31	108,8	100,0	0,07	0,1	0,90	1,0	1,0

Pro optimum benchmarkové úlohy jsou hodnoty pro omezení uvedeny v tabulce 7. Optimalizace byla provedena pomocí algoritmů ISRES, Pattern Search, NSGA-II a Nelder–Mead, výsledky jsou patrné i s optimálním řešením v tabulce 8. Průběhy jednotlivých optimalizačních algoritmů jsou zobrazeny na obr. 66. Jak je z grafů a z hodnot v tabulce patrné, nejhoršího výsledku dosáhla simplexová metoda Nelder–Mead. Vzhledem k tomu, že se jedná o multimodální úlohu a také vzhledem ke množství vstupních parametrů optimalizace, je patrné, že algoritmus nejspíše dosáhl jednoho z lokálních minim. Nejlepší shodu v tomto případě vykazuje algoritmus NSGA-II i s ohledem na přijatelnou dobu řešení.

Tab. 8: Optimální hodnoty vstupních parametrů pro jednotlivé algoritmy

	m_{total} (kg)	a (mm)	b (mm)	c (mm)	d (mm)	n_1 (-)	S_1 (mm ²)	S_2 (mm ²)	čas (s)
Optimum	2,31	12,92	50,12	16,61	43,26	640,8	0,33	2,92	-
ISRES	2,32	13,15	60,96	15,55	37,50	724,2	0,33	2,93	42,6
Pattern Search	3,09	19,77	44,46	9,87	47,85	409,0	0,15	2,40	8,3
NSGA-II	2,32	12,21	29,03	20,71	63,83	465,9	0,32	2,67	27,0
Nelder–Mead	5,12	15,20	94,23	34,37	33,68	746,1	0,43	9,90	163,4

1.3.2 Vícekriteriální optimalizace

V případě jednokriteriální optimalizace byla hlavním požadavkem co nejmenší celková hmotnost železných plechů a měděného vinutí m_{total} . Pro vícekriteriální optimalizaci bylo použito šest limitujících omezení—pro teplotu mědi a železa, hodnotu magnetizačního proudu, úbytku napětí a činitele plnění obou cívek. Účinnost η , která byla v předchozí kapitole použita jako jedna z vazeb, je nyní kromě celkové hmotnosti druhým optimalizačním kritériem. V tomto případě neomezujeme tedy možnosti řešení pro $\eta \geq 80\%$, ale snažíme se její hodnotu maximalizovat. Protože jsou obecně optimalizační problémy řešeny jako minimalizační, využijeme k popisu cílové funkce popisující účinnost její zápornou hodnotu.

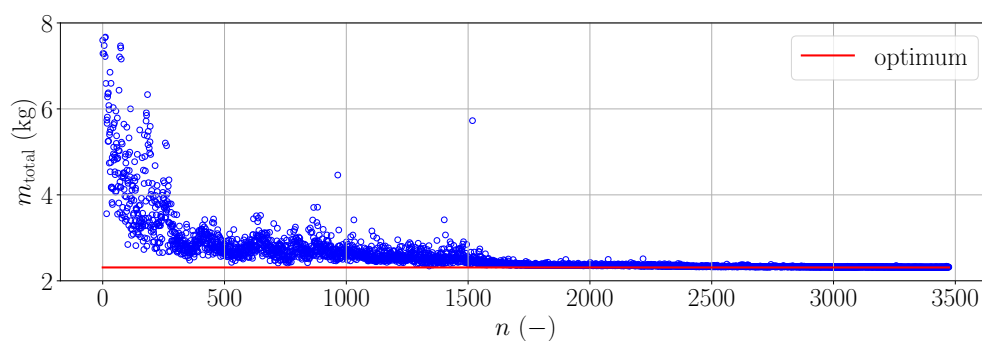
$$\begin{array}{ll}
 \min m_{\text{total}}, & \text{za podmínek} \\
 \min (-\eta) &
 \end{array}
 \quad
 \begin{array}{ll}
 T_{\text{Cu}} \leq 120 \text{ }^\circ\text{C} & T_{\text{Fe}} < 100 \text{ }^\circ\text{C} \\
 \frac{I_{10}}{I_1} \leq 0,1 & \frac{\Delta U_2}{U_2} \leq 0,1 \\
 f_1, f_2 \leq 1 &
 \end{array}$$

Vícekriteriální optimalizace byla řešena pomocí genetického algoritmu NSGA-II a AGE-MOEA. Výsledky na Paretově frontě porovnané s frontou z daného řešení testovací úlohy, je zobrazeno na obr. 67.

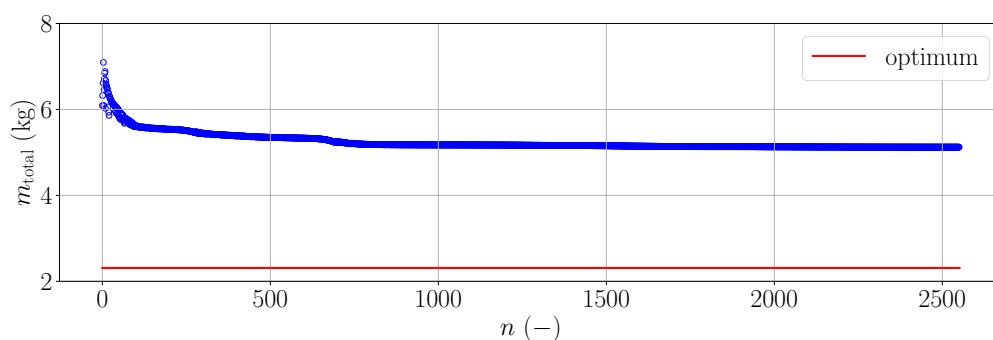
1.4 NUMERICKÉ ŘEŠENÍ

Daný problém byl řešen pro ilustraci v programu COMSOL pro ověření správnosti analytických předpokladů. Uvažujme nyní transformátor s konkrétními rozměry magnetického obvodu $a = 12,9$ mm, $b = 50,1$ mm a $c = 16,6$ mm a s tloušťkou $d = 43,3$ mm, viz obr. 64. Průřezy vodičů primárního a sekundárního vinutí jsou vyjádřeny parametry $S_1 = 32,5$ mm², $S_2 = 29,1$ mm². Počet závitů primární cívky je $n_1 = 641$. Tyto hodnoty představují optimum benchmarkového příkladu, viz [95].

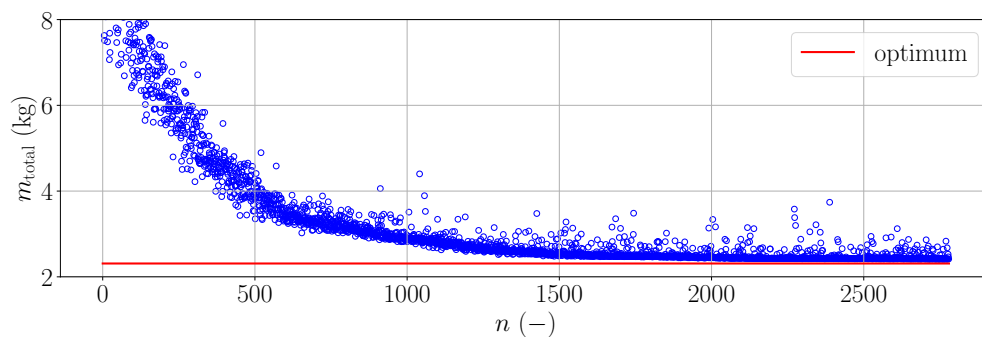
Řešení daného obvodu je zobrazeno na obr. 68 pro rozložení magnetické indukce v řezu transformátoru. Vzhledem k tomu, že u analytického řešení je velké množství jevů zanedbáno a výpočty jsou založeny na zjednodušených předpokladech, lze předpokládat, že numerické řešení bude více odpovídat realitě.



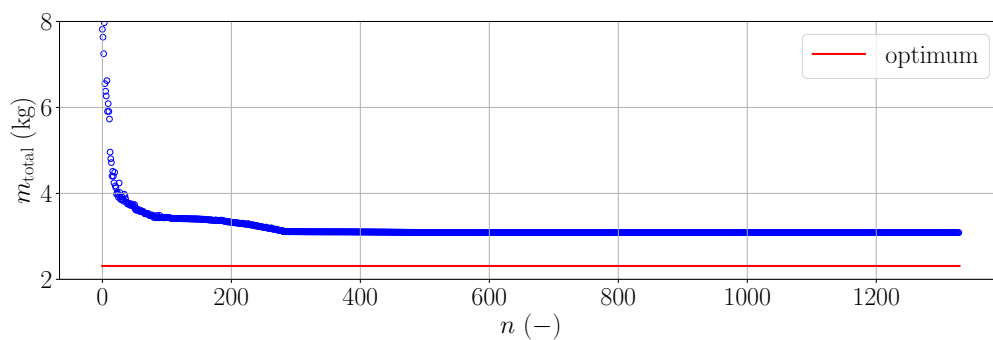
(a) Průběh optimalizační procesu řešeného metodou ISRES



(b) Průběh optimalizační procesu řešeného metodou Nelder-Mead

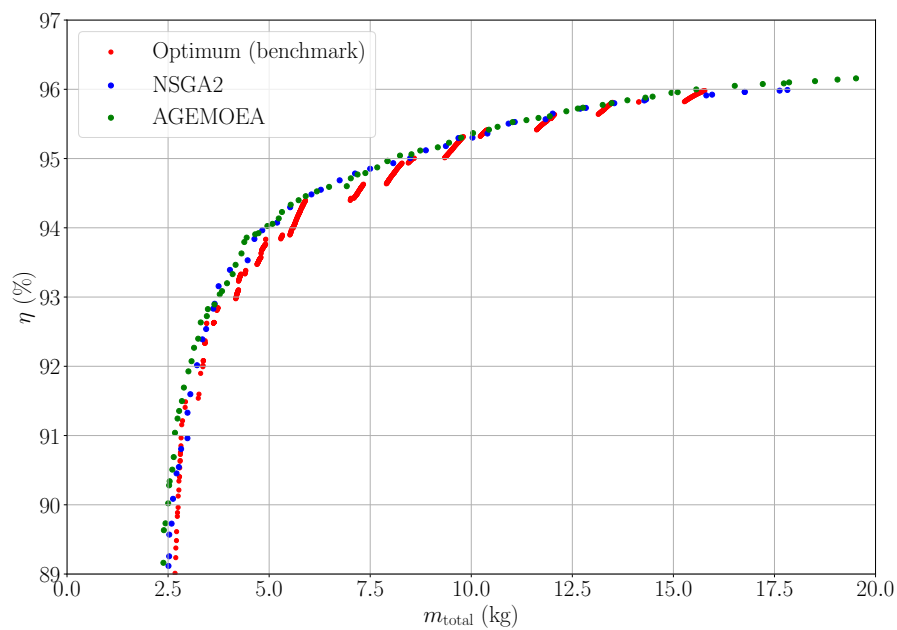


(c) Průběh optimalizační procesu řešeného metodou NSGA-II

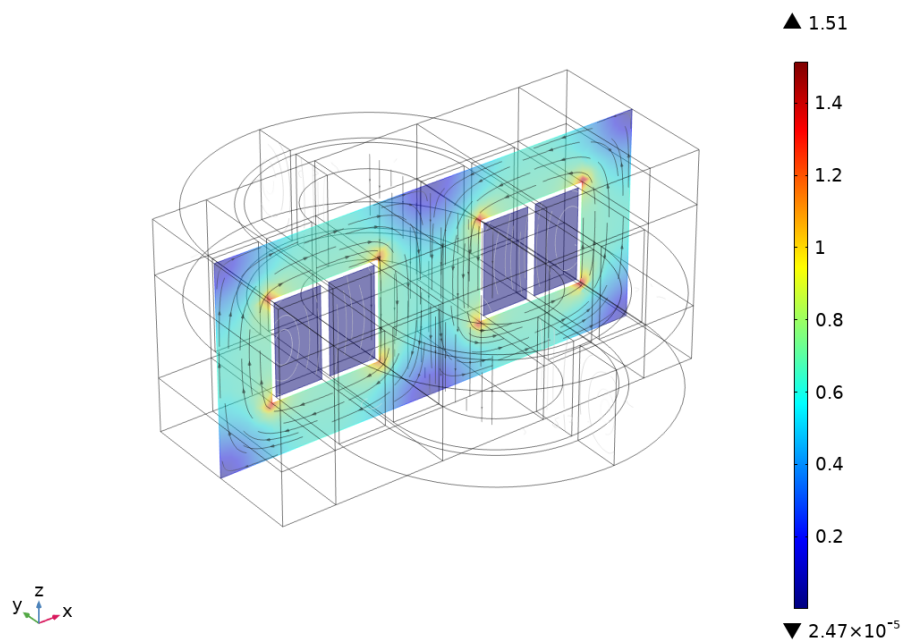


(d) Průběh optimalizační procesu řešeného metodou Pattern Search

Obr. 66: Porovnání několika metod při monokriteriální optimalizaci transformátoru



Obr. 67: Porovnání Paretoovy fronty s řešením testovacího problému



Obr. 68: Rozložení magnetické indukce v řezu transformátorem (ilustrativní řešení)

ANALYTICKÝ MODEL STEJNOSMĚRNÉHO MOTORU S PM

Tato kapitola studuje optimální návrh bezkomutátorového stejnosměrného motoru s permanentními magnety pro pohon solárního vozidla. Nejprve je představen analytický model problému, kdy je možné z jeho základních rozměrů a znalostí materiálových vlastností stanovit požadované parametry, v tomto případě minimální hmotnost motoru uloženého v kole vozidla s ohledem na maximální účinnost.

Celý analytický model lze vyjádřit pomocí 67 nelineárních rovnic obsahujících 78 parametrů, které popisují elektrické, magnetické a tepelné jevy. V této práci ovšem není dostatek prostoru se jimi zabývat, rovnice lze nalézt detailně popsané ve specifikaci benchmarku v [96]. Principiálně lze pomocí několika základních geometrických údajů a znalostí materiálových parametrů stanovit ostatní rozměry a požadovanou celkovou hmotnost.

2.1 ANALYTICKÝ MODEL

Pro výpočet účinnosti lze vycházet ze vztahu pro elektromagnetický výkon. Obecně lze elektromagnetický výkon definovat vztahem

$$P_{em} = M(t)\omega(t) = \sum_{i=1}^m e_{em,i}(t)i_i(t), \quad (17)$$

kde M značí elektromagnetický moment, ω vyjadřuje rychlost otáček, m je počet fází s elektromotorickou silou $e_{em,i}$ a zadaným proudem i_i . Pokud uvažujeme konstantní rychlost otáček, pak lze uvažovat elektromagnetický výkon rovněž konstantní (průběh jednotlivých napětí a proudů je patrný na obr. 69) a platí:

$$M\omega = 2E_{em}I, \quad (18)$$

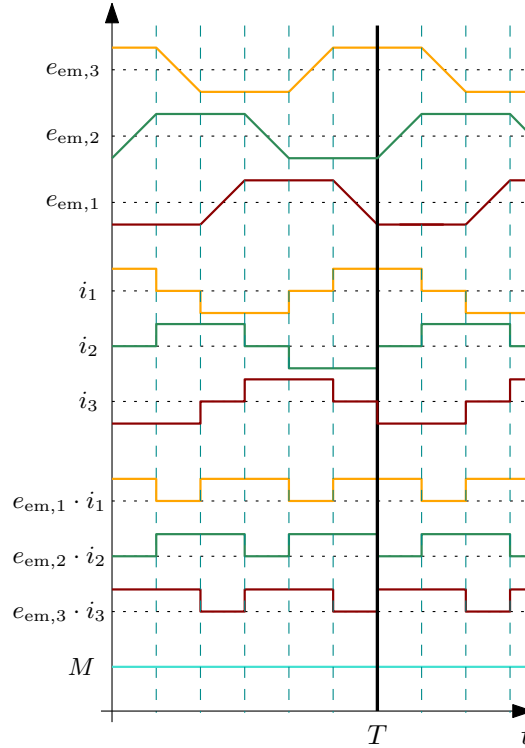
kde E_{em} značí elektromotorickou sílu jedné fáze a I proud jednou fází.

Nyní je možné určit elektromotorickou sílu, kterou lze po několika úpravách vyjádřit ve vztahu

$$E_{em} = \frac{n}{4}B_{air}D_sL_m\omega, \quad (19)$$

kde B_{air} je maximální magnetická indukce ve vzduchové mezeře, D_s je průměr statoru a L_m je magnetická délka motoru, tj. délka svazku plechů.

Z výše uvedených znalostí a znalostí materiálových parametrů je možné stanovit ostatní geometrické parametry, viz obr. 70, potřebné k výpočtu hmotnosti jednotlivých částí motoru a také hmotnosti jednotlivých částí, které nám dají celkovou požadovanou hmotnost zařízení m_{total} . Geometrické parametry a vztahy pro jejich výpočet jsou uvedeny podrobněji v [96].



Obr. 69: Elektromotorické síly, proudy a výkony pro tři fáze a výsledný moment

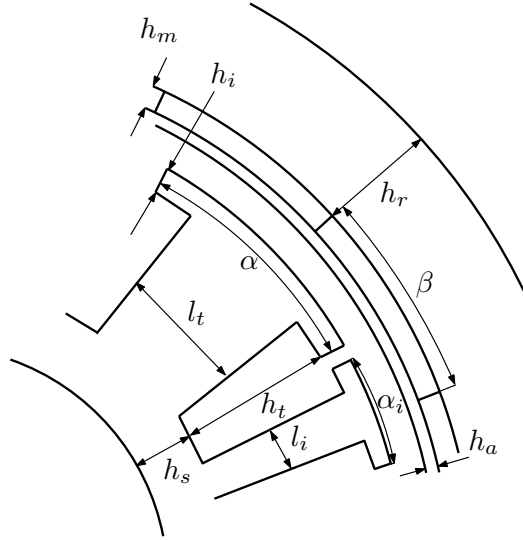
V případě, že je znám moment elektromagnetických sil, je možné určit magnetický tok, který je vyvolán permanentními magnety. Nejprve je potřeba určit hodnotu magnetického toku při chodu naprázdno vyvolaného jen pomocí permanentních magnetů. Pak lze určit magnetický tok na jednotlivých částech stroje. V případě, že je znám tok z magnetu, jsou potom díky znalosti základních rozměrů známy i hodnoty magnetické indukce mezi jednotlivými částmi stroje, podrobněji rozepsáno v [96].

Požadovanou hodnotu pro účinnost je možné určit v dalším kroku ze vztahu

$$\eta = \frac{M\omega - P_m}{M\omega + P_J + P_{Fe}}, \quad (20)$$

kde P_m jsou mechanické ztráty, P_J jsou Jouleovy ztráty, P_{Fe} jsou ztráty v železe a člen $M\omega$ představuje mechanický výkon.

Tepelný model je velmi jednoduchý. Na jedné straně se předpokládá, že tepelné odpory vedení jsou vždy velmi nižší než tepelné odpory konvekce. Teplota je tedy stejná ve všech částech statoru, které jsou v kontaktu, což jsou cívký, zuby a třmen statoru. Na druhou stranu, i když stator a rotor nejsou v kontaktu při otáčení, velké plochy jsou v kontaktu se vzduchem v řádu desetin milimetru. Za těchto podmínek je vzduch v mezeře modelován jako materiál se slabou tepelnou vodivostí. Navzdory tomu je tepelný odpor mezi rotorem a státorem zanedbatelný, protože se rovná tloušťce vzduchu dělené tepelnou vodivostí a povrchů, které jsou



Obr. 70: Rozměry vinutí, permanentních magnetů a vnějšího rotoru v radiálním řezu

v kontaktu. Gradient teploty uvnitř motoru lze tedy zanedbat s předpokladem, že všechny aktivní materiály mají stejnou teplotu:

$$T_a = T_{Cu} = T_{ext} + \frac{P_j + P_{Fe} + P_m}{hS_{ext}}, \quad (21)$$

kde h je koeficient přestupu tepla konvekcí, T_{ext} je počáteční teplota a S_{ext} je povrch motoru.

2.2 ILUSTRATIVNÍ PŘÍKLAD

Uvažujme nyní stejnosměrný motor s hodnotou magnetické indukce B_t v zubech motoru, B_{air} ve vzduchové mezeře a B_s v tělese statoru s průměrem otvoru D_s a proudovou hustotu označme J .

Dále uvažujme hodnotu pro elektromagnetický moment $M = 20 \text{ N} \cdot \text{m}$. a otáčky $\omega = 721 \text{ ot} \cdot \text{min}^{-1}$. Koeficient plnění je $k_r = 0,5$, koeficient roztažnosti $k_0 = 0,95$, koeficient úniku tepla mezi magnety a vzduchem $k_{m0} = 0,8$. Hodnota rezistivity pro měď je $\rho_{Cu} = 1,7 \cdot 10^{-8} \Omega \cdot \text{m}$. Hodnota remanentní indukce je $B_{r_m} = 1,045 \text{ T}$, $B_c = 1,045 \text{ T}$ značí mezní demagnetizační hodnotu indukce. Dále je známa měrná hmotnost statorové části $\rho_s = 7850 \text{ kg} \cdot \text{m}^{-3}$, magnetů $\rho_m = 7400 \text{ kg} \cdot \text{m}^{-3}$, cívek $\rho_L = 8950 \text{ kg} \cdot \text{m}^{-3}$ a hlavy rotoru $\rho_r = 7850 \text{ kg} \cdot \text{m}^{-3}$. Známý jsou také mechanické ztráty $P_m = 15 \text{ W}$, hodnota permeability permanentních magnetů $\mu_m = 1,05$, dále odchylky rezistivity magnetů $\alpha_a = -5 \cdot 10^{-4} \text{ K}^{-1}$ a cívek $\alpha_{Cu} = -3,8 \cdot 10^{-3} \text{ K}^{-1}$. Počáteční teplotu lze uvažovat $T_{ext} = 50 \text{ }^\circ\text{C}$, koeficient přestupu tepla konvekcí je $h = 10 \text{ W} \cdot \text{m}^{-2} \cdot \text{K}^{-1}$. Hodnota napětí je $U_{DC} = 120 \text{ V}$, magnetická indukce rotoru $B_r = 1,2 \text{ T}$, poměr délky rotoru a statoru $r_{rs} = 1,11$, počet pólů magnetu je $p = 6$. Frekvence je uvažována $f = 50 \text{ Hz}$.

2.3 OPTIMALIZACE

Autoři ve svém článku řeší optimalizační problém nejprve pro celkem 5, 7 a 11 návrhových proměnných. Pro testovací účely je v této práci optimalizace provedena jen pro variantu s pěti

návrhovými proměnnými, konkrétně pro magnetickou indukci jednotlivých zubů motoru B_t , vzduchové mezery B_{air} a statoru B_s , pro průměr otvoru statoru D_s a pro hodnotu proudové hustoty J . Tyto parametry budou předmětem optimalizace a jejich spodní a horní meze a počáteční hodnoty jsou uvedeny v tabulce 9.

Tab. 9: Vstupní parametry pro optimalizaci

Parametr	B_t (T)	B_{air} (T)	B_s (T)	D_s (mm)	J (A · mm ⁻²)
výchozí hodnota	1,0	0,65	1,1	250	3,0
spodní mez	0,9	0,50	0,6	150	2,0
horní mez	1,8	0,76	1,6	330	5,0

2.3.1 Vzorkování návrhového prostoru

Vzorkování návrhového prostoru bylo nejprve provedeno pomocí tradičních metod pro experimentální design – rovnoměrné rozdělení, LHS a faktoriální rozdělení. Náhradní modely (polynomiální aproximace druhého řádu, kriging, marginální Gaussův proces (MGP)) vyvinuté těmito experimenty byly silně nejednoznačné a výsledky se významně lišily od analytického řešení.

Vzhledem k tomu, že se jedná o problém, kdy jednotlivé návrhové proměnné spolu velmi nelineárně souvisí z důvodu několika omezení množiny možných řešení, byl nakonec použit přístup Mixture of Experts (MoE), který rozděluje návrhový prostor pro jeden globální model na několik lokálních expertů. Designový prostor je tak rozdělen do několika dílčích částí, které lze předvídat nezávisle. Výsledky prezentované v [98] ukazují, že výkon dynamické optimalizace lze tímto způsobem výrazně zlepšit, zejména u problémů s různými dynamickými pareto-optimálními řešeními v rozhodovacím prostoru v čase a problémů s vysoce nelineární vazbou rozhodovacích proměnných.

Prezentovaná implementace využívá algoritmus Expectation-Maximization (EM) pro Gaussian Mixture Models (GMM). Regresi lze rozdělit (viz [99]) do několika kroků:

1. **Clustering** – prostor parametrů je rozdělen na základě zvolené metody DoE.
2. **Školení expertů** – pro každý shluk je trénován vhodný náhradní model.
3. **Rekombinace** – vhodný náhradní model v každém shluku je kombinován pomocí GMM pro nejlepší výsledek regrese.

Uvažujme funkci $y_i(\mathbf{x})$ na shluku i . Jeho náhradním modelem je pak $\hat{y}_i(\mathbf{x})$. Pokud jsou tedy známy místní modely $\hat{y}_i(\mathbf{x})$, pak bude globální model ve tvaru

$$\hat{y}(\mathbf{x}) = \sum_{i=1}^K \mathbb{P}(\kappa = i | X = \mathbf{x}) \hat{y}_i(\mathbf{x}), \quad (22)$$

což je klasický pravděpodobnostní výraz pro MoE, kde K je počet jednotlivých lokálních modelů. $\mathbb{P}(\kappa = i | X = \mathbf{x})$, značeno hradlovou sítí, je pravděpodobnost, že leží ve shluku i . V případě MoE bylo použito vzorkování prostoru parametrů založené na klasické Haltonově

metodě.

Polynomiální aproximace druhého řádu byla použita jako náhradní model pro regresní problém s jedním výstupním parametrem. Kriging byl přidán pro dva výstupní parametry. V obou případech je prostor rozdělen na 10 expertů a výsledná interpolace pak dává vhodnou kombinaci. Ve výsledném modelu pro problém se dvěma výstupními parametry bylo použito 6 krigingových expertů a 4 polynomiální aproximace druhého řádu.

2.3.2 Jednokriteriální optimalizace

Cílem jednokriteriální optimalizace je maximalizovat účinnost η motoru. Protože jsou obecně optimalizační problémy řešeny jako minimalizační, využijeme k popisu cílové funkce popisující účinnost její zápornou hodnotu. Na tento problém je použito 5 omezujících vazeb – pro celkovou hmotnost m_{total} , vnější průměr D_{ext} , vnitřní průměr D_{int} , maximální proud I_{max} , teplotu motoru T_{Cu} . Omezme tedy obecnou cílovou funkci pro maximální účinnost η :

$$\begin{aligned} \min(-\eta), \quad \text{za podmíněk} \quad & m_{\text{total}} \leq 15 \text{ kg} & D_{\text{ext}} \leq 340 \text{ mm} \\ & D_{\text{int}} \geq 76 \text{ mm} & I_{\text{max}} \geq 125 \text{ A} \\ & T_{\text{Cu}} \leq 120 \text{ }^\circ\text{C} \end{aligned}$$

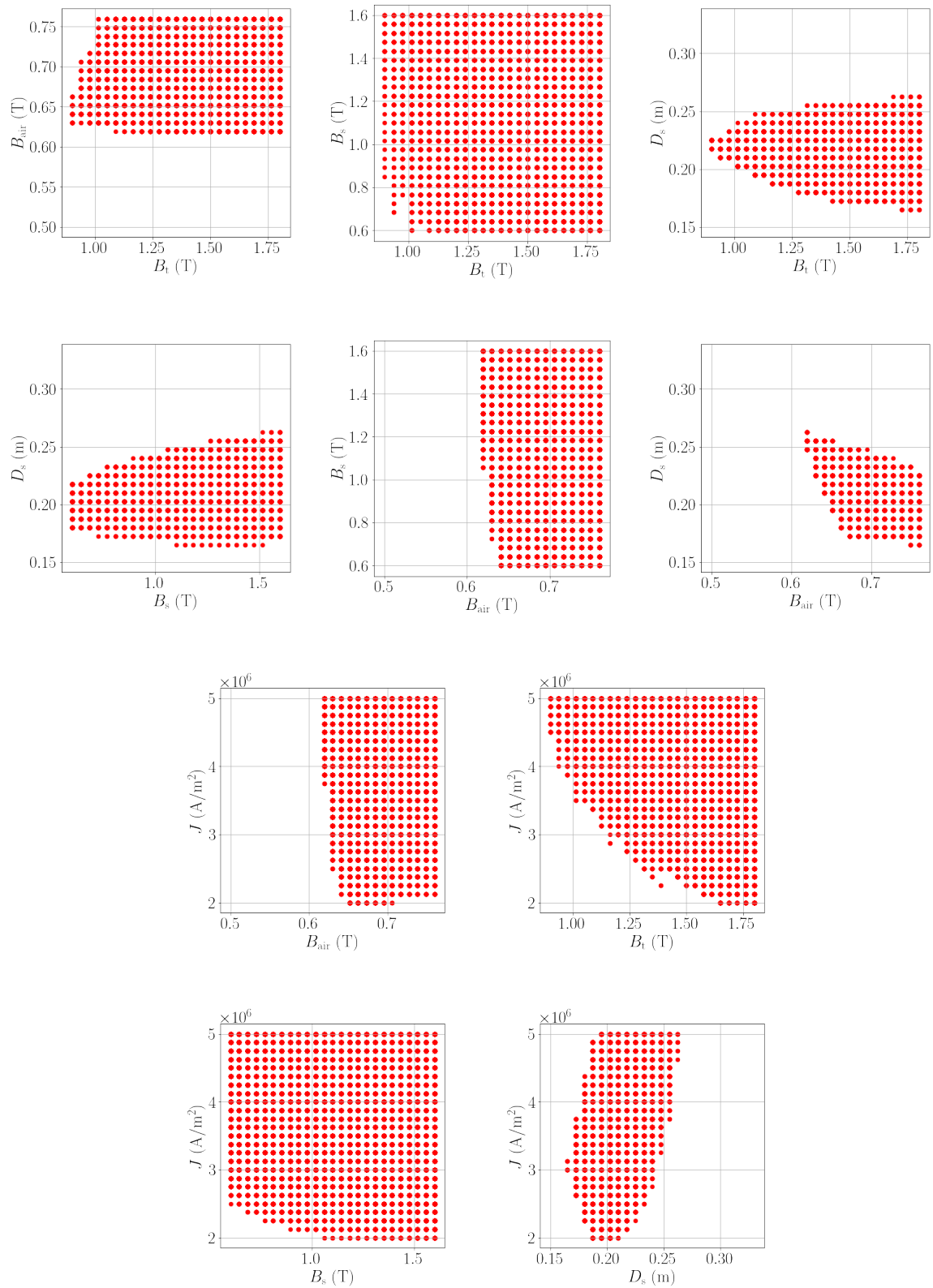
Na obr. 71 jsou zobrazeny povolené rozsahy jednotlivých parametrů v závislosti na daných omezeních.

Pro optimum benchmarkové úlohy jsou hodnoty pro omezení uvedeny v tabulce 9. Optimalizace byla nejprve provedena pro analytický model pomocí algoritmů Nelder–Mead, NSGA-II a Pattern Search. Následně byl optimalizován predikovaný model pomocí algoritmu Pattern Search. V tabulce 10 jsou hodnoty pro vstupní parametry s optimálními výsledky pro jednotlivé algoritmy s odchylkou od definovaného optima, kterým je v tomto případě původní analytické řešení. V tabulce 11 lze nalézt, jak jednotlivá řešení splňují požadovaná kritéria i s požadovanou hodnotou účinnosti η .

Tab. 10: Vstupní parametry a odchylka od cílové funkce

Metoda	B_t (T)	B_{air} (T)	B_s (T)	D_s (mm)	J (A · mm ⁻²)	odchylka (T)
Optimum	1,8000	0,648	0,896	201	2,04	0,000
Nelder Mead	1,8000	0,699	0,864	198	2,23	0,050
NSGA-II	1,7997	0,648	0,898	202	2,08	0,004
Pattern Search	1,8000	0,656	0,992	201	2,01	0,010
Pattern Search (predikce)	1,8000	0,735	0,929	195	2,36	0,009

Na obr. 72 je vždy v horním grafu znázorněna hodnota účinnosti η v závislosti pro jednotlivé parametry. Modrá křivka znázorňuje původní analytické řešení, které je bráno jako optimum úlohy a červená křivka znázorňuje predikovaný výsledek. Ve spodním grafu je pak zobrazena odchylka a ve spodním grafu je odchylka ϵ predikovaného a analytického řešení. Světle modrá oblast zohledňuje oblast možných řešení definovaných podle počátečního omezení.



Obr. 71: Povolený rozsah vstupních parametrů po aplikaci vazby

Tab. 11: Výstupní veličiny

Metoda	m_{total} (kg)	D_{int} (mm)	D_{ext} (mm)	T_{Cu} (°C)	I_{max} (A)	η (%)
Optimum	15,00	76,001	238,900	95,35	124,990	95,318
Nelder Mead	15,00	76,007	239,303	95,60	181,940	95,268
NSGA-II	15,00	79,954	240,049	94,90	125,660	95,314
Pattern Search	15,00	76,001	238,931	95,59	131,908	95,308
Pattern Search (predikce)	14,84	76,002	238,507	97,04	245,036	95,189

Průběh jednotlivých algoritmů s vyznačenou optimální hodnotou je pak patrný na obr. 73.

Vícekritériální optimalizace

Pro jednokritériální optimalizaci byla hlavním kritériem co největší účinnost stroje. Vzhledem ke konstrukčním požadavkům je druhým důležitým kritériem také celková hmotnost zařízení m_{total} , která by měla být co možná nejmenší. Konečné řešení pro návrh motoru je opět limitováno několika kritérii - pro vnější průměr D_{ext} , vnitřní průměr D_{int} , maximální proud I_{max} , teplotu motoru T_{Cu} . Omezme tedy obecné cílové funkce pro maximalizaci účinnosti η a minimalizaci hmotnosti m_{total} . Optimalizační problém lze nyní zapsat

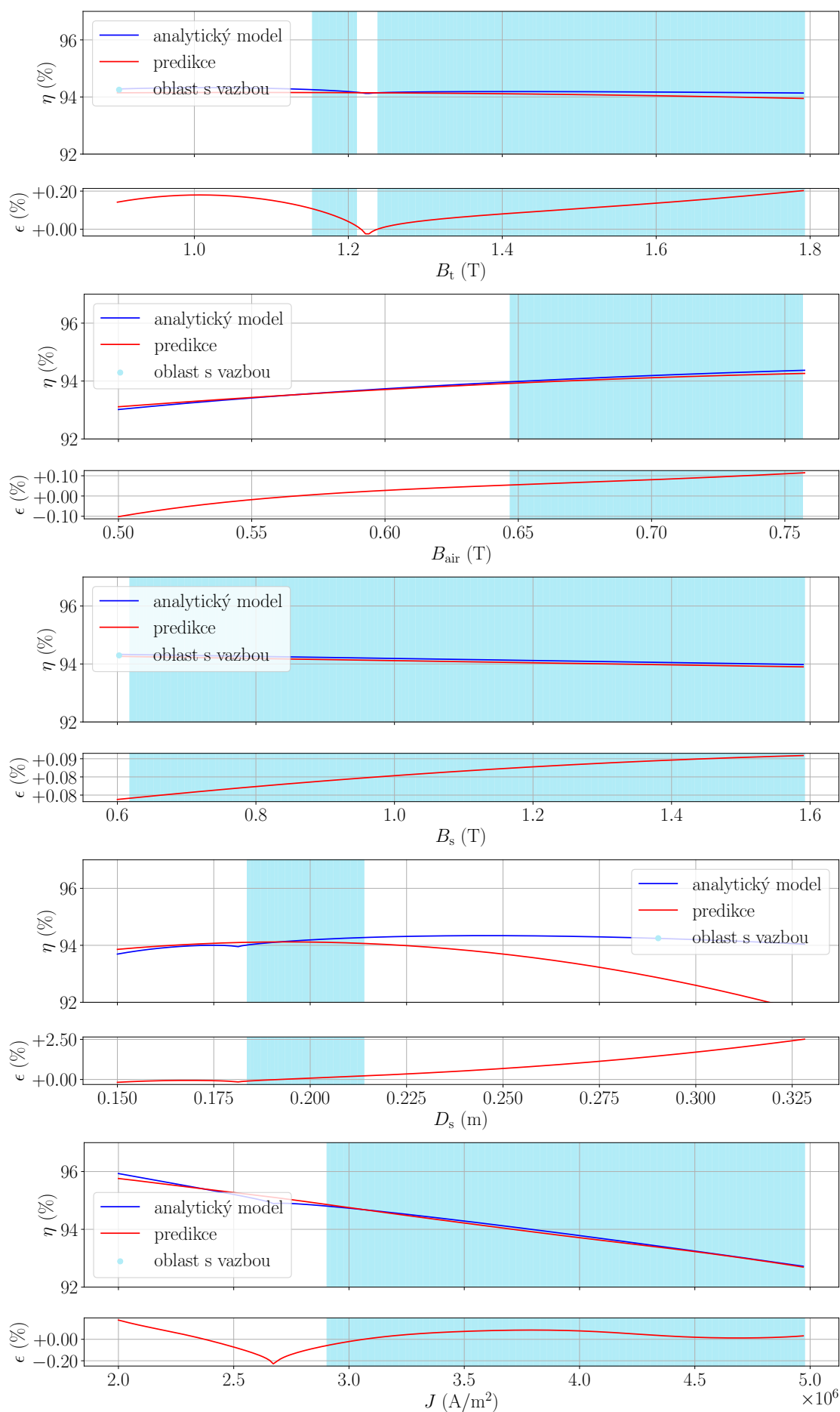
$$\begin{array}{llll} \min(-\eta), & \text{za podmínek} & T_{\text{Cu}} \leq 120 \text{ °C} & D_{\text{ext}} \leq 340 \text{ mm} \\ \min m_{\text{total}} & & D_{\text{int}} \geq 76 \text{ mm} & I_{\text{max}} \geq 125 \text{ A.} \end{array}$$

Řešení vícekritériální optimalizační úlohy bylo v tomto případě vyřešeno pomocí genetických optimalizačních algoritmů AGEMOEA a NSGA-II. Na obr. 74 je tentokrát v horním grafu znázorněno kromě řešení pro účinnost η (vlevo) také řešení pro požadovanou hodnotu hmotnosti m_{total} (vpravo) v závislosti na jednotlivých návrhových proměnných. Modrá křivka je opět pro původní analytické řešení, které je bráno jako optimum úlohy a červená křivka znázorňuje predikovaný výsledek. Ve spodním grafu je pak zobrazena odchylka a ve spodním grafu je odchylka ϵ predikovaného a analytického řešení. Světle modrá oblast zohledňuje oblast možných řešení definovaných podle počátečního omezení. Na obr. 75 je vykreslena Paretova fronta pro algoritmus AGEMOEA a NSGA-II pro výpočet založený na analytickém řešení v porovnání s výpočtem založený na predikovaném řešení pomocí algoritmu NSGA-II.

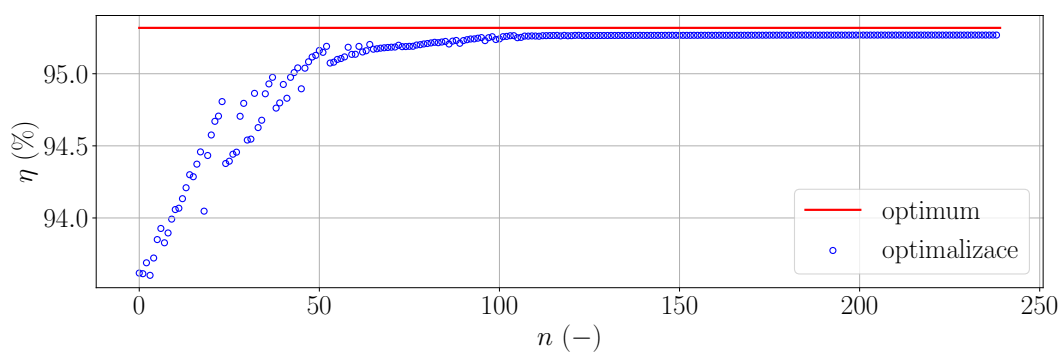
2.4 NUMERICKÉ ŘEŠENÍ

Daný problém byl řešen pro ilustraci v programu COMSOL pro ověření správnosti analytických předpokladů. Uvažujme nyní stejnosměrný motor s konkrétními hodnotami pro magnetickou indukci $B_t = 1,8 \text{ T}$ pro zuby motoru, $B_{\text{air}} = 0,65 \text{ T}$ pro vzduchovou mezeru a $B_s = 0,896 \text{ T}$ pro stator s průměrem otvoru $D_s = 201,24 \text{ mm}$. Rovněž je známa proudová hustota $J = 2,04 \text{ A/mm}^2$. Tyto hodnoty představují optimum získané z analytického řešení, uvedeném v [96].

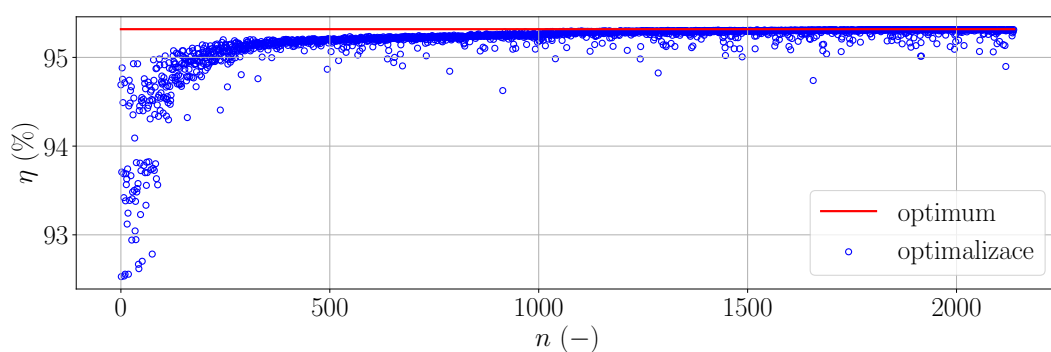
Řešení daného obvodu je zobrazeno na obr. 76 pro rozložení magnetické indukce v řezu motoru pro stav naprázdno. Vzhledem k tomu, že u analytického řešení je velké množství jevů zanedbáno a výpočty jsou založeny na zjednodušujících předpokladech, lze předpokládat, že numerické řešení bude více odpovídat realitě.



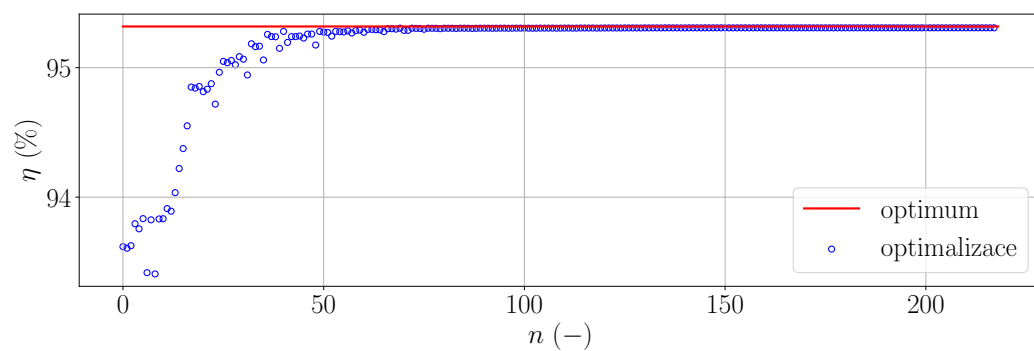
Obr. 72: Přesné řešení a predikce ve vybraných hladinách pro jedno kritérium



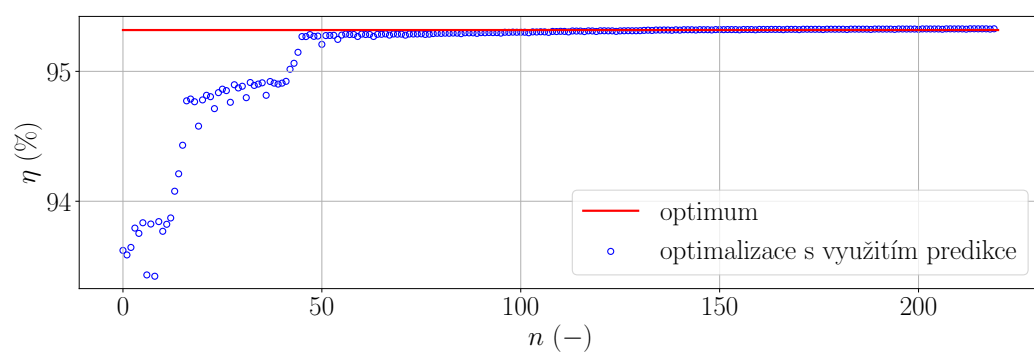
(a) Průběh optimalizační procesu řešeného metodou Nelder–Mead



(b) Průběh optimalizační procesu řešeného metodou NSGA–II

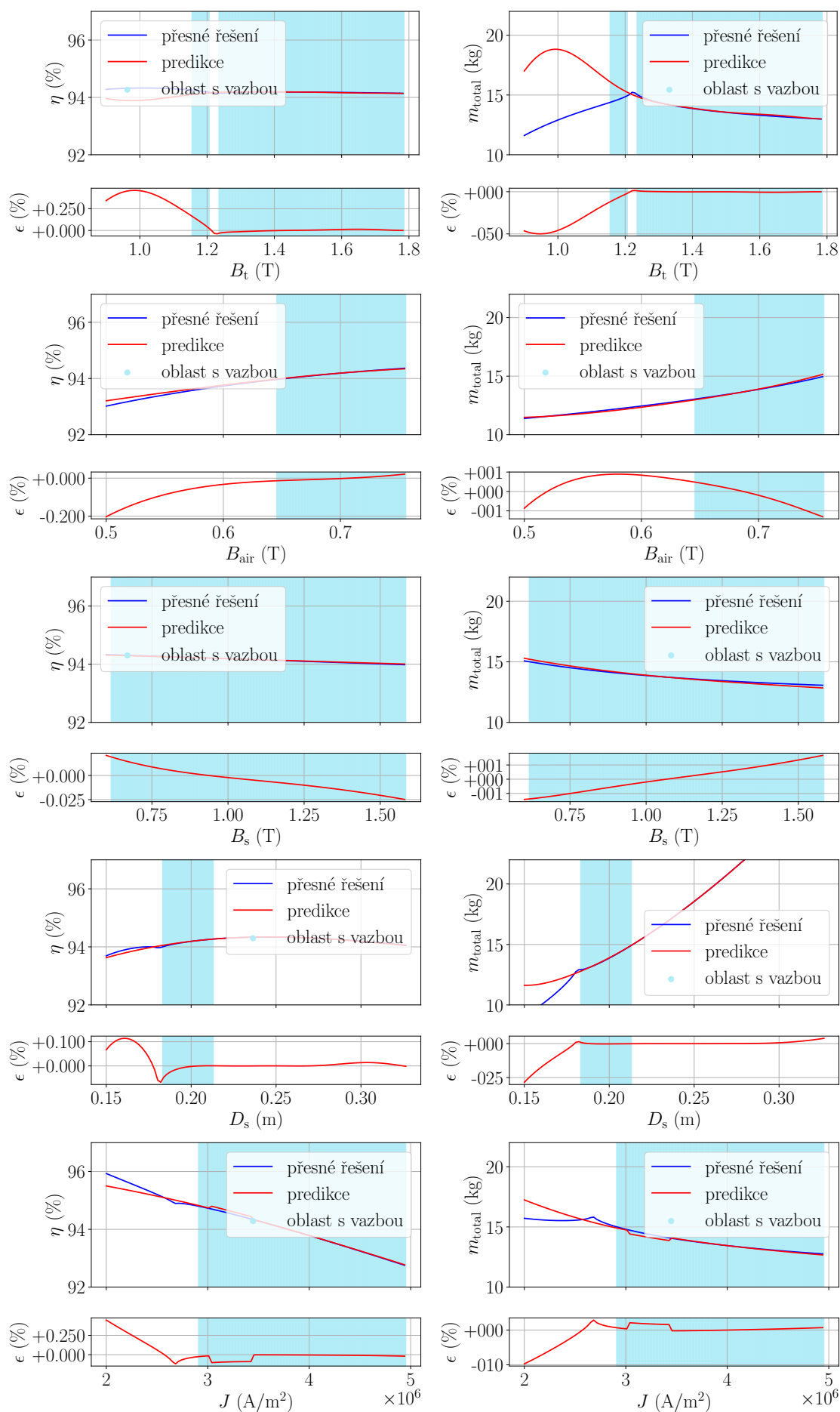


(c) Průběh optimalizační procesu řešeného metodou Pattern Search

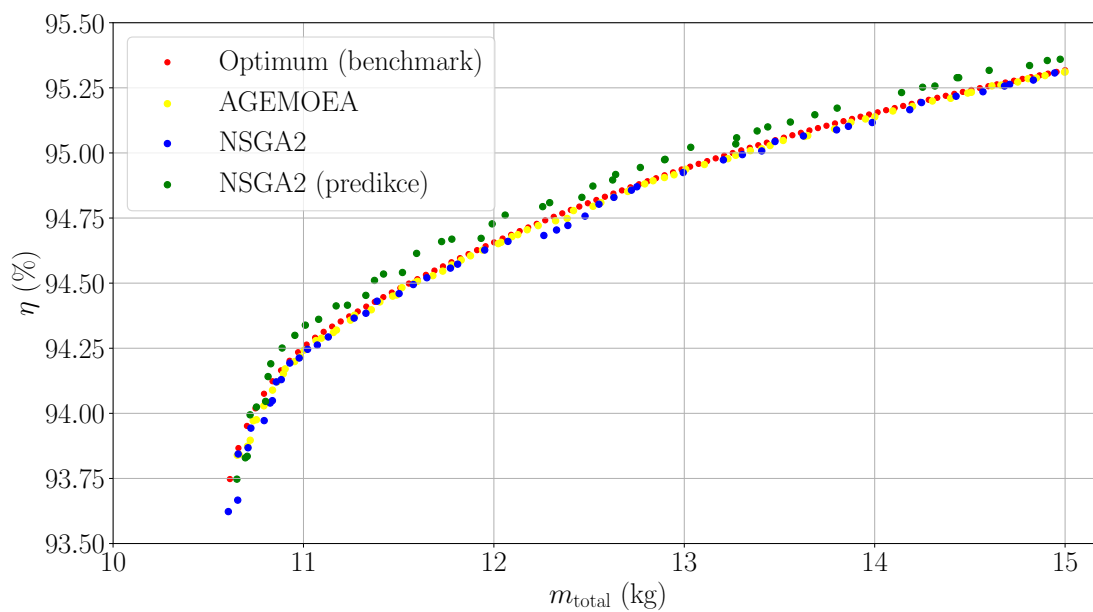


(d) Průběh optimalizační procesu řešeného metodou Pattern Search s využitím predikce

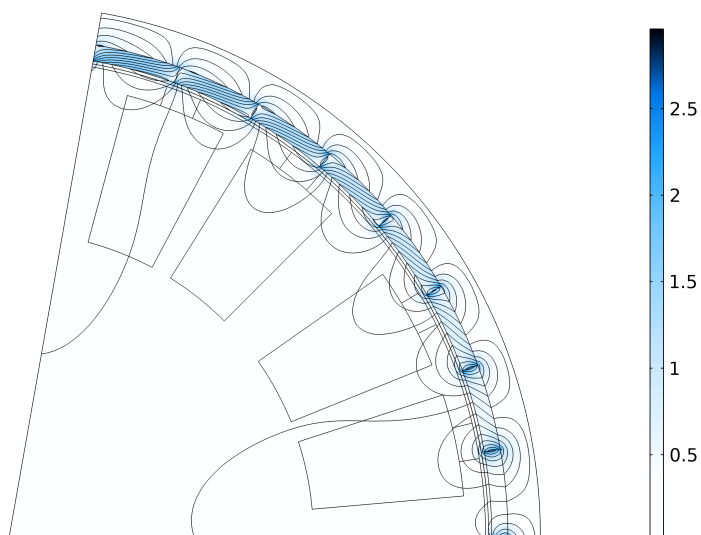
Obr. 73: Porovnání několika metod při jednokriteriální optimalizaci motoru



Obr. 74: Přesné řešení a predikce ve vybraných hladinách pro dvě kritéria



Obr. 75: Porovnání Paretoovy fronty s řešením testovacího problému



Obr. 76: Řešení numerického modelu v programu COMSOL

ANALYTICKÝ MODEL SÉRIOVÉHO RL OBVODU

V následující kapitole je řešen testovací příklad pro predikci průběhu proudu jednoduchého RL obvodu napájeného ze zdroje napětí. Při aproximaci je nutné použít vhodný náhradní model. K tomu je nutné vzít v úvahu následující aspekty a požadavky [100], [101]:

- nelinearita velmi ztěžuje předpovídání chování systému a odhad nejistoty pro model a výpočet,
- dynamické časoprostorové modely (DSTM) vícerozměrných systémů se pokoušejí charakterizovat interakce mezi různými stupnicemi variability, ale musí se vyrovnat s prokletím dimenzionality (hlavně v nelineárních systémech),
- struktura neurálních modelů by měla být přizpůsobena dynamice problému,
- modelování složitých časových závislostí v sekvenčních datech pomocí RNN, zejména dlouhodobých závislostí, zůstává otevřenou výzvou.

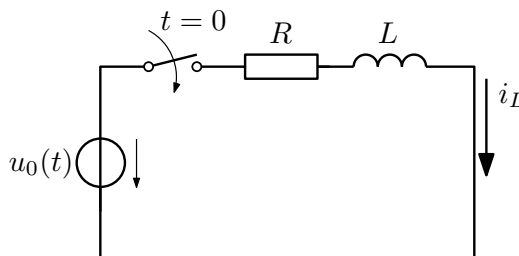
V následující kapitole je na základě těchto aspektů aplikován algoritmus rekurentní neuronové sítě pro predikci průběhu proudu v obvodu na základě znalosti vstupního napětí.

3.1 ANALYTICKÝ MODEL

Pro přechodný děj lze naformulovat diferenciální rovnici 1. řádu pro stavovou veličinu, která se mění spojitě, v tomto případě pro proud induktorem. Pro stavovou veličinu lze snáze určit počáteční podmínky, pro obvod na obrázku 77 je počáteční podmínka $i_0 = 0$ A. Matematicky je možné tuto rovnici zapsat následovně

$$L \frac{di_L}{dt} + Ri_L = u_0(t), \quad (23)$$

kde L představuje indukčnost, R rezistivitu, i_L proudu induktorem a $u(t)$ napětí zdroje.



Obr. 77: Schéma simulovaného obvodu

Řešením diferenciální rovnice je obecná rovnice ve tvaru

$$i_L(t) = K e^{\lambda t} + i_L(\infty).$$

Řešení této rovnice se skládá z

- řešení *homogenní rovnice* (na pravé straně je nula), tato část respektuje konfiguraci obvodu a lze ji zapsat ve tvaru součinu integrační konstanty K a exponenciální funkce, jejíž argument je dán kořenem charakteristické rovnice,
- z *partikulárního řešení*, které zohledňuje typ zdroje v obvodu. V případě obvodů se stejnosměrnými a harmonickými zdroji jej lze snadno určit z nového ustáleného stavu. Pro proud v obvodu v ustáleném stavu ($t \rightarrow \infty$) tedy platí

$$i_L(\infty) = \frac{U_{\text{stac}}}{R},$$

kde U_{stac} je po částech konstantní funkce (obdélníkový průběh) v intervalu od 0 do U_0 . Integrační konstantu K určíme z počátečních podmínek a její hodnota bude záviset na stavu, ve kterém se obvod nacházel před počátkem přechodného děje.

Levá strana rovnice je tedy homogenní rovnicí, jejíž tvar závisí na konfiguraci obvodu. K homogenní rovnici sestavíme charakteristickou rovnici a určíme její kořen

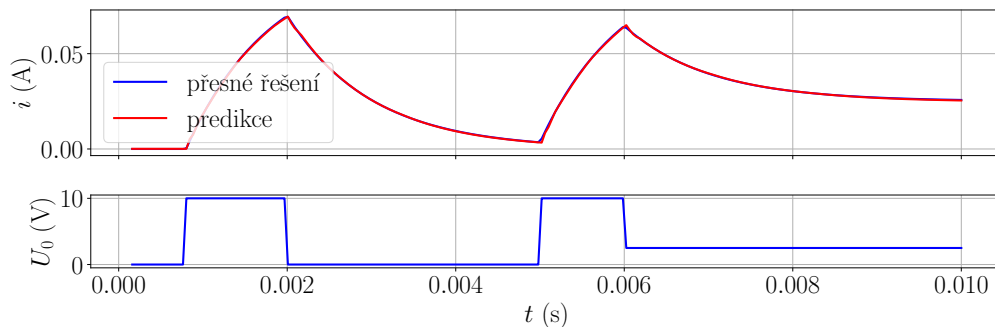
$$L\lambda + R = 0 \quad \rightarrow \quad \lambda = -\frac{R}{L}. \quad (24)$$

Konečné řešení je pak superpozicí řešení pro jednotlivé hodnoty napětí U_{stac} , které lze zapsat ve tvaru

$$i_L(t) = -\frac{U_{\text{stac}}}{R}e^{-\frac{R}{L}t} + \frac{U_{\text{stac}}}{R} = \frac{U_{\text{stac}}}{R} \left(1 - e^{-\frac{R}{L}t}\right).$$

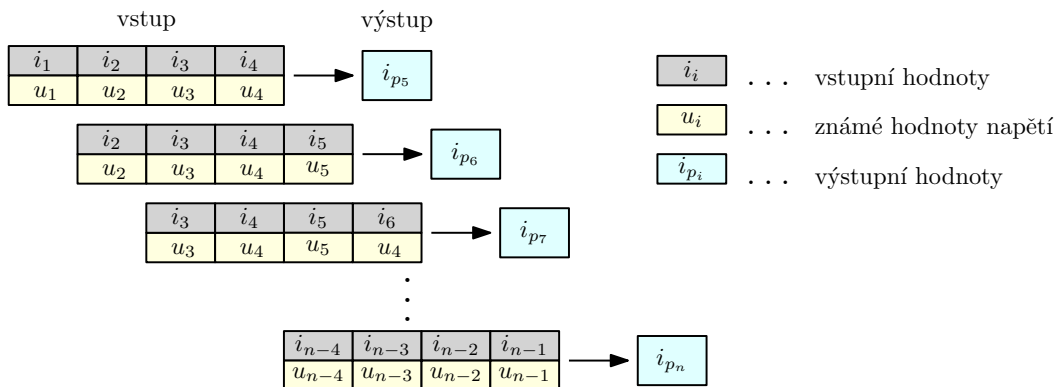
3.2 TRÉNOVÁNÍ NÁHRADNÍHO MODELU

Pro tvorbu modelu byla uvažována funkce pro trénování, jejíž průběh je znázorněn na obr. 78. Modrou čarou v horním grafu je znázorněna hodnota přesného řešení proudu induktorem $i_L(t)$, kterým je v tomto případě řešení diferenciální rovnice z předchozí kapitoly pro zvolené hodnoty napětí U_{stac} znázorněného ve spodním grafu.

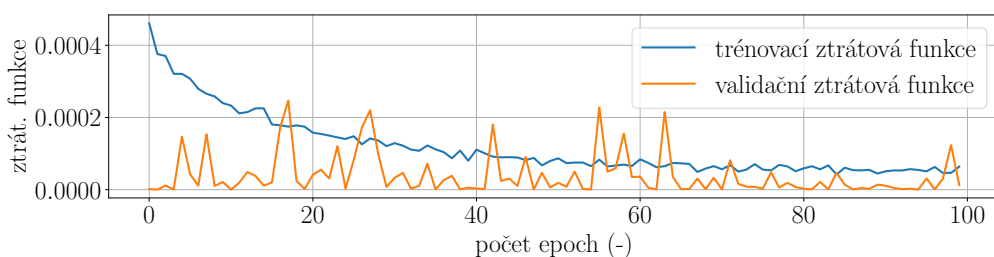


Obr. 78: Porovnání analytického řešení a trénovací funkce pomocí LSTM

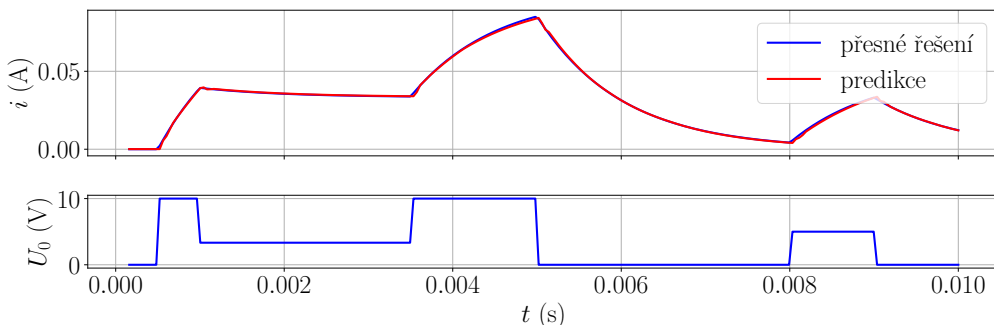
Červenou čarou na obr. 78 je vykreslena hodnota pro predikované řešení trénovací funkce. Princip trénování algoritmu RNN je znázorněn na obr. 79. Pro proud i_i pro $i = 1, \dots, n$ jsou známy první čtyři hodnoty i_1, i_2, i_3, i_4 na vstupu a výstupem je následující hodnota i_5 pro známé hodnoty napětí U_{stac} . V dalším cyklu je okno opět se čtyřmi vstupními hodnotami



Obr. 80: Princip testování natrénované sítě



Obr. 81: Trénovací a validační ztrátová funkce

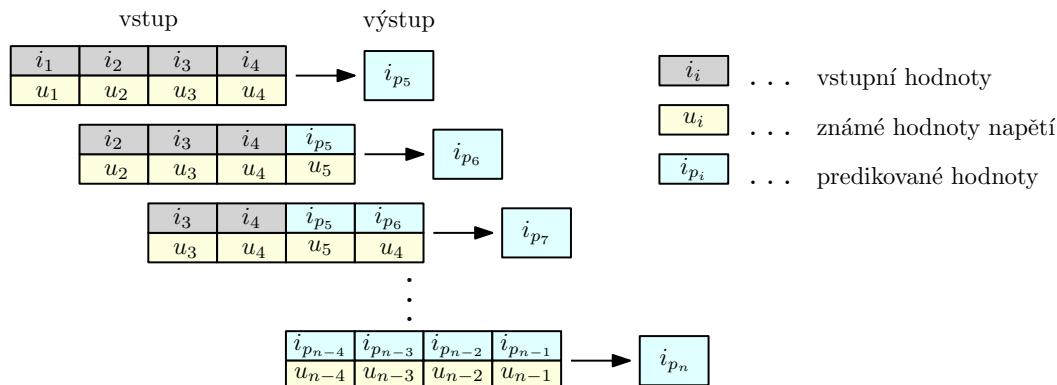


Obr. 82: Porovnání analytického řešení a validační funkce pomocí LSTM

3.3 PREDIKCE PROUDU PRO ZVOLENÉ NAPĚTÍ ZDROJE

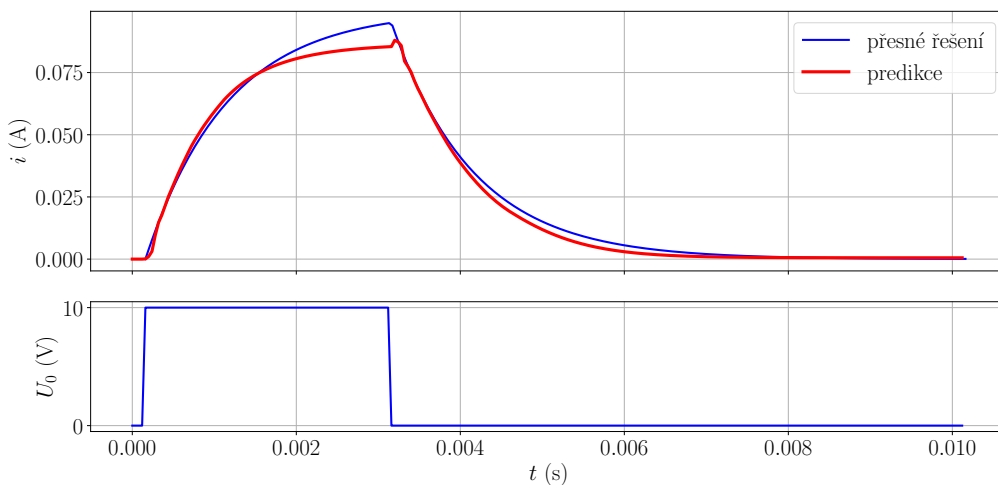
Nyní lze uvažovat opět známou časovou závislost napětí zdroje $u_0(t)$. Pro určení predikovaného průběhu jsou nyní uvažovány jako počáteční podmínka opět první čtyři hodnoty proudu i_1, i_2, i_3, i_4 a napětí $u_{0,1}, u_{0,2}, u_{0,3}, u_{0,4}$ a výstupem je predikovaná následující hodnota proudu i_{p_5} . Následně se opět posune celé okno stejně jako v předchozích případech, ovšem nyní v následujícím kroku již uvažujeme ve vstupním okně hodnoty proudu i_2, i_3, i_4 s predikovanou hodnotou i_{p_5} získanou na výstupu v předchozím kroku. Výstupní hodnotou je nyní predikovaná hodnota i_{p_6} , která se v dalším kroku opět přidá do vstupní vrstvy. Princip je velmi dobře patrný na obr. 83. Pro predikci jsou tedy použity počáteční hodnoty proudu

a napětí a následně predikované hodnoty proudu ve spojení se známou hodnotou napětí.



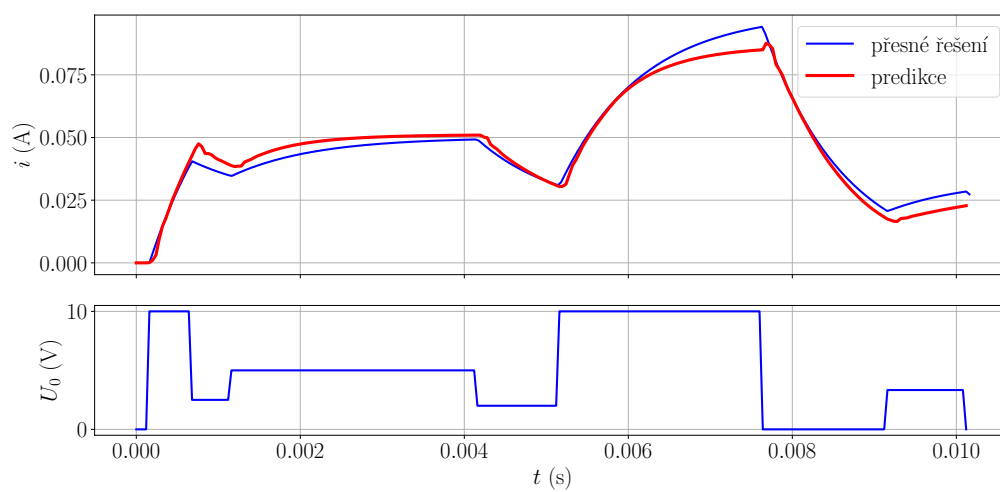
Obr. 83: Základní princip predikce proudu

Řešení predikovaného průběhu je simulováno pro dva testovací průběhy, viz obr. 84 a 85. Ve spodní části grafů je znázorněna hodnota časově závislého napětí $u_0(t)$ a v horním grafu je v obou případech vykreslen průběh proudu. Modrou čarou je opět znázorněno analytické řešení diferenciální rovnice a červenou barvou predikované řešení.



Obr. 84: Porovnání analytického řešení s predikcí pomocí LSTM (testovací příklad 1)

Řešení pomocí predikce s využitím rekurentních neuronových sítí (RNN) lze chápat za specifických podmínek jako možnou alternativu výpočtu časového průběhu problémů, které lze popsat pomocí obyčejných diferenciálních rovnic. Svě využití může najít u zařízení s omezenými zdroji.



Obr. 85: Porovnání analytického řešení s predikcí pomocí LSTM (testovací příklad 2)

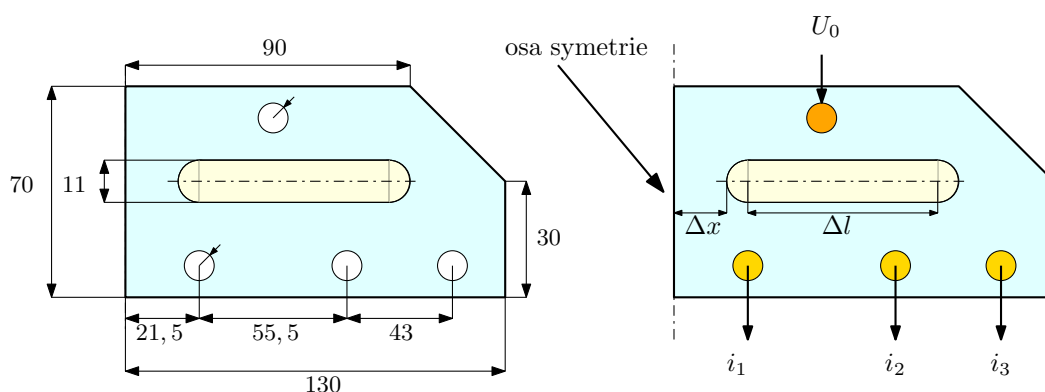
Část V

ILUSTRATIVNÍ PŘÍKLADY

LAMELOVÝ PŘÍVOD

Jedním z ilustrativních příkladů je návrh lamel frekvenčního měniče. V této kapitole bude demonstrováno využití aproximačního modelu a následné porovnání celkové doby optimalizace plného modelu a optimalizace provedené na modelu náhradním.

Průtokem elektrického proudu se ohřívají přívodní lamely stejnosměrné části měniče nerovnoměrně Jouleovými ztrátami, přičemž teplota samotného zařízení nesmí přesáhnout danou mez (u mědi je to zpravidla $150\text{ }^{\circ}\text{C}$). Fyzikální parametry lamel se mění s teplotou (elektrická vodivost, tepelná vodivost, specifické teplo, měrná hmotnost), jedná se proto o sdružený problém teplotního a proudového pole. To znamená, že v každém časovém kroku by měl být brán ohled na změny materiálových vlastností, což je vzhledem ke komplikovanému uspořádání lamely výpočetně poměrně náročná úloha.



Obr. 86: Obecné schéma lamely frekvenčního měniče s parametry a rozměry v mm

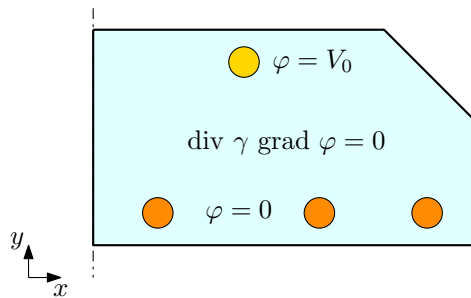
Nerovnoměrný ohřev zařízení je dán nerovnoměrným rozložením proudu na výstupních svorkách lamely, jejíž hlavní rozměry jsou vyznačené v levé části obr. 86. V jeho pravé části jsou naznačené rozměry Δl a Δx , jež je nutno optimalizovat, dále vstupní svorka připojená k napětí U_0 a výstupní svorky, z nichž se odebírají proudy i_1 , i_2 a i_3 . Tyto proudy musí být pokud možno stejné, jejich velikosti se mohou lišit jen minimálně. Problematika tvarové a topologické optimalizace byla prezentována jako příspěvek na mezinárodní konferenci *19th International Conference Computational Problems of Electrical Engineering (CPEE)*, viz [MK.3].

1.1 MATEMATICKÝ MODEL A SIMULACE ÚLOHY

Matematický model proudového pole v lamelle lze popsat pomocí parciální diferenciální rovnice (viz rovněž obr. 87)

$$\operatorname{div}(\gamma \operatorname{grad} \varphi) = 0, \quad (25)$$

kde γ značí elektrickou vodivost jejího materiálu.



Obr. 87: Základní uspořádání a matematický model

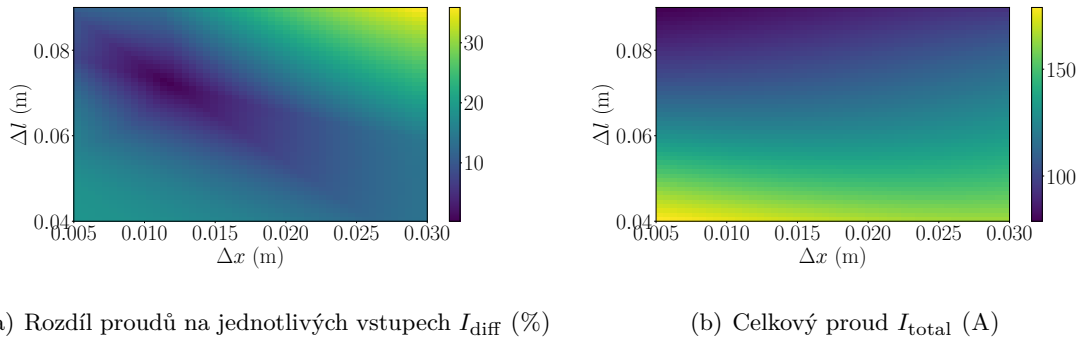
Jednoznačnost řešení je dána okrajovými podmínkami. V tomto případě jsou podmínky určeny pro hranici lamely, a její vstupní i výstupní svorky. Podél hranice lamely platí pro potenciál Neumannova okrajová podmínka, zatímco potenciály výstupních svorek již byly zmíněny. Potenciál vstupní svorky budiž 1 mV.

Simulace byla provedena v programu Agros Suite pro rozměry již vyrobeného zařízení, viz obr. 88. Jak bylo zmíněno výše, cílem je najít takové hodnoty parametrů Δx a Δl , které budou splňovat požadavky pro minimální odpor vodivých cest a zároveň rovnoměrné rozložení proudů na výstupních svorkách lamely, aby se zamezilo jejímu nadměrnému lokálnímu ohřevu. Aby bylo možné ukázat vliv obou parametrů na výsledné řešení, byla v první fázi provedena parametrická analýza, kdy parametr délky otvoru Δl nabýval hodnoty v rozmezí $\langle 0,04, 0,09 \rangle$ a parametr vzdálenosti otvoru od středu zařízení Δx hodnot $\langle 0,005, 0030 \rangle$. Výsledné řešení je znázorněno na obr. 89.



Obr. 88: Rozložení proudových hustot na vyrobené lamelle

Dalším ukazatelem, který lze sledovat, je odpor proudových cest, jenž má být vhodným způsobem omezen. Vstupními parametry jsou:

Obr. 89: Závislost cílových funkcí na hledaných parametrech Δx a Δl

- rozměr výřezu pro ukotvení měniče, přesněji vzdálenost od osy symetrie Δx a
- délka výřezu Δl , viz obr. 86.

Celková analýza zařízení byla pak dále rozšířena o výpočet oteplení lamelových přívodů. Tento výpočet nevyužívá pokročilé metody analýzy a proto není součástí předložené dizertační práce.

1.2 OPTIMALIZACE PLNÉHO MODELU

Na plném modelu byla provedena monokriteriální optimalizace cílové funkce pro celkový proud a vícekriteriální optimalizace celkového proudu a jeho rozložení na jednotlivé výstupy.

1.2.1 Monokriteriální optimalizace

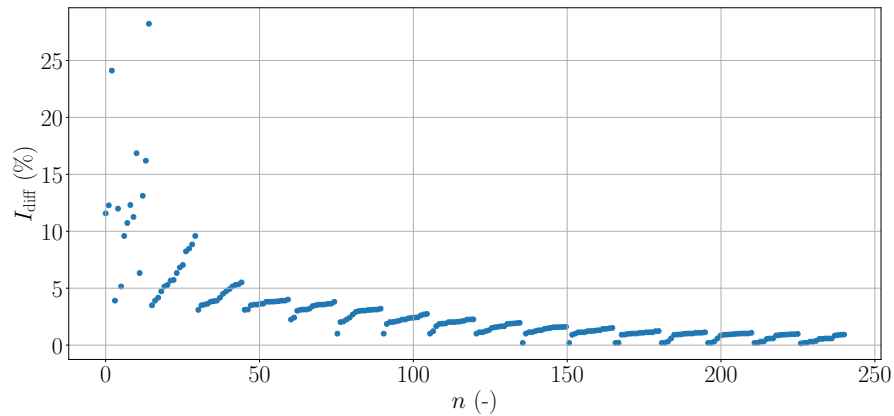
Hlavním ukazatelem pro rovnoměrný ohřev zařízení by mělo být co nejrovnoměrnější rozložení proudu na výstupech zařízení. V prvním kroku byla tedy provedena monokriteriální optimalizace na plném modelu, kdy lze cílovou funkci vyjádřit jako řešení co nejmenšího rozdílu mezi hodnotami i_1 , i_2 a i_3 . Cílová funkce má tedy tvar

$$i_{\text{diff}}(\%) = \frac{|(i_1 - i_2)| + |(i_2 - i_3)| + |(i_3 - i_1)|}{3} \cdot 100. \quad (26)$$

Optimalizace pro jednu cílovou funkci je znázorněna na obr. 90 a byla provedena pomocí genetického optimalizačního algoritmu NSGA-II pro 15 populací o 15-ti jedincích. Ačkoliv je rozvážení hodnot proudů na jednotlivých výstupech velmi malé, znamená podstatnou změnu hodnoty celkového proudu. Proto bylo v dalším kroku potřeba přistoupit k vícekriteriální optimalizaci.

1.2.2 Vícekriteriální optimalizace

Optimalizaci tvaru lamely frekvenčního měniče lze tedy definovat jako multikriteriální optimalizaci se dvěma cílovými funkcemi. Prvním kritériem je tvar a rozteč výstupních svorek na základě minimálního rozdílu hodnoty proudů na výstupu. Druhé kritérium je stanovení



Obr. 90: Monokriteriální optimalizace plného modelu

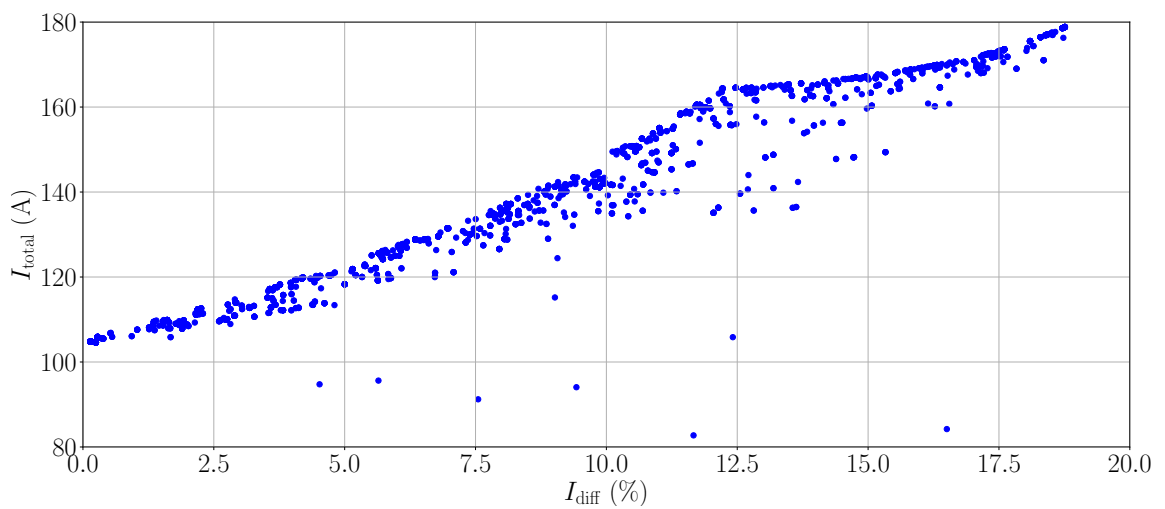
závislosti vyvážení proudových cest v závislosti na jejich minimálním odporu, to lze vyjádřit ve tvaru

$$R_{\text{int}} = \frac{U_0}{i_{\text{total}}}, \quad (27)$$

kde i_{total} značí celkový proud na výstupech zařízení. Ke snížení hodnoty odporu vodivých cest bude potřeba maximalizovat právě hodnotu celkového proudu. V tomto případě hledáme tedy minimum záporné hodnoty i_{total} .

Optimálním řešením bude tedy sada hodnot Pareto fronty pro dvě cílové funkce. Vstupními parametry budou opět vzdálenost od osy symetrie Δx a délka výřezu Δl .

Pro plnou optimalizaci byl opět využit genetický algoritmus NSGA-II se 150 generacemi o 15-ti jedincích. Průběh optimalizačního procesu se zobrazou Pareto frontou je znázorněn na obr. 91.



Obr. 91: Vícekriteriální optimalizace plného modelu

1.3 OPTIMALIZACE S VYUŽITÍM NÁHRADNÍHO MODELU

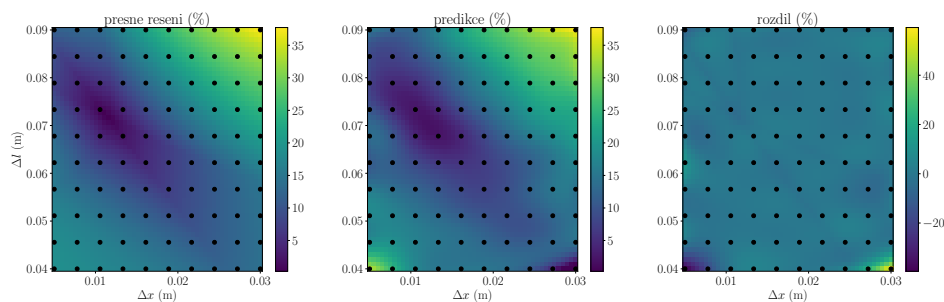
Vzhledem k rychlosti počítačového řešení tohoto problému v aplikaci Agros Suite bylo provedeno plné vyčíslení obou cílových funkcí pro všechny kombinace parametrů návrhového prostoru. Obrázky 89 (a) a (b) slouží pro ilustraci vývoje obou cílových funkcí. Je třeba poznamenat, že takto detailní analýza není u složitějších modelů možná.

Vzorkování návrhového prostoru

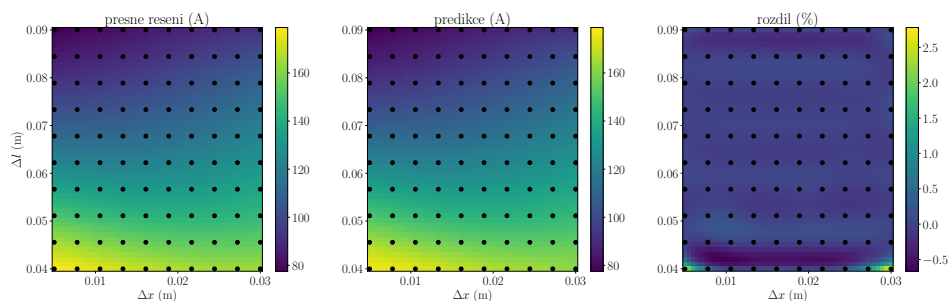
Vzorkování návrhového prostoru bylo provedeno pomocí statických i adaptivních metod návrhu experimentu.

Statické vzorkování

Na obr.92–95 jsou znázorněna řešení pro statické vzorkování. Aproximace prostoru byla provedena pro obě cílové funkce, tedy pro řešení rozdílu proudů na výstupech a pro řešení celkového proudu. Počet návrhových bodů je 100 a pro návrh bylo zvoleno rovnoměrné vzorkování v celém návrhovém prostoru. Na obrázcích vlevo je vždy znázorněno přesné řešení. Prostřední obrázek zobrazuje predikci jednotlivých aproximačních algoritmů, obrázek napravo vyjadřuje procentuální rozdíl (%) predikce a původního řešení. Postupně bylo porovnáno několik metod aproximace cílové funkce a v dalším textu jsou pak zobrazeny pouze některé z nich. Jedná se o gaussovský proces (obr. 92), rozhodovací stromy (obr. 93), Kriging (obr. 94) a KPLSK (obr. 95).

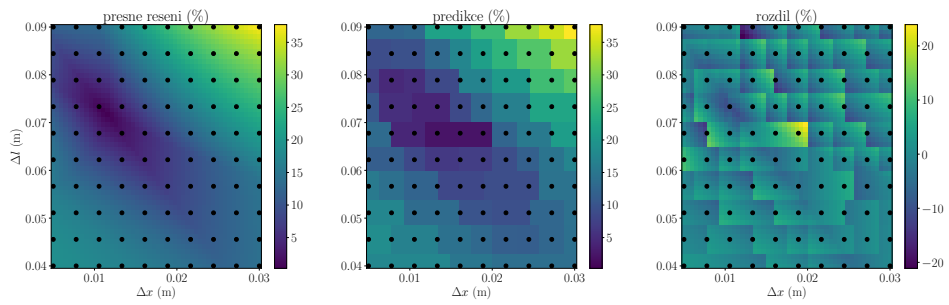


(a) Rozdíl proudů na jednotlivých výstupech

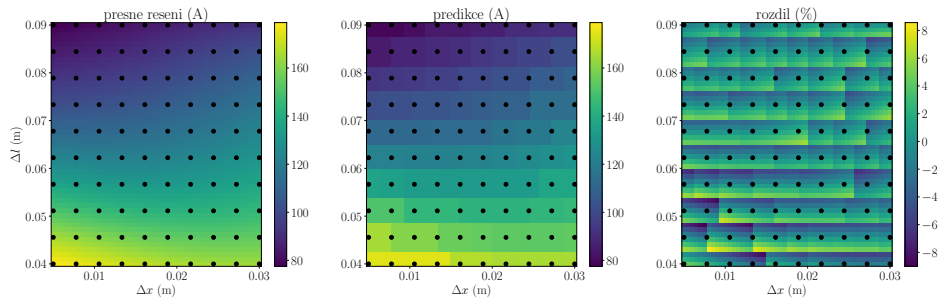


(b) Celkový proud

Obr. 92: Statická aproximace navzorkovaného prostoru - gaussovský proces

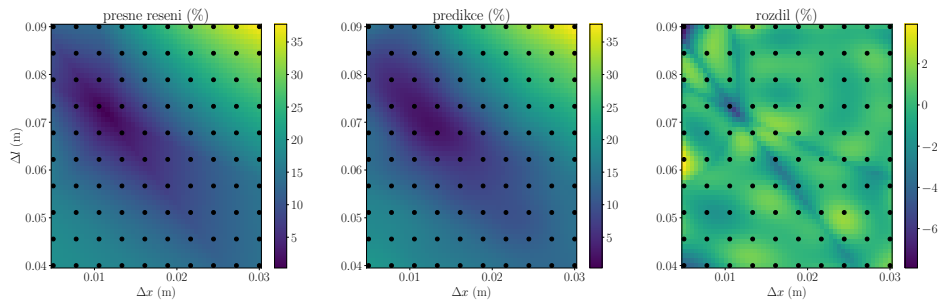


(a) Rozdíl proudů na jednotlivých výstupech

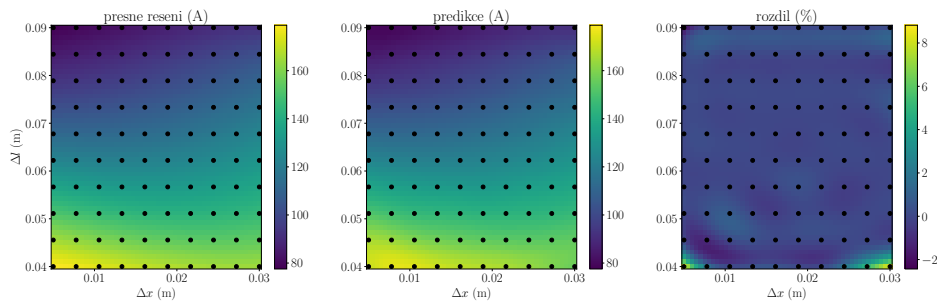


(b) Celkový proud

Obr. 93: Statická aproximace navzorkovaného prostoru - rozhodovací stromy

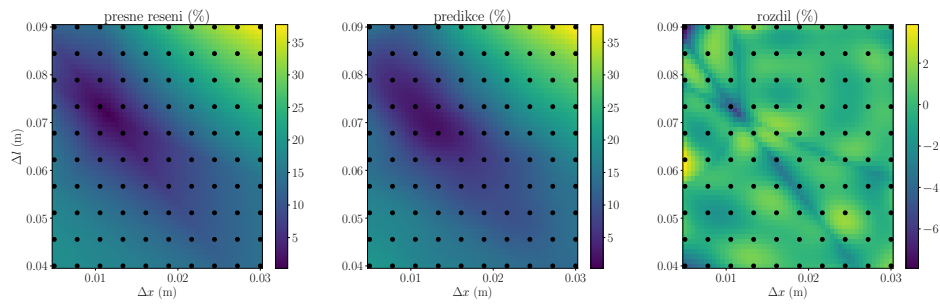


(a) Rozdíl proudů na jednotlivých výstupech

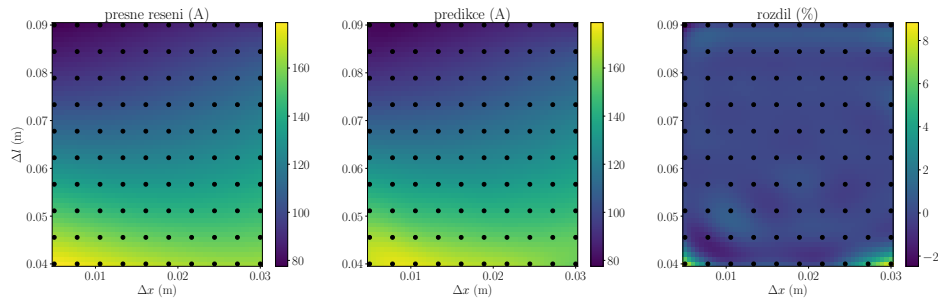


(b) Celkový proud

Obr. 94: Statická aproximace navzorkovaného prostoru - Kriging



(a) Rozdíl proudů na jednotlivých výstupech



(b) Celkový proud

Obr. 95: Statická aproximace navzorkovaného prostoru - KPLSK

V tabulce 13 jsou porovnány jednotlivé aproximace z hlediska určení průměrné absolutní chyby (MAE), střední chyby čtverce (MSE), variační míry (V), která určuje rozptyl daného souboru dat, a koeficientu rozptylu (R^2), který poskytuje zpětnou vazbu o tom, jak úspěšně bude model odhadovat budoucí hodnoty. Celkový počet 100 návrhových bodů byl rozdělen na sadu pro trénovací data (67 bodů) a testovací data (33 bodů).

Tab. 13: Porovnání aproximačních algoritmů pro rovnoměrně rozdělený návrhový prostor

Aproximace	Trénovací data				Testovací data			
	MAE	MSE	V	R^2	MAE	MSE	V	R^2
Gaussův proces	0,0	0,0	1,0	1,0	1,37	13,69	0,67	0,67
Rozhodovací stromy	0,0	0,0	1,0	1,0	1,92	5,19	0,93	0,92
Kriging	0,0	0,0	1,0	1,0	0,83	3,35	0,99	0,99
KPLSK	0,0	0,0	1,0	1,0	0,83	3,35	0,99	0,99

Adaptivní vzorkování

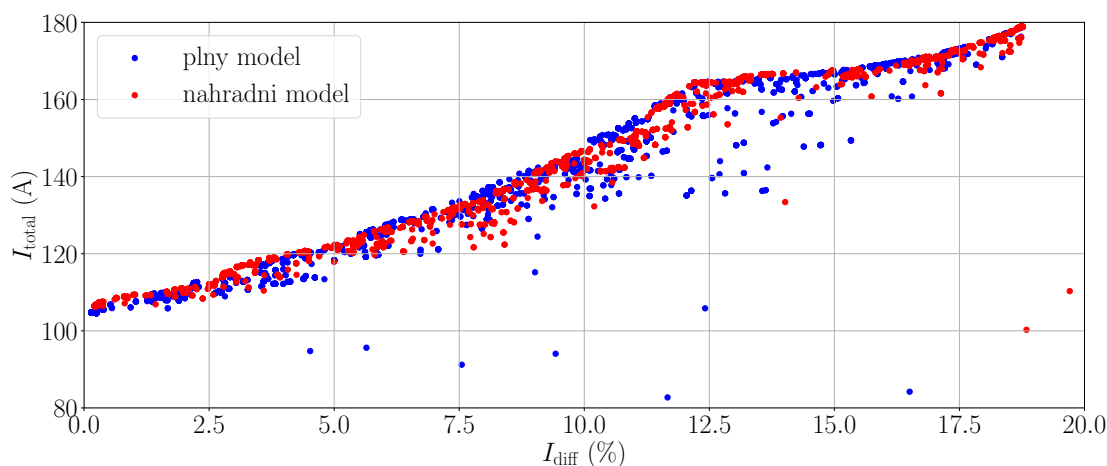
Jednou z variant možné tvorby náhradního modelu je také adaptivní vzorkování návrhového prostoru. Na obr. 97 je znázorněn proces adaptivního návrhu experimentu založený na gaussovském procesu pro první cílovou funkci – rozdílu proudů na jednotlivých výstupech. Během návrhu je sledována hodnota maximální směrodatné odchylky cílové funkce v návrhovém prostoru a na jejím základě určen nový bod, který je následně přidán do zmíněného prostoru. Výsledkem je poměrně rychlá a přesná aproximace cílové funkce náhradním modelem. Obrázek znázorňuje výběr z prvních 150 bodů. V dalších iteracích se již aproximace výrazně neměnila.

Tab. 14: Porovnání optimalizačního procesu s náhradním modelem a bez něho

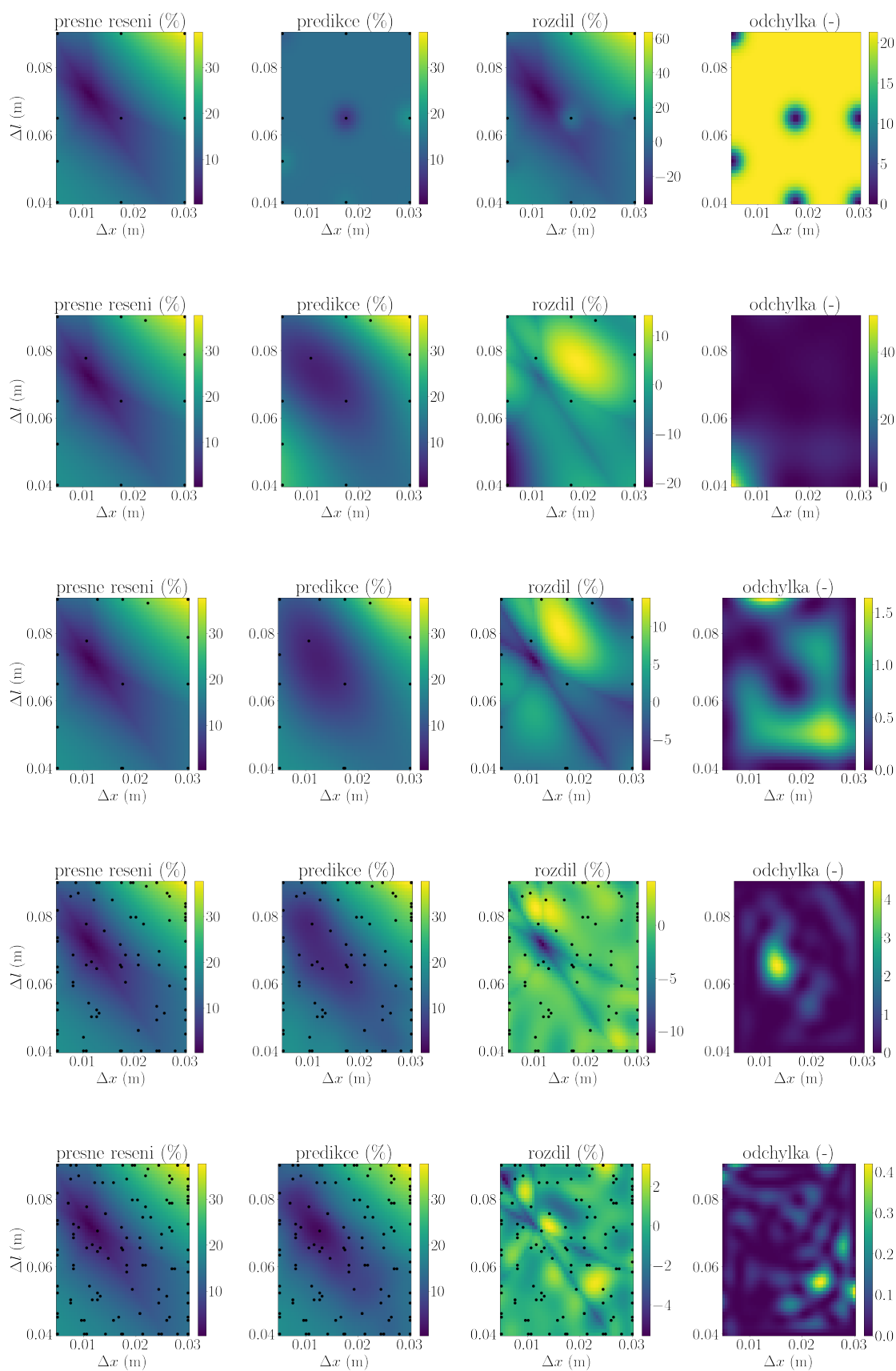
optimalizace	tvorba modelu (s)	čas optimalizace (s)	zrychlení
bez náhradního modelu	0	2584	
s náhradním modelem	309	7	8x

Optimalizace

K optimalizaci takto aproximovaného modelu byl využit opět algoritmus NSGA-II pro 150 generací o 15-ti jedincích. Na obr. 96 jsou znázorněny výsledky optimalizace plného modelu a optimalizace na aproximovaném modelu. Jak je z obrázku patrné, aproximovaný model v tomto případě velmi dobře simuluje plný model, ovšem s poměrně zásadní časovou úsporou. V tabulce 14 jsou porovnány hodnoty a čas jednotlivých optimalizačních procesů. Pro plný model trvala optimalizace 2584 s. U optimalizace na náhradním modelu (Kriging) bylo zapotřebí nejprve aproximovat navzorkovaný prostor vhodnou funkcí, což trvalo 309 s, ale samotná optimalizace na takto proložené realitě trvala jen 7 s. Celkový výpočet na zjednodušeném modelu byl tak přibližně 8-krát rychlejší.



Obr. 96: Porovnání výsledků optimalizačního procesu s a bez náhradního modelu

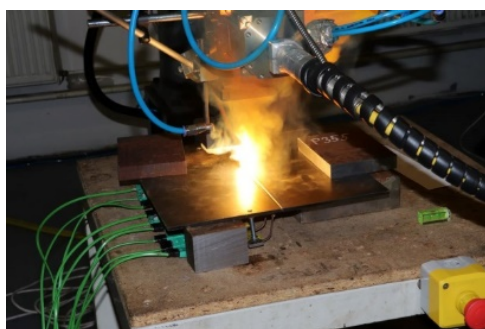


Obr. 97: Adaptivní tvorba náhradního modelu (10, 15, 20, 100 a 150 bodů)

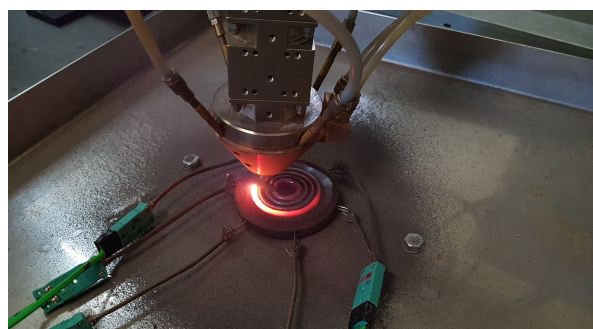
LASEROVÝ OHŘEV

Tato kapitola je zaměřena na možnosti kalibrace modelů laserového ohřevu na základě údajů získaných experimentálně. Vzhledem k tomu, že plný model je sdruženým problémem několika fyzikálních polí (teplotní pole, fázové změny, proudění), je takřka nemožné jej řešit optimalizací na jeho plné verzi. Řešením je aproximace reality modelem, který by zahrnoval všechny jevy spojené s problémy laserového ohřevu a byl uspokojivě výpočetně náročný. V této kapitole jsou představeny dva přístupy pro

- laserové svařování a
- laserové depozice kovů.



(a) Laserové svařování



(b) Laserová depozice kovů

Obr. 98: Příklad a) laserového svařování a b) laserového depozice kovů

Počítačové modelování laserového svařování a laserové depozice kovů představuje velmi složitý a náročný úkol. Laserové svařování, viz obr. 98 a), patří k technologiím používaným více než dvacet let v mnoha průmyslových a laboratorních aplikacích [102]. Při svařování se desky pomalu pohybují pod statickou laserovou hlavou. Laserový paprsek, který desky ozáří, dodává do daného místa vysoký tepelný výkon, který velmi rychle roztaví okolní materiál a vytvoří kapiláru, kterou částice odpařeného kovu unikají na jeho povrch. Vlivem kontinuálního posunu substrátu začne následně materiál znovu tuhnout – v tomto místě vznikne svar. Kvalita svaru se obvykle hodnotí podle jeho hloubky, šířky a mikrostruktury. Tyto parametry jsou funkcí relativní rychlosti mezi laserovou hlavou a svařovaným materiálem a dále výkonem laserového paprsku dodávaného na povrch ohřívaného materiálu.

Schéma procesu laserové depozice kovů, která je zobrazena na obr. 98 b), je podobné jako u svařování, ale místo vytvoření svaru je na substrát aplikována nová vrstva materiálu za účelem zlepšení povrchových vlastností daného kovu (zejména tvrdosti). Tato vrstva je tvořena speciálním kovovým práškem, který se přidává tryskou (místo prášku lze ale použít i tyčinku nebo drát). Laserový paprsek v tomto případě roztaví jak povrch materiálu, tak prášek, a

jejich směs vytvoří po ochlazení návar (housesenku). Návary jsou vytvářeny vedle sebe (a to i v několika vrstvách), dokud není požadovaná část povrchu zcela pokryta.

Hlavním problémem aproximace reality u problémů elektrotepelného ohřevu je vytvořit model, který by zahrnoval vliv ohřevu i na okolí svaru a zároveň nebyl extrémně výpočetně náročný. Je potřeba najít uspokojivou alternativu pro aproximaci fyzikální reality, která by byla schopna respektovat výše uvedené jevy. V první fázi je nutné zohlednit fázové změny a proudění, aby model mohl být řešen jen pro teplotní pole, díky čemuž lze snížit mnohonásobně čas výpočtu.

V prvním kroku bylo nutné porovnat tepelně ovlivněné části v okolí svaru připravené ze skutečného materiálu se stejnou oblastí modelovanou pomocí materiálu dostupného v knihovně v programu COMSOL Multiphysics. Experimenty byly provedeny s ocelí AM-355, jejíž charakteristika je taktéž dostupná v programu COMSOL. Bohužel však nebyly známy přesné teplotní závislosti měrného tepla ρc_p a tepelné vodivosti λ tohoto materiálu, které se mohou lišit díky vlivu neznámých příměsí. Cílem je tedy určit tyto charakteristiky co možná nej přesněji pro daný problém na základě dat získaných z experimentu.

Příspěvky na problematiku kalibrace modelů laserového svařování byly prezentovány na mezinárodních konferencích [MK.5], [MK.6] a [MK.8] s následným publikováním v impaktovaných časopisech *Compel-the International journal for computation and mathematics in electrical and electronic engineering* ([IF.1]), *IEEE Transactions on magnetics* ([IF.2]) a *Journal of Computational and Applied Mathematics* ([IF.3]).

2.1 MATEMATICKÝ MODEL

Problémy laserového svařování a depozice jsou sdruženým problémem, kde je zapotřebí vzít v úvahu nejen vliv teplotního pole, ale i proudění, keyhole efekt, případně u depozice vliv nárůstu objemu materiálu.

2.1.1 Teplotní pole produkované laserovým paprskem

Rovnice pro přenos tepla lze zapsat ve tvaru

$$\operatorname{div} (\lambda \operatorname{grad} T) = \rho c_p \left(\frac{\partial T}{\partial t} + \mathbf{v} \cdot \operatorname{grad} T \right). \quad (28)$$

kde T představuje teplotu, \mathbf{v} je vektor rychlosti pohybu laserového paprsku a t je čas. Příslušnou okrajovou podmínku lze zapsat ve tvaru

$$-\frac{\partial T}{\partial n} = -q_{\text{in}} + \alpha (T - T_0) + \sigma C (T^4 - T_r^4). \quad (29)$$

Zde T je teplota povrchu desky, α označuje součinitel přenosu tepla konvekcí, T_0 udává teplotu okolního média (vzduchu), $\sigma = 5,67 \cdot 10^{-8} \text{ W} \cdot \text{m}^{-2} \cdot \text{K}^{-4}$ je Stefan–Boltzmannova konstanta, C představuje emisivitu vyhřívané části, T_r je teplota povrchu, z něhož se tepelné záření odráží, a q_{in} odpovídá tepelnému toku dodanému laserovým paprskem. Předpokládá se, že tento tok v ozářené oblasti podléhá normálnímu rozdělení.

2.1.2 *Keyhole efekt*

Tavení materiálu na ozařovaném místě vede k vytvoření miniaturní lázně roztaveného kovu charakterizované přítomností kapilárního a turbulentního proudění. Proudění v lázni je poháněno gravitačními, vztlakovými a setrvačnými silami. Pole jeho rychlostí ovlivňuje rozložení prostupu tepla a lokální teplotní pole jak v místě roztaveného materiálu, tak i v jeho okolí. Poloha hranice mezi tuhou a tekutou fází se přitom spojitě mění. Některé částice kovu na nejteplejších místech v roztaveném materiálu se vypařují a během tohoto procesu se v důsledku jejich tlaku v roztaveném materiálu vytvoří kapilára, kterou unikají do okolí. Nad ozařeným místem tak vznikne obláček plazmatu, což vede ke snížení výkonu laserového paprsku v důsledku odrazu a absorpce světelných vln. Všechny tyto efekty jsou charakterizovány nelineární interakcí všech přítomných polí, které je třeba řešit v silně sdružené formulaci [103], [104].

2.1.3 *Nárůst objemu materiálu*

Jedním z dalších faktorů, který je důležité zohlednit v případě laserové depozice kovu, je spojitý nárůst objemu materiálu. Hmotnost zahřátého substrátu se neustále zvyšuje díky dodávanému kovovému prášku. To také znamená neustále se měnící geometrii, která vyžaduje přesíťování modelu v každém časovém kroku [103]. Tento postup však zabere velké množství času a je extrémně drahý. Proto lze využít alternativní algoritmus založený na deformaci geometrie [105], [106]. Nejprve je definována funkce, která poté, co materiál překročí svůj bod tavení, umožňuje změnu hranice mezi substrátem, ke kterému je prášek přidán, a vzduchem. Jeho koeficienty je pak nutné korelovat s ohledem na experimentální data a diskretizaci hranice oblasti.

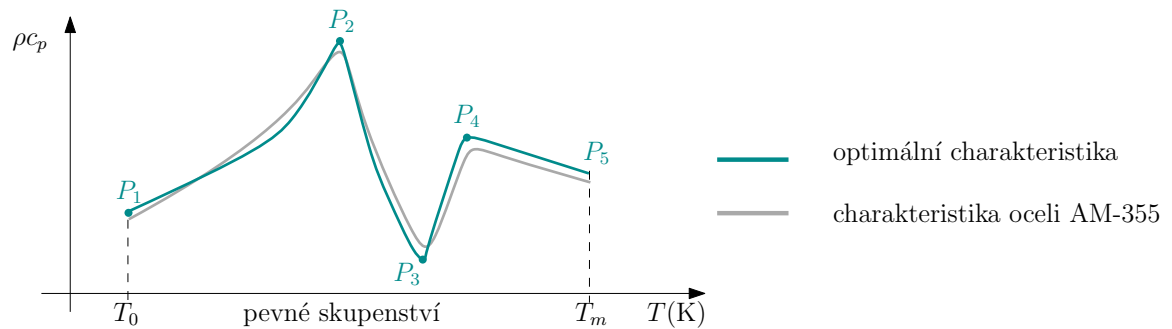
2.2 POČÍTAČOVÝ MODEL

Bohužel, plný 3D model popsáný v předchozí části je stále prakticky neřešitelný. Ale poněvadž hlavním cílem je ukázat techniku pro nalezení ekvivalentních teplotně závislých materiálových charakteristik, které poskytují výsledky dobře souhlasící s experimenty, lze použít silně zjednodušený model náhradní. Také geometrii uspořádání lze částečně zjednodušit použitím menšího vzorku. Model účinků laserového paprsku je realizován přímo na povrchu substrátu. Tato volba na jedné straně znemožňuje srovnání výsledků výpočtů s experimentálními daty, na druhé straně však dovoluje značnou úsporu času potřebného pro testovací a optimalizační výpočty. Poté není problém navrženou metodiku implementovat zpět do plného modelu a následně správnost porovnat s daty z měření.

Pro řešení je použit substrát pro ocel AM-355 s dobře známými nelineárními materiálovými charakteristikami pro teplotní rozsah 273–1489 K. Pro určení cílové funkce pro optimalizaci byly použity výsledky založené na výpočtech s charakteristikou pro měrnou tepelnou kapacitu $\rho c_p(T)$, které jsou aproximovány čtyřmi body s neznámými souřadnicemi spojenými křivkou, viz obr. 99.

Cílem metodiky je vyvinout algoritmus pro rychlé nahrazení charakteristiky nebo rekonstrukci její chybějící části. Ukáže se, že taková náhrada je velmi užitečná pro praktické aplikace. Není důležité najít dokonalou shodu se skutečnou charakteristikou, ale aproximovat ji jednodušší charakteristikou, která je přijatelná a poskytuje dobrou shodu v teplotním

modelu s experimentem nebo s referenčním řešením.

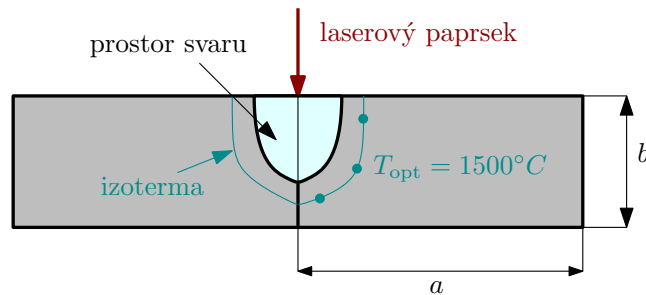


Obr. 99: Princip optimalizace funkce aproximující měrné teplo

Počítačové modelování bylo realizováno v programu COMSOL Multiphysics 5.6.

2.2.1 Laserové svařování

Předpokládejme dvě paralelní desky vyrobené z neznámé oceli (princiální schéma viz obr. 100) s neznámou charakteristikou pro měrné teplo ρc_p . Rozměry desek uvažujme $a = 0,15\text{m}$, $b = 0,005\text{m}$. Zahřívání je realizováno laserovým paprskem o výkonu 3 kW. Modrá čára na obrázku znázorňuje izotermu s teplotou $T = 1500^\circ\text{C}$, která byla experimentálně změřena na výbrusu kolmého řezu svarem ve znázorněných bodech. V těchto bodech bude zároveň odečítána teplota během simulačního procesu. Principiální schéma je znázorněno na obr. 100.

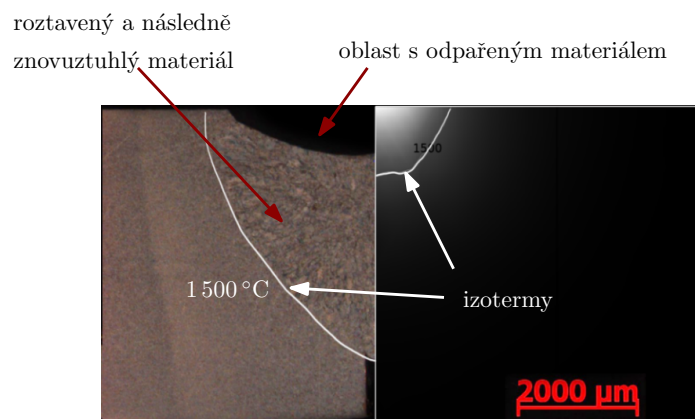


Obr. 100: Principiální schéma laserového svařování

Pro výchozí hodnoty bude použita charakteristika pro ocel AM-355. Na obrázku 101 je znázorněno porovnání řešení měřeného profilu s namodelovanou deskou, kde je model řešen jen jako teplotní problém se zanedbáním fázových změn a proudění. Vlevo je zobrazen profil svaru z experimentu. Je zde velmi dobře patrný tzv. *keyhole* efekt, který je doprovázen tokem roztaveného kovu v místě ozářeném laserovým paprskem a následně jeho částečné odpaření.

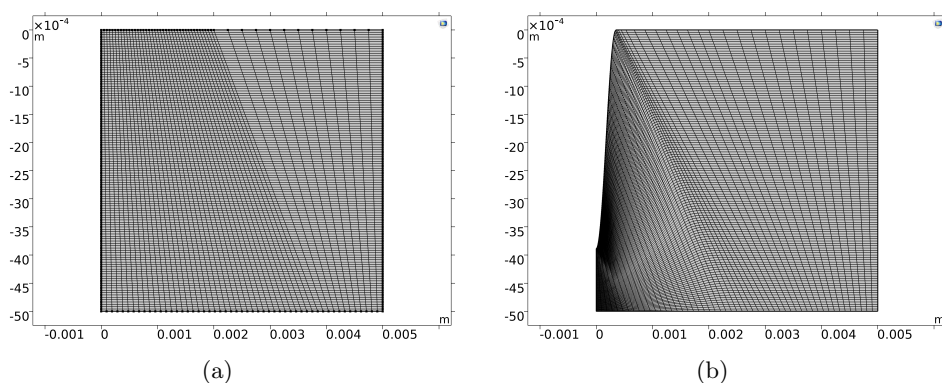
Je zde patrná i změna skupenství části materiálu zasaženého teplem z pevného na kapalné a naopak při chladnutí opět z kapalného na pevné skupenství. Tvar takto zasažené oblasti je kopírován izotermou s hodnotou $T_{\text{opt}} = 1500^\circ\text{C}$. V pravé části obrázku jsou znázorněny výsledky simulace laserového svařování v programu COMSOL. Pro lepší porovnání je zde

znázorněna izoterma opět pro hodnotu $T_{\text{opt}} = 1500\text{ }^{\circ}\text{C}$.



Obr. 101: Profil svaru před optimalizací - naměřené hodnoty (vlevo) a modelované hodnoty (vpravo)

V prvním kroku byl zanedbán vliv proudění a změny materiálu a simulace byla provedena jen pro teplotní pole. Během procesu dochází k deformaci a změně materiálu, proto je nutné vyřešit dynamickou deformaci oceli v místě svaru. Pokud nezohledníme deformaci geometrie, bude se teplo rovnoměrně šířit materiálem, viz obr. 101, a tvar izotermy v modelu nebude moci uspokojivě simulovat realitu.



Obr. 102: Výpočetní síť profilu a) se zanedbáním „keyhole“ efektu a b) s vlivem „keyhole“ efektu

Tento problém lze vyřešit pomocí vhodného popisu pro deformaci sítě během optimalizačního procesu.

$$f(x, T) = -Y \cos\left(\frac{x}{X\pi}\right)^2 \cdot F(T), \quad (30)$$

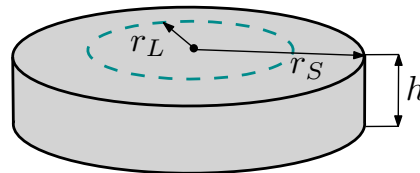
kde $Y = 4,2 \cdot 10^{-3}$ m a $X = 0,7 \cdot 10^{-3}$ m jsou parametry deformace sítě, které jsou nastaveny tak, aby tvarově odpovídaly deformaci při změně skupenství v prostoru svaru (viz obr. 100), a $F(T)$ je teplotně závislá funkce respektující fázovou změnu materiálu během experimentu.

Výsledné řešení deformace sítě je znázorněno na obr. 102. V levé části obrázku je vidět model s výpočetní sítí pro původní model. V pravé části je finální deformovaný tvar výpočetní sítě.

2.2.2 Laserová depozice kovů

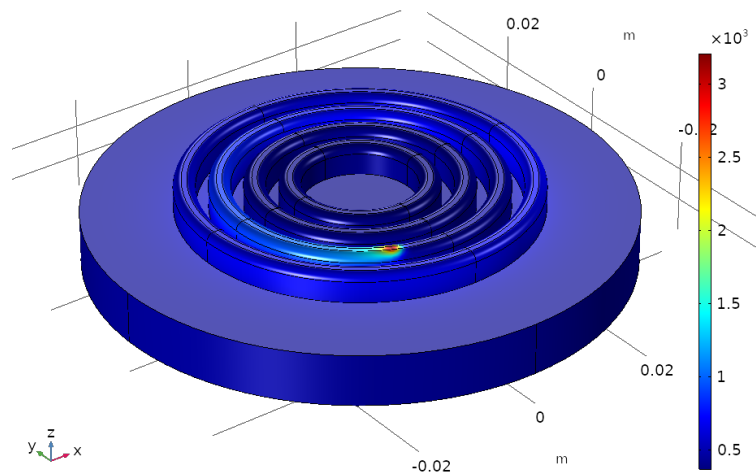
Prvním příkladem je laserové opláštění čtyř návarů na ocelovém disku. Model procesu byl experimentálně ověřen s dobrou shodou. Disk byl vyroben z oceli AM-355. Jeho poloměr je 40 mm a výška 8 mm. Návary byly provedeny ocelovým práškem značky H13. Jejich tvary byly kruhové o poloměrech 25, 20, 15 a 10 mm. Každý návar sestával z devíti vrstev. Proces navařování probíhal postupně. Na začátku byly realizovány první vrstvy každého návaru, pak druhé vrstvy atd. Výpočet provedený na plném matematickém modelu však trval více než deset hodin. Z tohoto důvodu je takto komplikovaný model nevhodný zejména v počáteční fázi, kdy je nutné vyzkoušet jeho správné nastavení a konkrétní metody. Například optimační metoda NSGA-II vyžadovala testování asi 1500 variant. Proto musel být problém zredukován, jak je uvedeno v podsekcí.

Základní princip je znázorněn na obr. 103.



Obr. 103: Principiální schéma laserové depozice

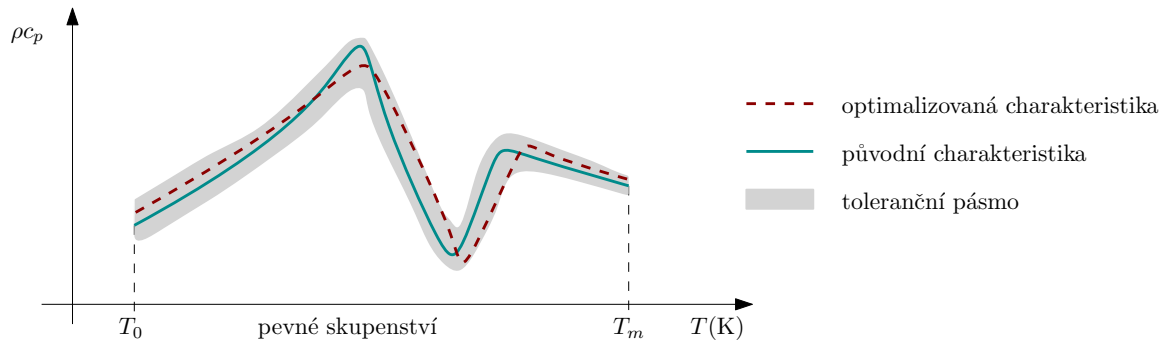
Laserový paprsek o průměru 2 mm se pohybuje po kružnici o poloměru $r_L = 10$ mm. Poloměr substrátového disku je 20 mm. Proces depozice byl sledován 50 sekund, což odpovídá neúplným 16 vrstvám při rychlosti $v_L = 20 \text{ mm} \cdot \text{s}^{-1}$. Na obr. 104 je patrný hotový počítačový model s již patrnou navařenou vrstvou.



Obr. 104: Numerický model laserové depozice

2.3 OPTIMALIZACE

Izoterma teploty tavení ozářeného materiálu získaná pomocí obou přístupů se obvykle liší, i když by teoreticky v obou případech měla být stejná. Za předpokladu, že všechny vstupní údaje, jako je geometrie a parametry laseru, jsou totožné, důvod vysoce pravděpodobně spočívá v rozdílech mezi použitými a skutečnými teplotními charakteristikami materiálu.



Obr. 105: Princip optimalizace funkce aproximující měrné teplo

Pro teplotně závislou charakteristiku měrného tepla ρc_p ocelového materiálu je předpokládaný průběh na obr.105 znázorněn zelenou čarou. Pokud tuto ocel použijeme v modelu, získáme nesprávný výsledek:

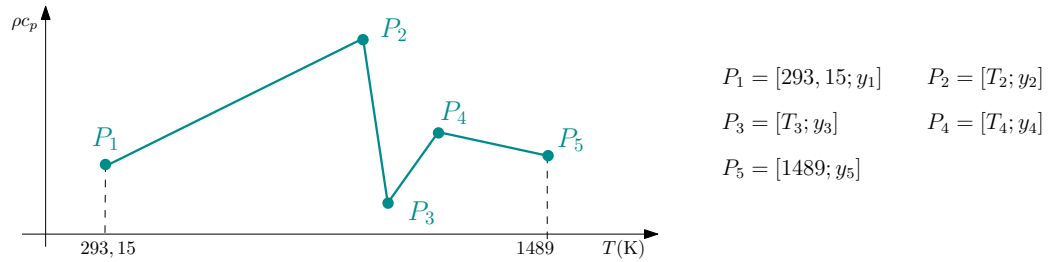
1. Uvažujme toleranční pásmo pro tuto funkci v rozmezí $\pm 20\%$. Lze předpokládat, že hledaná charakteristika pro daný ocelový materiál bude mít podobný průběh a lze ji aproximovat funkcí, jež se bude nacházet v tolerančním pásmu.
2. Tuto charakteristiku lze nahradit interpolační funkcí po částech lineární tak, aby se souřadnice jednotlivých bodů nacházely v tolerančním pásmu, viz obr. 99.
3. Vzhledem ke známé poloze izotermy s teplotou $T_{\text{opt}} = 1500^\circ\text{C}$ lze vybrat několik bodů, kde byla tato teplota změřena a na nichž bude odečítána hodnota během optimalizačního procesu T_i .
4. Nyní je možné formulovat cílovou funkci optimalizačního problému, která bude minimalizována. Hledáme tedy takové materiálové parametry, pro něž bude teplota v předem zvolených bodech odpovídat požadované hodnotě, což lze zapsat ve tvaru

$$OF = \sum_{i=1}^n |T_{\text{opt}} - T_i|. \quad (31)$$

Funkci ρc_p lze aproximovat interpolační funkcí, přičemž parametry podléhající optimalizaci jsou v této chvíli souřadnice bodů, které funkce prokládá, viz obr.106. V tabulce 15 jsou tyto souřadnice P_1, \dots, P_5 uvedeny pro původní ocel AM-355. Ve druhé části tabulky jsou tyto souřadnice nahrazeny koeficienty T_2, T_3, T_4, T_5 a y_1, y_2, y_3, y_4, y_5 , které představují parametry optimalizace a budou měněny v rozmezí $\pm 30\%$ své původní hodnoty.

2.3.1 Laserové svařování

Původní myšlenku pro optimalizaci jak materiálových parametrů tak i parametrů pro deformaci sítě bylo nutné opustit kvůli velké nestabilitě výpočtu. Vzhledem k rychlé deformaci

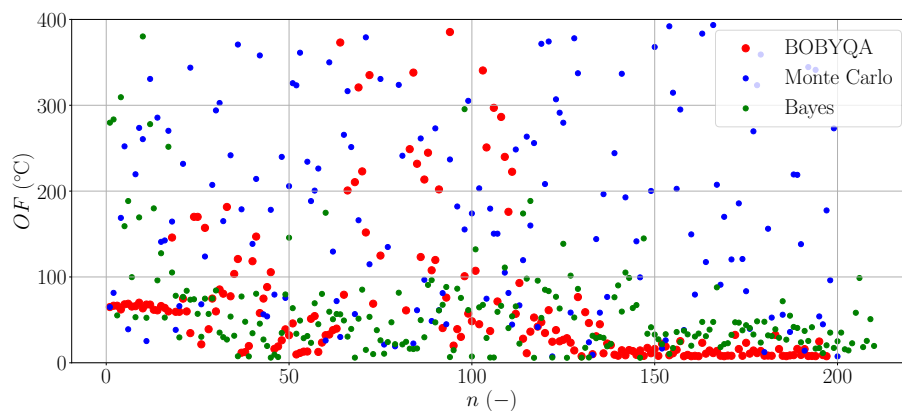


Obr. 106: Aproximační funkce materiálové charakteristiky měrného tepla ρc_p pro ocel AM-355

Tab. 15: Souřadnice bodů křivky aproximující materiálovou charakteristiku měrného tepla ρc_p oceli AM-355

ocel AM-355	P_1	P_2	P_3	P_4	P_5
T (K)	273	1171	1207	1408	1489
ρc_p (J/(m ³ · K))	$4,1 \cdot 10^6$	$5,9 \cdot 10^6$	$3,4 \cdot 10^6$	$4,7 \cdot 10^6$	$4,6 \cdot 10^6$
optimalizovaná funkce	P_1	P_2	P_3	P_4	P_5
T (K)	273	T_2	T_3	T_4	1489
ρc_p (J/(m ³ · K))	y_1	y_2	y_3	y_4	y_5

výpočetní síť je model nestabilní a i nepatrné změny mají vliv na stabilitu výpočtu. Proto byl vyřešen tvar výpočetní sítě v první fázi a následně již na této geometrii probíhala optimalizace materiálových parametrů. Požadavky na tvar a rychlost deformace byly uzpůsobeny tvaru *keyhole* v čase $t = 0,6$ s a zároveň izotermě $T = 1500$ °C, která tvar *keyhole* kopíruje.



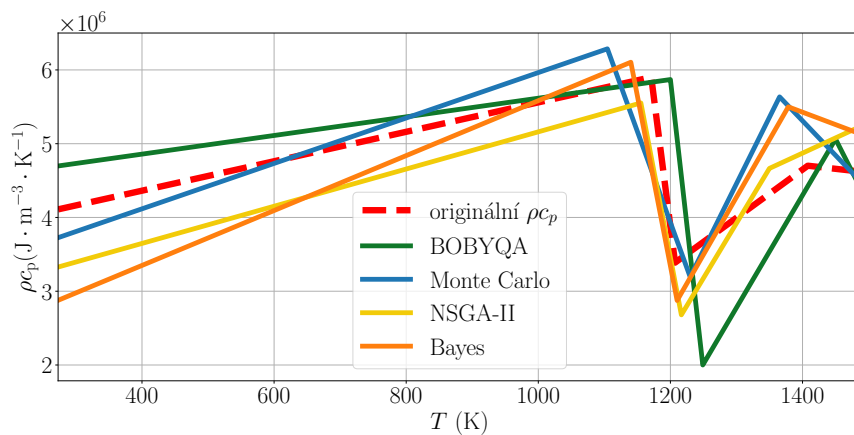
Obr. 107: Průběh optimalizačního procesu

Po nalezení optimálního tvaru geometrie bylo možné spustit optimalizační proces pro identifikaci materiálových parametrů. Optimalizace probíhala v programu COMSOL pomocí implementovaných algoritmů BOBYQA a Monte Carlo. Souběžně s tím probíhaly výpočty na Artap serveru s využitím genetického algoritmu NSGA-II a Bayesovské optimalizace. V tabulce 16 jsou porovnány algoritmy BOBYQA, Monte Carlo a Bayesovská optimalizace z hlediska počtu iterací a výsledné cílové funkce, definované v rovnici 31. Jednotlivé průběhy

jsou znázorněny na obr. 107. Jak je z tabulky patrné, ač nejmenší hodnoty dosáhl algoritmus NSGA-II s 1 510 iteracemi, tak v tomto případě byl problém vyřešen nejefektivněji pomocí algoritmu Bayesovské optimalizace, která po 210 iteracích dokázala najít cílovou funkci s poměrně uspokojivou přesností.

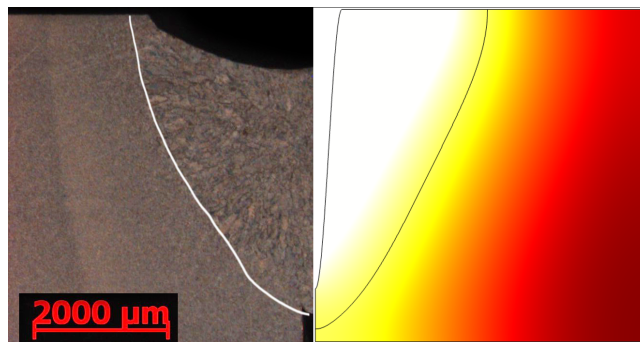
Tab. 16: Výsledné hodnoty cílové funkce po ukončení optimalizačního procesu

		BOBYQA	Monte Carlo	NSGA-II	Bayes
počet iterací	(-)	196	200	1510	210
cílová funkce	(°C)	7,7	7,4	4,2	5,8



Obr. 108: Výsledné průběhy materiálové charakteristiky pro měrné teplo ρ_{c_p}

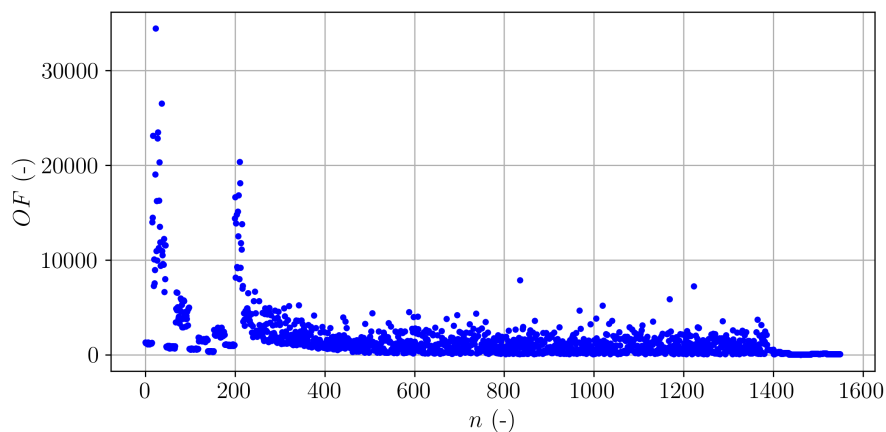
Výsledné průběhy funkce pro měrné teplo ρ_{c_p} jsou znázorněny na 108. Charakteristika původní oceli je zobrazena červenou přerušovanou čarou, dále jsou zobrazeny výsledky získané pomocí jednotlivých optimalizačních algoritmů. Na obr. 109 je porovnání profilu svaru získaného z naměřených hodnot a profilu získaného pomocí simulace s funkcí pro měrné teplo získanou pomocí algoritmu Bayesovské optimalizace.



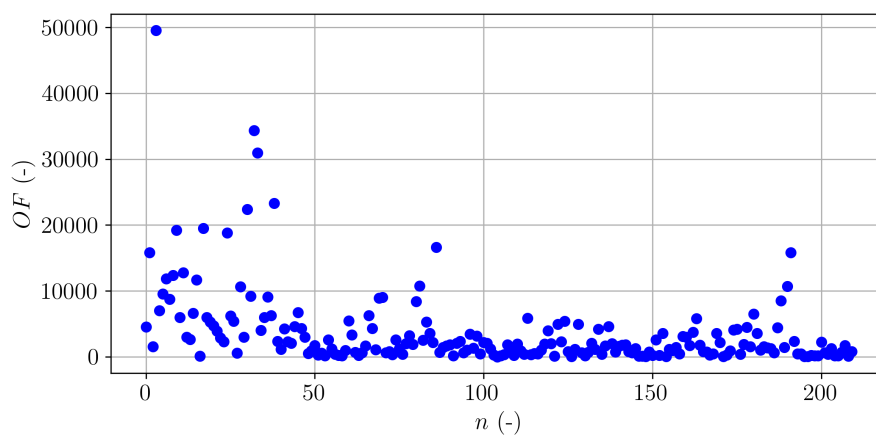
Obr. 109: Porovnání profilu svaru z experimentu (vlevo) a modelovaného profilu s optimalizovanou charakteristikou (vpravo)

2.3.2 Laserová depozice kovů

Obrázky 110, 111 a 112 ukazují konkrétní řešení optimalizace s algoritmy NSGA-II, Bayes a BOBYQA. Jednotlivé iterace technik nejsou přímo srovnatelné s ohledem na aplikaci genetického algoritmu NSGA-II řešícího jednotlivé generace.



Obr. 110: Výsledky optimalizačního procesu pomocí algoritmu NSGA-II

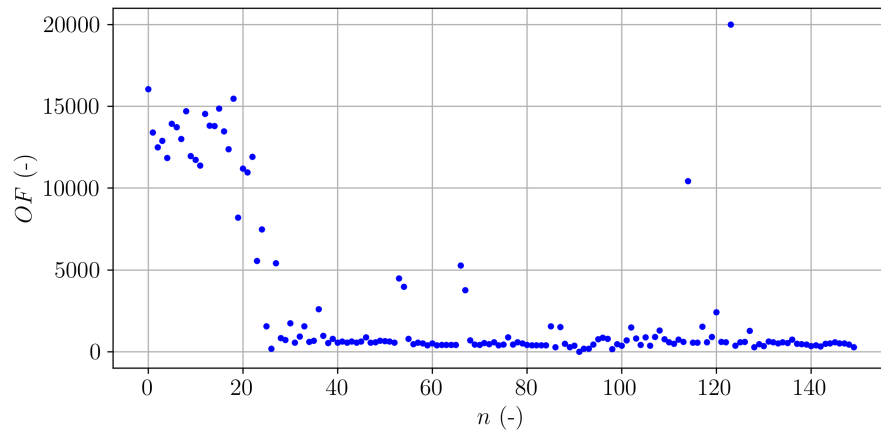


Obr. 111: Výsledky optimalizačního procesu pomocí Bayesovské optimalizace

Tabulka 17 ukazuje počty iterací nutných pro konvergenci jednotlivých optimalizačních metod.

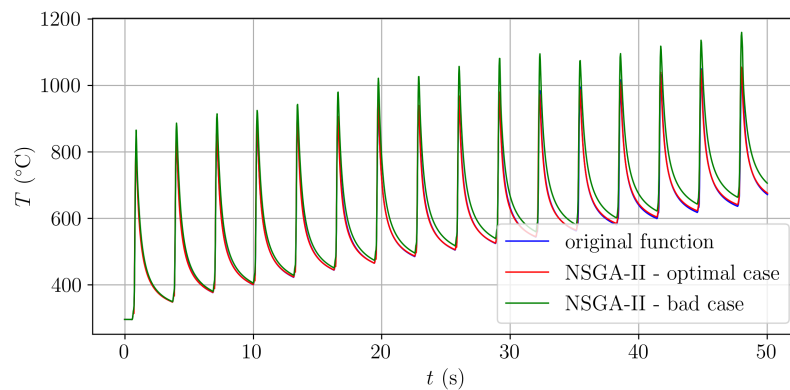
Tab. 17: Výsledné hodnoty cílové funkce po ukončení optimalizačního procesu

		BOBYQA	Bayes	NSGA-II
počet iterací	(-)	149	209	1500
cílová funkce	(°C)	7,23	36,17	58,76



Obr. 112: Výsledky optimalizačního procesu pomocí algoritmu BOBYQA

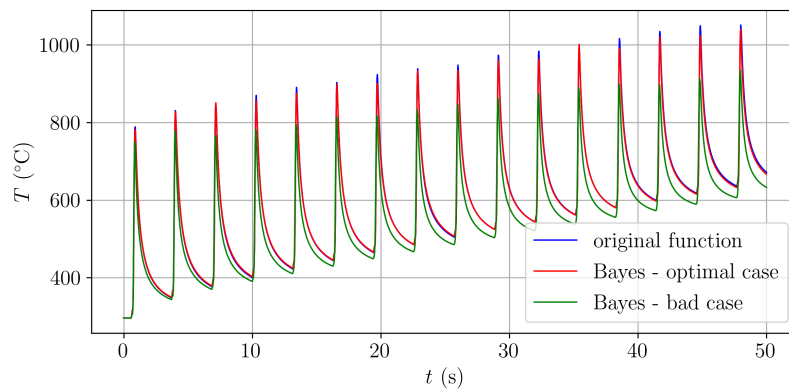
Obrázky 113, 114 a 115 porovnávají časový vývoj teplot pro charakteristiky referenčního materiálu a páry ekvivalentních řešení. Ve všech případech jsme zvolili jedno řešení jako nejnižší hodnotu účelové funkce pro danou techniku a jiné řešení jako velmi vysokou hodnotu účelové funkce (nevhodné řešení). Je zřejmé, že rozdíl mezi časovými průběhy v definovaném bodě v hloubce 1 mm pod povrchem substrátu v místě jeho ozáření laserovým paprskem je poměrně výrazný. Lepší možnost srovnání aplikovaných optimalizačních technik je znázorněna na obr. 116, jenž ukazuje jejich velmi dobrou shodu.



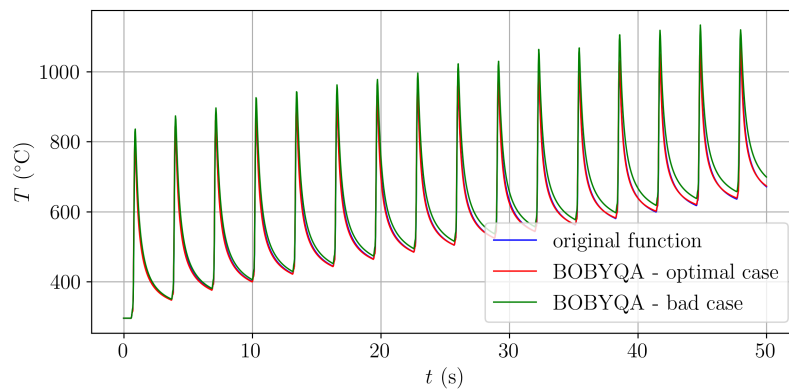
Obr. 113: Porovnání výsledků optimalizace NSGA-II s původními charakteristikami materiálu na časovém průběhu teploty ve vybraném bodě

Čas mezi dvěma sousedními vrcholy na obrázku odpovídá době jednoho oběhu kruhového návaru, viz obr. 103. Vrcholy ukazují teploty bodu, když je laserový paprsek přímo nad ním.

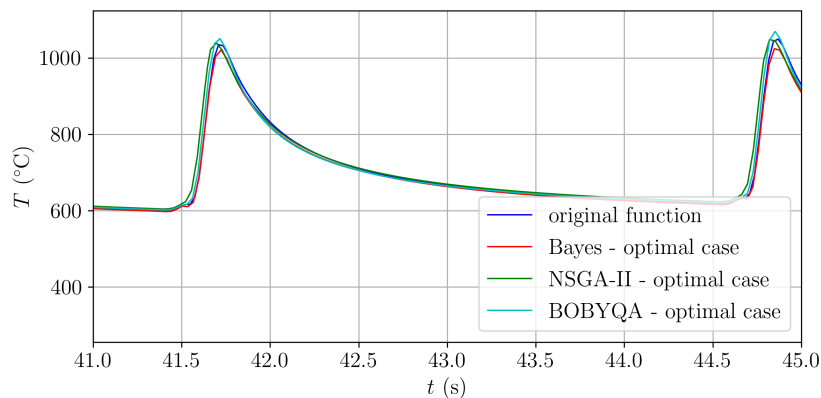
Obrázek 116 ukazuje přiblížení koncových částí všech tří předchozích obrázků. Shoda mezi teplotami získanými všemi optimalizačními metodami je mimořádně dobrá (maximální rozdíly jsou jen cca 3%).



Obr. 114: Srovnání výsledků Bayesovy optimalizace s původními charakteristikami materiálu na časovém průběhu teploty ve vybraném bodě



Obr. 115: Srovnání výsledků optimalizace metodou BOBYQA s původními charakteristikami materiálu na časovém průběhu teploty ve vybraném bodě



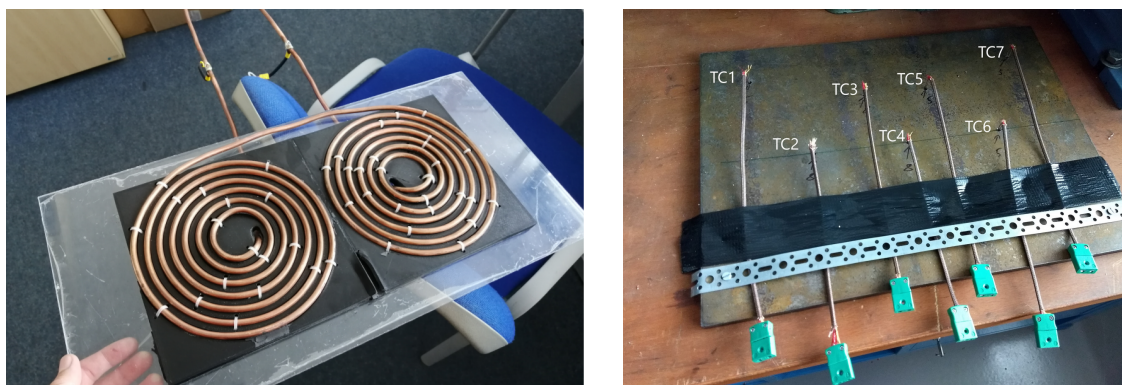
Obr. 116: Výsledky optimalizovaných variant zobrazené v detailu posledního návaru

VYPALOVÁNÍ NEVODIVÝCH VRSTEV

V této kapitole je prezentována aplikace algoritmu rekurentních neuronových sítí na predikci průběhu požadované výstupní teploty, jehož princip je patrný z ilustrativního příkladu v kapitole 2.4. Problém je řešen pro možnost vypalování tenkých vrstev na bázi nepřímého indukčního ohřevu.

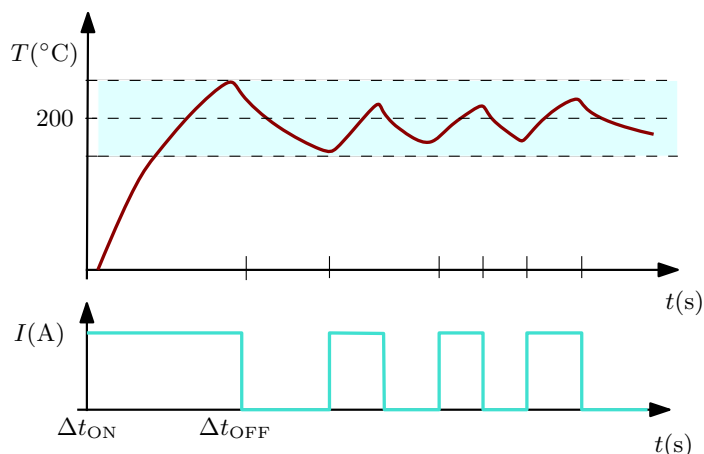
V mnoha aplikacích musí být povrchy kovových těles pokryty tenkou ochrannou vrstvou, která zabrání jejich případné degradaci vlivem okolního prostředí. Takové tenké vrstvy (umělé pryskyřice, laky nebo barvy) jsou většinou elektricky nevodivé a jejich šířka se pohybuje od desítek mikronů do desetin milimetru. Většinou jsou tyto vrstvy po nanesení zahřáté a následně ochlazené v předepsaném režimu, který zajistí jejich vypálení a vytvrzení.

Základní myšlenky pro tento přístup byly uvedeny v [107]. Teplo vzniká v povrchových vrstvách substrátu a následně je převedeno do tenké vrstvy vedením [108]. Účinnost této technologie je velmi vysoká (až 90 %), roste s magnetickou permeabilitou substrátu a klesá s jeho elektrickou vodivostí. Samotný ohřev lze velmi dobře lokalizovat. Jedinou nevýhodou je výroba induktorů, jejichž tvary se musí přizpůsobit tvaru povrchu ohřívajícího tělesa. Kmitočet napájecího proudu by měl přesahovat alespoň 10 kHz. Celé uspořádání je vyobrazeno na obr. 117.



Obr. 117: Systém pro vypékání nevodivých vrstev (induktor a systém termočlánků)

Vzhledem k tomu, že požadovaný teplotní rozsah je nutné udržovat po určitou dobu, je kontinuální ohřev substrátu poté, co se dosáhne teploty vypékání, nevhodný. Jedna z možností řízení vychází z pulsně šířkové modulace a rozdělení pulsů na interval, kdy je napájení zapnuto na čas Δt_{on} a vypnuto na čas Δt_{off} . Délka intervalů by pak mohla být řízena s ohledem na požadovanou teplotu vypékání (viz obr. 118). Další možností je spojitá regulace výkonu napájecího zdroje.



Obr. 118: Časový vývoj teploty v určitém bodě tenké vrstvy (nahore) a proudu dodávaného do induktoru (dole)

Požadavky na dobu vypalování pro udržení konkrétní teploty udává výrobce barvy a úkolem je navrhnout parametry ohřevu - amplitudu proudu, frekvenci a dobu spínání [2]. U problémů tohoto typu často není možné přímo měřit zahřátou plochu ve všech potřebných bodech, ale pouze ve vybraných oblastech. Teplotu v požadovaných místech je tedy nutné zjistit z modelu. Řešení plného modelu je však časově velmi náročné a nelze jej provést v reálném čase. Vzhledem k tomu, že nanášená ochranná vrstva je velmi tenká, její konečněprvková diskretizace vede k tomu, že počet stupňů volnosti je několik milionů a výpočet jediné varianty trvá několik hodin. Standardní on-line řízení procesu je prakticky nemožné. Jedním z možných přístupů je vytvořit náhradní model pomocí algoritmu rekurentní neuronové sítě a na základě stanovení teploty v měřicích bodech následně vyhodnocovat rozložení teploty v reálném čase. Tento přístup by mohl výrazně zpřesnit realistický odhad teploty a proces celého výpočtu podstatně urychlit.

Daná problematika je jedním z příspěvků, který byl přijat na mezinárodní konferenci COM-PUMAG 2021, viz [PČ.1].

3.1 MATEMATICKÝ MODEL

Sdružený model se skládá ze dvou obecně nelineárních parciálních diferenciálních rovnic. Rovnice popisující elektromagnetické pole má tvar

$$\operatorname{rot} \left(\frac{1}{\mu} \operatorname{rot} \mathbf{A} \right) + \gamma \frac{\partial \mathbf{A}}{\partial t} = \mathbf{J}_{\text{ext}},$$

kde \mathbf{A} je magnetický vektorový potenciál, μ je magnetická permeabilita, γ je elektrická vodivost a \mathbf{J}_{ext} je proudová hustota.

Teplotní pole je popsáno pomocí rovnice

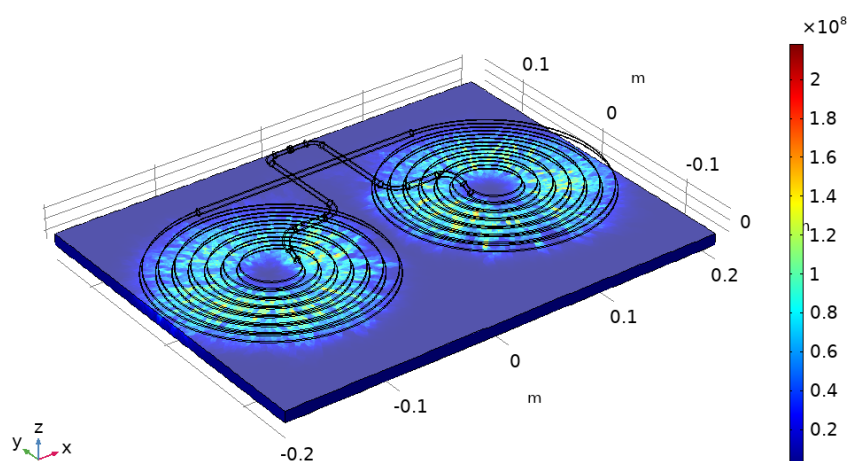
$$\operatorname{div} (\lambda \operatorname{grad} T) = \rho c_p \frac{\partial T}{\partial t} - w_j,$$

kde T je teplota, λ je tepelná vodivost, ρ je měrná hmotnost, c_p je měrné teplo a w_j označuje měrné Jouleovy ztráty.

3.2 NUMERICKÉ ŘEŠENÍ

Ohřívaná deska je vyrobena z oceli S355. Rozměry desky jsou uvažovány $400 \times 300 \times 10$ mm s tloušťkou krycí vrstvy 0,5 mm. Induktory představují dvě šroubovitě měděné trubice s vnitřním průměrem 4 mm a vnějším průměrem 6 mm. Jednotlivé induktory jsou tvořeny ze sedmi závitů a generují magnetické pole ve stejném směru. Vzdálenost mezi induktory a ohřívanou deskou je 2 mm.

Tenká vrstva laku byla vypálena na rovinném substrátu vyrobeném z nemagnetické oceli. Uspořádání obsahující desku a dvojšroubovicový induktor je znázorněno na obr. 117 vlevo. Na tomto zařízení probíhalo měření teploty v několika bodech (viz obr. 117 vpravo) a tato teplota byla následně porovnávána s daným numerickým modelem.

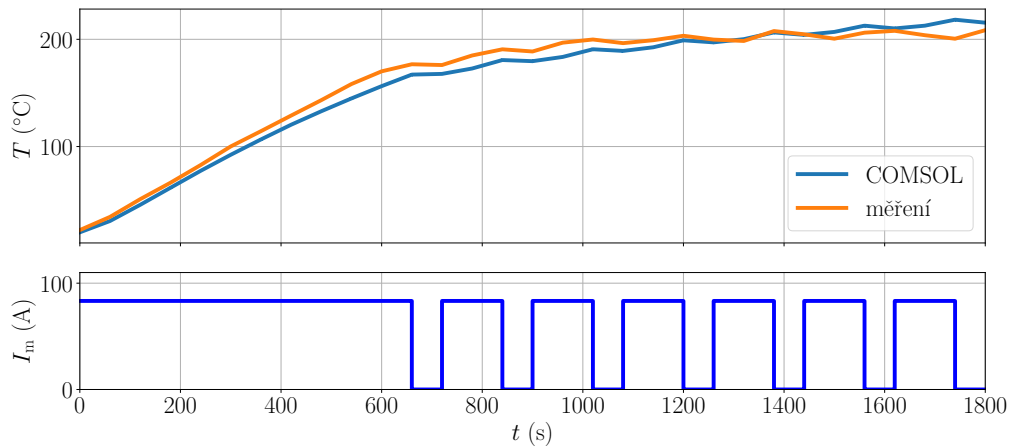


Obr. 119: Ilustrativní řešení numerického modelu se zobrazenou hustotou objemových ztrát

Model byl řešen v programu COMSOL Multiphysics. Počet stupňů volnosti je 7 000 000. Teplota je snímána v několika bodech. Rozložení senzorů odpovídá poloze termočláneků na obr. 117. Teplota je následně využita jako zpětná vazba pro regulátor v modelu založeném na RNN. Řízenou veličinou byl v tomto případě výkon zdroje. Na obr. 119 je znázorněno řešení v programu COMSOL Multiphysics s rozložením objemových tepelných ztrát. Řešení teplotního přechodného děje je zobrazeno na obr. 120, kde modrá čára znázorňuje řešení získané z počítačové simulace a modrá čára je pro hodnoty získané z měření v bodě TC2. V dolní části grafu je zobrazen spínací cyklus pro zdrojový proud. V tomto případě bylo nastaveno $I_m = 83$ A a $f = 13$ kHz.

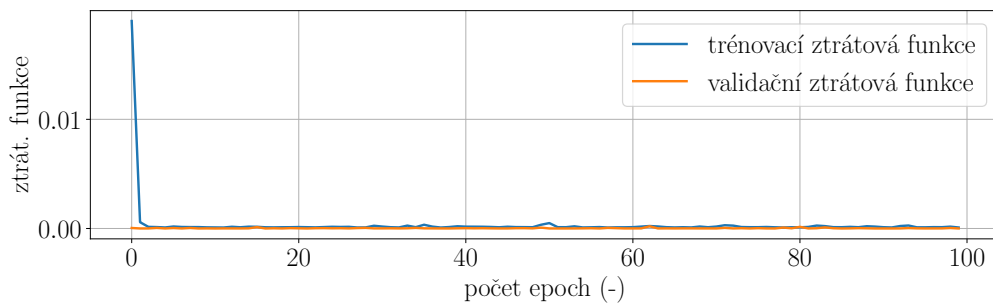
3.3 NÁHRADNÍ MODEL

Vzhledem k výše uvedené složitosti je na tomto modelu výpočetně nemožné provést optimální řízení pro průběh spínání. Jak je patrné z obr. 118, jedná se o podobný průběh sledované veličiny jako u průběhu proudu v ilustrativním příkladu 2.4. Proto budou v tomto případě



Obr. 120: Porovnání teplotního průběhu získaného z měření a z řešení numerického modelu

pro možnost predikce nepřímého ohřevu ocelové desky uvažovány algoritmy rekurentních neuronových sítí - LSTM, GRU a SimpleRNN. Princip trénování modelu bude uvažován stejně jako v předchozím případě, kdy bude na základě znalosti čtyř předchozích hodnot predikována následující hodnota. Následně bude systém o krok posunut a cyklus se bude opakovat. Na obr. 121 je zobrazena trénovací a validační ztrátová funkce v závislosti na počtu epoch. Obrázek 122 znázorňuje řešení zkušebního průběhu ohřevu v programu COMSOL Multiphysics a validační funkce provedené pomocí algoritmu LSTM. Z obrázku je patrná velmi dobrá shoda modelu s predikcí.

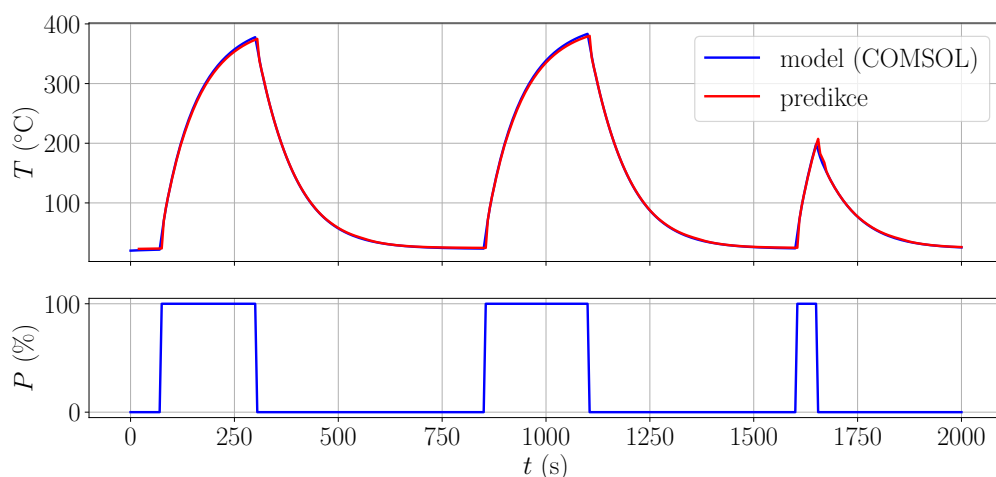


Obr. 121: Trénovací a validační ztrátová funkce

V tabulce 3.3 je zobrazena kontrola testovacích průběhů pro jednotlivé algoritmy rekurentních sítí. Porovnávána byla kvadratická odchylka σ^2 predikovaných průběhů od počítačového modelu v COMSOLu, která byla počítána vždy v jednotlivých krocích cyklu. Lze ji vyjádřit

$$\sigma^2 = \sum_0^n (T_{\text{COMSOL}} - T_{\text{predikce}})^2, \quad (32)$$

kde $n = 400$ je počet kroků v daném časovém intervalu, T_{COMSOL} je teplota získaná z programu COMSOL a T_{predikce} je predikovaná teplota. Tabulka obsahuje odchylky σ_1^2 a σ_2^2 pro testovací příklady na obr. 123 a 124.

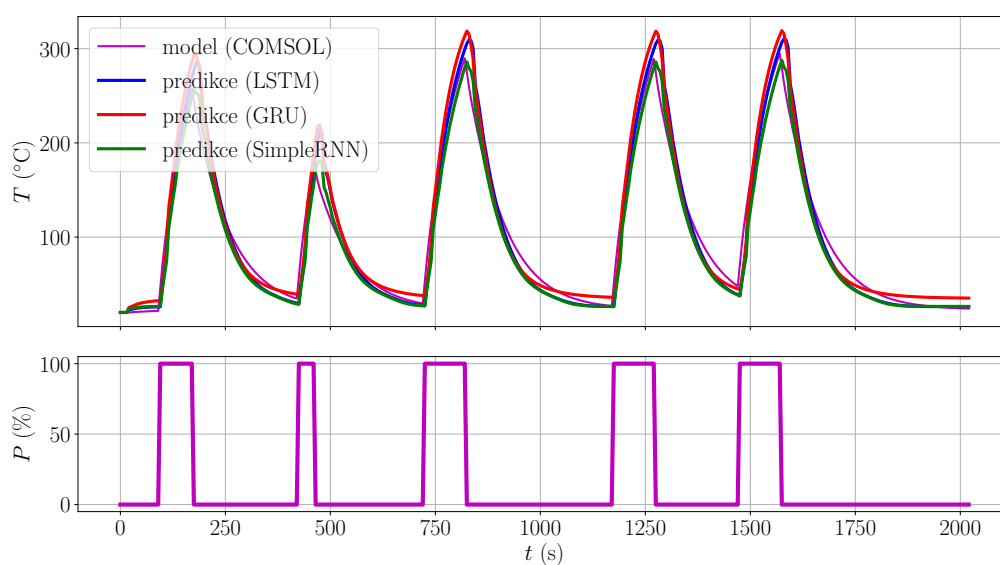


Obr. 122: Porovnání řešení z COMSOLu a validační funkce pomocí LSTM

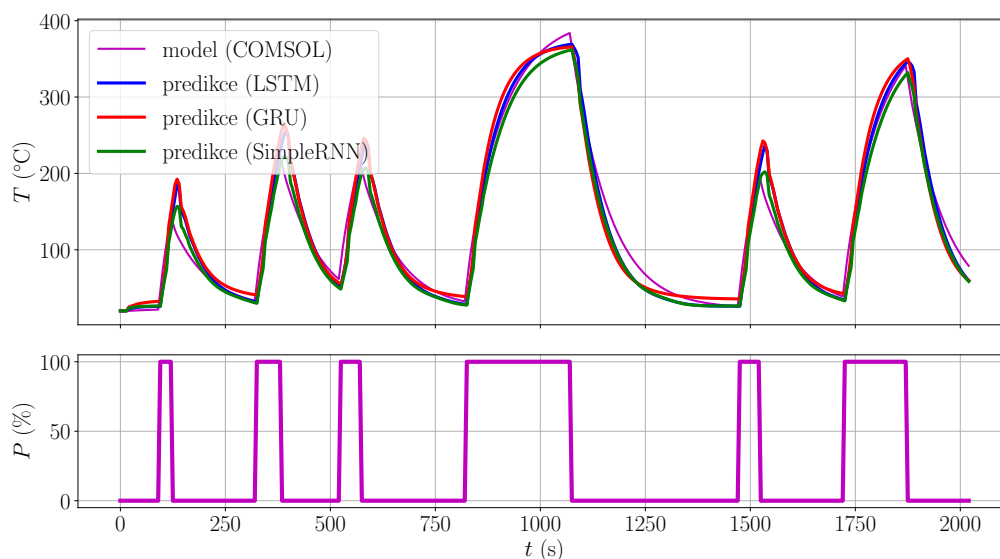
Tab. 18: Odchylky řešení pro jednotlivé algoritmy

Metoda	σ_1^2 ($^{\circ}\text{C}^2$)	σ_2^2 ($^{\circ}\text{C}^2$)
SimpleRNN	620038	678964
GRU	645480	662720
LSTM	723905	751752

Na obr. 123 a 124 je znázorněno porovnání dvou testovacích průběhů nepřímého ohřevu pro jednotlivé algoritmy v porovnání s řešením získaným z řešení numerického modelu v programu COMSOL. Je zřejmé, že všechny zvolené metody predikce velmi dobře odpovídají přesnému řešení.



Obr. 123: Porovnání řešení z COMSOLu s predikcí (testovací příklad 1)



Obr. 124: Porovnání z COMSOLu s predikcí (testovací příklad 2)

3.4 ŘÍZENÍ PROCESU POMOCÍ PID REGULÁTORU

Řízení procesu bylo provedeno pomocí PID regulátoru, který patří mezi spojité regulátory. Je složený z **p**roporcionální, **i**ntegrační a **d**erivační části (PID). Principiálně lze vyjádřit přenos regulátoru jako poměr akční veličiny $x(t)$ a regulační odchylky $e(t)$

$$f_R(t) = \frac{x(t)}{e(t)}.$$

Jako vstup regulátoru bude v tomto případě regulační odchylka teploty

$$e_T(t) = T_{\text{setpoint}} - T,$$

T_{setpoint} je cílová teplota a T je aktuální výstupní teplota. Výstupem z regulátoru (zároveň i vstupem do systému) je výkon P . Hodnota výkonu je omezena v rozsahu od 0 do 100 %. Pro realizaci algoritmu pro PID regulátor lze využít rovnici pro součet všech tří členů regulátoru (proporcionálního, integračního a derivačního členu) v následujícím tvaru

$$P(t) = P_p + P_i + P_d. \quad (33)$$

kde

$$P_p = K_p \cdot e_T(t), \quad (34)$$

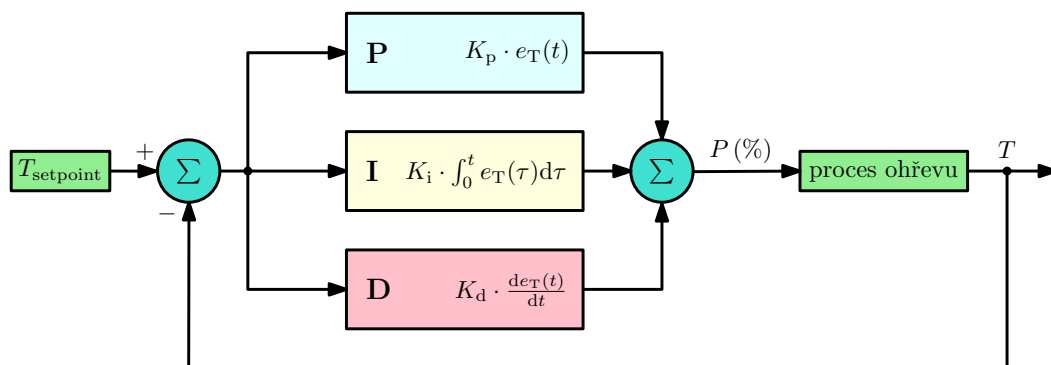
je proporcionální složka regulátoru,

$$P_i = K_i \cdot \int_0^t e_T(\tau) d\tau \quad (35)$$

je integrační složka a

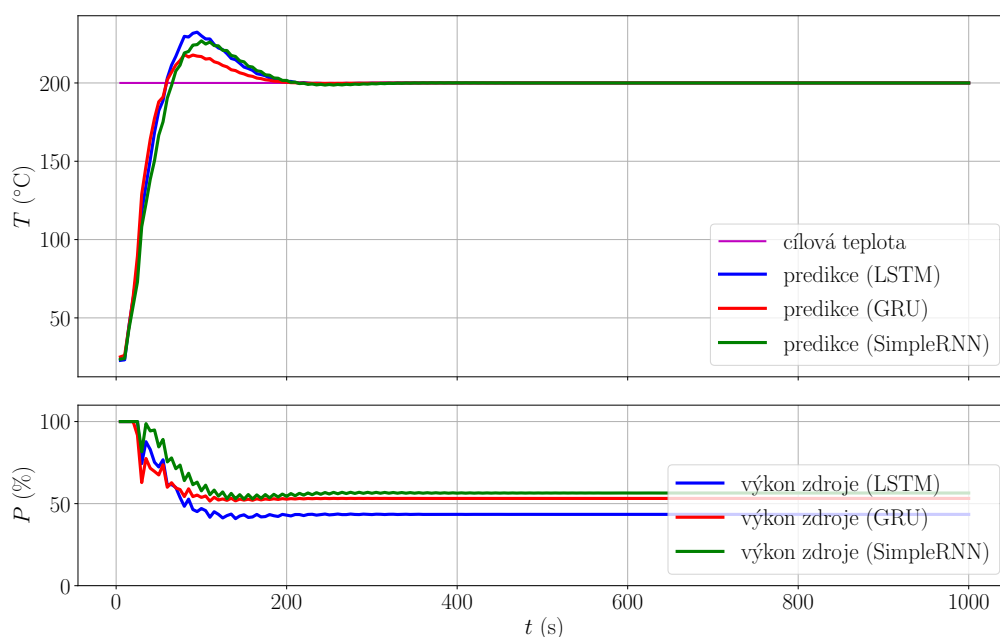
$$P_d = K_d \cdot \frac{de_T(t)}{dt} \quad (36)$$

je derivační složka. Složky P_p , P_i a P_d označují relativní příspěvky proporcionálního, integrálního a derivačního členu k řídicí akci. Konstanty $K_p = 0,54$, $K_i = 0,015$ a $K_d = 4,725$ byly určeny pomocí Zieglerovy–Nicholsovy metody. Pro přesnější určení těchto konstant by bylo možné použít některý z automatických algoritmů. Princip je patrný na obr. 125.

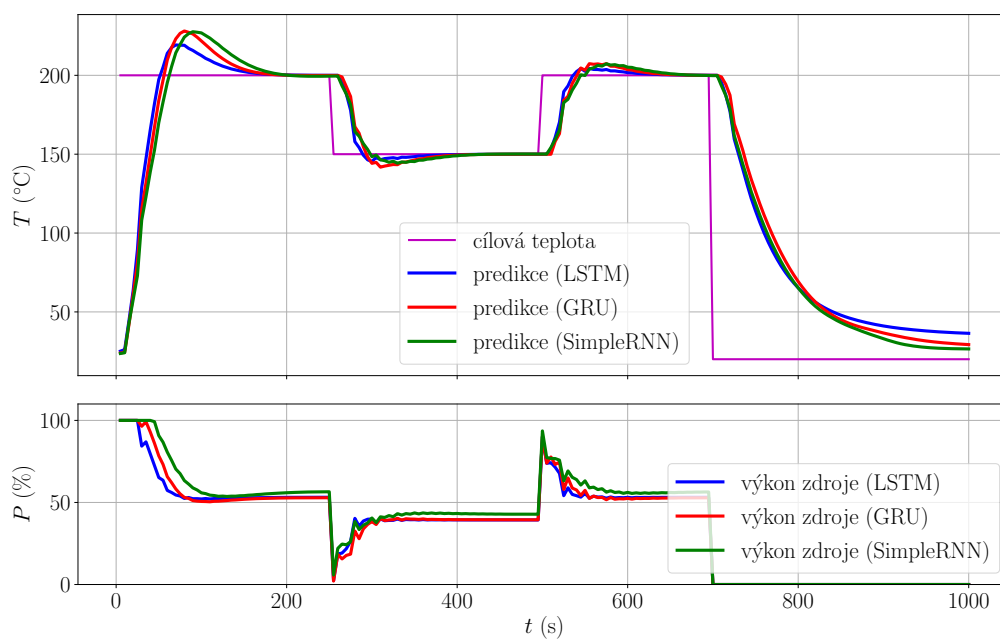


Obr. 125: Základní princip PID regulátoru

Na obr. 126 a 127 je zobrazena regulace pro dva testovací průběhy. V horní části grafu je zobrazen průběh teploty prvního testovacího příkladu, kde požadovaná hodnota teploty je $200\text{ }^\circ\text{C}$ s povolenou tolerancí $\pm 20\text{ }^\circ\text{C}$. Ve spodní části grafu je pak zobrazen průběh regulovaného výkonu. Výsledky jsou vykresleny pro cílovou hodnotu teploty fialovou čarou v horních grafech a následně pro jednotlivé algoritmy, pomocí kterých byl výpočet proveden. U druhého příkladu byl ukázán průběh teploty pro více požadovaných hodnot v různých časových intervalech. Jak je na první pohled patrné, je zde poměrně dobrá shoda pro regulovanou hodnotu predikovaného průběhu.



Obr. 126: Časová závislost teploty a výkonu s PID regulátorem (testovací příklad 1)



Obr. 127: Časová závislost teploty a výkonu s PID regulátorem (testovací příklad 2)

V tomto případě by bylo možné využití náhradního modelu jako digitálního dvojčete reálného zařízení používaného pro vytvrzování barev. Prediktivní výpočet potřebného výkonu zdroje na základě požadovaného teplotního profilu lze implementovat na mikrokontroléru s omezenými zdroji.

Část VI

ZÁVĚR A VÝHLED DO BUDOUCNA

ZÁVĚR A DALŠÍ VÝVOJ

Počítačové modelování a simulace jsou dnes využívány ve většině vědních disciplín a hrají důležitou roli při optimálním návrhu zařízení nejen v elektrotechnice. S rozmachem využití počítačových simulací a požadavků na parametry navrhovaných zařízení však stoupá i složitost plných modelů, které jsou mnohdy jen velmi těžko upočitatelné v rozumném čase, takže případná optimalizace je stěží proveditelná. K zajištění požadovaných vlastností je možné časově drahý a náročný model nahradit modelem zjednodušeným, který bude uspokojivě aproximovat realitu. Tyto modely jsou zpravidla sestaveny z dat, která jsou získána z chování původního systému. Jedná se o velmi efektivní přístup zpravidla v prvotních fázích návrhu, kdy je potřeba, aby byly splněny základní požadavky chování dané reality. V pozdějších fázích již lze uplatnit složitější počítačové modely, na kterých lze doladit požadované detailnější vlastnosti.

Předložená práce se zabývá právě optimalizací složitějších problémů s využitím zjednodušené reality v podobě vhodného náhradního modelu a je rozdělena do pěti hlavních částí.

V první části je uvedena motivace práce a vysvětleny základní pojmy, se kterými je možné se v práci setkat. Dále jsou stručně popsány řešené problémy pro vybrané oblasti elektrotechniky. Cílem této kapitoly je, aby si čtenář udělal představu, jakým způsobem je možné realizovat optimalizační proces pomocí náhradních modelů. Jsou zde vysvětleny pojmy procesu návrhu experimentu a vzorkování návrhového prostoru, na jehož základě je následně sestaven náhradní model aproximující realitu, který je pak předmětem optimalizace. Vše je aplikováno na jednoduchém ilustrativním příkladu s analytickým řešením, pomocí kterého lze ověřit použité metody. V další kapitole v první části je pak shrnut současný stav poznání, z něhož vycházejí i hlavní cíle práce.

Ve druhé části jsou již detailněji rozepsány základní pojmy, které byly stručně uvedeny v předchozí části. Jsou zde představeny metody návrhů experimentu a metody pro aproximaci návrhového prostoru. Dále je popsána možnost klasifikace optimalizačních úloh a metody pro jejich řešení, které jsou v současnosti využívány. Vše je podrobně rozebráno a předvedeno na druhém analyticky řešeném ilustrativním příkladu.

Třetí část pak seznamuje čtenáře se základními softwarovými balíky, na nichž byly realizovány počítačové simulace předložených problémů vymezených v motivaci práce. Jedná se zejména o komerční program COMSOL Multiphysics a dále programy Agros Suite a Artap, které jsou dlouhodobě vyvíjeny na pracovišti katedry elektrotechniky a počítačového modelování.

Ve čtvrté části je přístup pro optimalizace s využitím náhradních modelů (podrobně rozebrán v teoretické druhé části) aplikován na konkrétní testovací příklady s analytickým řešením, na němž je možné poměrně rychle a efektivně ověřit funkčnost uvažovaných metod. Konkrétně se jedná o dva benchmarkové příklady a řešení průběhu proudu v RL obvodu, který lze popsat pomocí obyčejné diferenciální rovnice.

V závěrečné části práce je možné již ověřené metody implementovat na vybrané složitější počítačové simulace reálných problémů z elektrotechniky. Konkrétně se jedná o optimalizaci tvaru lamelové přípojnice, problémy nejistot v materiálových charakteristikách používaných u laserového ohřevu a možnosti predikce řízení ohřevu pro vytvrzování ochranné elektricky nevodivé vrstvy nanesené na ocelovou desku. Problematika optimalizace tvaru lamelové přípojnice byla prezentována na mezinárodní konferenci *19th International Conference Computational Problems of Electrical Engineering (CPEE)* (viz. [MK.3]), možnosti identifikace parametrů pro charakteristiky materiálů využívaných u laserového ohřevu byly prezentovány na mezinárodních konferencích *HES - 19*, *COMPUMAG 2019* a *ESCO 2020* s následnou publikací ve formě článku v impaktovaných časopisech, viz [IF.1], [IF.2] a [IF.3]. Poslední uvedený příspěvek na predikci řízení ohřevu byl přijat na mezinárodní konferenci *COMPUMAG 2021* (viz. [PČ.1]), která se uskuteční v lednu 2022. Tato práce prezentuje využití rekurentních modelů neuronových sítí pro predikci veličin v průmyslových procesech. Cílem je vytvoření digitálního dvojčete u problému charakterizovaného složitou komplikovanou fyzikou.

SPLNĚNÍ CÍLŮ PRÁCE

Hlavní cíle práce byly stanoveny na základě současného stavu poznání pro řešení složitějších optimalizačních problémů. Cílem předložené dizertační práce byla ucelená formulace a následná implementace dostatečně přesných náhradních modelů, které lze při řešení optimalizace náročnějších počítačových simulací využít. Nejprve bylo zapotřebí seznámit se podrobně s metodami, které jsou v současnosti využívány pro vzorkování parametrického prostoru, které je nezbytným předpokladem pro aproximaci složité reality vhodným modelem náhradním. Posledním krokem při řešení těchto problémů je výběr vhodné optimalizační metody. Vybrané metody bylo zapotřebí úspěšně implementovat. S ohledem na výpočetní náročnost úloh prezentovaných v páté části práce byla možnost optimalizace s využitím náhradních modelů nejprve testována na několika příkladech se známým analytickým řešením. Po ověření funkčnosti vybraných postupů bylo možné uspokojivě aproximovat dané konkrétní složité aplikace vhodným náhradním modelem a nalézt požadované optimální řešení. Splnění cílů je vždy prezentováno na porovnání jednotlivých příkladů s daty získanými z experimentu nebo řešení plného počítačového modelu. Výstupy z řešení problematiky byly prezentovány na mezinárodních konferencích a v několika impaktovaných časopisech.

Předložené simulace byly provedeny buď pomocí využití komerčního softwaru COMSOL Multiphysics nebo pomocí vlastního softwaru Agros Suite s následným využitím prostředí Artap. Tyto nástroje jsou dlouhodobě vyvíjeny na katedře elektrotechniky a počítačového modelování. Prezentované řešení je v současnosti využíváno i v řadě dalších projektů a jeví se poměrně perspektivně i s ohledem na současný stav poznání a směr výzkumu v oboru. S úspěchem byly poznatky uvedené v práci využity také při výuce v navazujícím magisterském studiu.

SEZNAM PUBLIKACÍ AUTORA

ČASOPISY S IMPAKT FAKTOREM

- [IF.1] Petrášová, I., Kotlan, V., Šroubová, L., Karban, P., Doležel, I., *Calibration of laser welding model based on optimization techniques*, COMPEL - The international journal for computation and mathematics in electrical and electronic engineering, vol. 39, no. 1, pp. 43–52, 2020. DOI: 10.1108/COMPEL-05-2019-0210.
- [IF.2] Petrášová, I., Karban, P., Kotlan, V., Doležel, I., *Optimization of selected electroheat models based on calibration of material properties*, IEEE Transactions on Magnetics, vol. 56, no. 1, pp. 1–4, 2020. DOI: 10.1109/TMAG.2019.2950088.
- [IF.3] Petrášová, I., Kotlan, V., Pánek, D., Doležel, I., *Methodology of determining material parameters based on optimization techniques*, Journal of Computational and Applied Mathematics, vol. 398, paper 113671, 2021. DOI: 10.1016/j.cam.2021.113671.
- [IF.4] Petrášová, I., Karban, P., Kropík, P., Pánek, D., Doležel, I., *Optimization of selected operation characteristics of array antennas*, Journal of Computational and Applied Mathematics, vol. 399, paper 113726, 2022. DOI:10.1016/j.cam.2021.113726.
- [IF.5] Karban, P., Pánek, D., Orosz, T., Petrášová, I., Doležel, I., *FEM based robust design optimization with Agros and Ārtap*, Computers & Mathematics with Applications, vol. 81, pp. 618–633, 2021. DOI: 10.1016/j.camwa.2020.02.010.

MEZINÁRODNÍ KONFERENCE

- [MK.1] Petrášová, I., Karban, P., Tureček, O., Bouzek, S., Zuzjak, L., Schlosser, M., Kroft, R., *Numerical modeling of acoustic field of loudspeaker*, Proc. 18th International Conference Computational Problems of Electrical Engineering (CPEE). Kutná Hora, Czech Republic, pp. 1–4, 2017.
- [MK.2] Petrášová, I., Karban, P., Bouzek, S., Sýkora, M., *Numerical models of array of electroacoustic transducers in anechoic chamber*, Proc. 12th International Conference Elektro, Mikulov, Czech Republic, pp. 1–4, 2018.
- [MK.3] Petrášová, I., Karban, P., Pánek, D., Šroubová, L., *Shape and topology optimization of high power converter busbar*, Proc. 19th International Conference Computational Problems of Electrical Engineering (CPEE), Banská Štiavnica, Slovakia, pp. 1–4, 2018.
- [MK.4] Petrášová, I., Karban, P., Pánek, D., *Design of electrothermal self-sensing microactuators with embedded sensors*, Proc. 18th Biennial IEEE Conference on Electromagnetic Field Computation CEFC 2018, Hangzhou, China, 2018, pp. 1–4.
- [MK.5] Petrášová, I., Kotlan, V., Šroubová, L., Karban, P., Doležel, I., *Calibration of laser welding model based on optimization techniques*, HES-19, Padua, Italy, 2019.
- [MK.6] Petrášová, I., Karban, P., Kotlan, V., Doležel, I., *Optimization of selected electroheat models based on calibration of material properties*, COMPUMAG 2019, Paris, France, 2019.

- [MK.7] Petrášová, I., Karban, P., Pánek, D., Doležel, I., *Model of self-sensing microactuator with embedded sensor and its control*, ISEF 2019, Nancy, France, 2019.
- [MK.8] Petrášová, I., Kotlan, V., Pánek, D., Doležel, I., *Methodology of determining material parameters based on optimization techniques*, ESCO2020, Czech Republic, Plzen, 2020.
- [MK.9] Petrášová, I., Karban, P., Kropík, P., Pánek, D., Doležel, I., *Optimization of selected operation characteristics of array antennas*, ESCO2020, Plzen, Czech Republic, 2020.

PŘIJATÉ PUBLIKACE NA MEZINÁRODNÍCH KONFERENCÍCH

- [PČ.1] Petrášová, I., Karban, P., Kotlan, V., *Indirect induction baking of thin layers controlled by digital twin technique based on RNN*, COMPUMAG 2021, Cancun, Mexico, 2022.
- [PČ.2] Karban, P., Petrášová, I., Kropík, P., Stachová, L., *Novel inductive flow meter design with feedback based on measurement of chemical properties*, COMPUMAG 2021, Cancun, Mexico, 2022.

DOMÁCÍ KONFERENCE

- [DK.1] Petrášová, I., *Vyšetření elektrického pole svazkových vodičů*, Sborník konference Elektrotechnika a informatika 2016, Západočeská univerzita v Plzni, 2016, pp.189–192, ISBN: 978-80-261-0516-9.
- [DK.2] Petrášová, I., *Numerické modelování akustického pole reproduktoru*, Sborník konference Elektrotechnika a informatika 2016, Západočeská univerzita v Plzni, 2017, pp. 55–58, ISBN: 978-80-261-0712-5.
- [DK.3] Petrášová, I., *Optimalizační metody pro lokalizaci zdroje hluku*, Sborník konference Elektrotechnika a informatika 2016, Západočeská univerzita v Plzni, 2018, pp.109–112, ISBN: 978-80-261-0785-9.

REFERENCE

- [1] Alotto, P., Molfino, P., Molinari, G., *Optimisation of electromagnetic devices with uncertain parameters and tolerances in the design variables*, COMPEL—The International Journal for Computation and Mathematics in Electrical and Electronic Engineering, vol. 20, no. 3, pp. 808–812, 2001.
- [2] Kotlan, V., Doležel, I., *Baking of thin electrically non-conductive layers by indirect induction heating*, Proc. 21st International Conference on Computational Problems of Electrical Engineering (CPEE), online, 16–19 Sept. 2020, Poland, pp. 1–4.
- [3] Queipo, N.V., Haftka, R.T., Shy, W., Goel, T., Vaidyanathan, R., Tucher, P.K., *Surrogate-based analysis and optimization*, Progress in Aerospace Sciences, vol. 41, pp. 1–28, 2005.
- [4] Kotti, M., Fakhfakh, M., Tlelo-Cuautle, E., *Effect of the design space sampling on the design performances*, Proc. IEEE 9th Latin American Symposium on Circuits and Systems (LASCAS), Feb. 25–28, 2018, Puerto Vallarta, Mexico.
- [5] Giunta, A.A., Wojtkiewicz Jr, S.F., Eldred, M.S., *Overview of modern design of experiments methods for computational simulations*, Proc. 41st AIAA aerospace sciences meeting and exhibit, Jan. 6–9, 2003, Reno, Nevada, USA.
- [6] Montgomery, D.C., *Design and analysis of experiments, Eight edition*, Wiley, 2013.
- [7] Sushant S.G., Karimi, I.A., Kraft, M., *LEAPS2: Learning based evolutionary assistive paradigm for surrogate selection*, Computers and Chemical Engineering, vol. 119, pp. 352–370, 2018.
- [8] Stigler, S.M., *Georgenne’s 1815 paper on the design and analysis of polynomial regression experiment*, Historia Mathematica, vol. 1, pp. 431–447, 1974.
- [9] Walzer, R., *Galen, on medical experience*, Philosophy vol. 21, no. 78, pp. 93–94, 1946.
- [10] Rabinovitch, N.L., *Probability and statistical inference in ancient and medieval Jewish literature*, University of Toronto Press, 1973.
- [11] Sheynin, O.B., *Mathematical treatment of astronomical observations, a historical essay*, Archive for History of Exact Sciences, vol. 11, no. 2–3, pp. 97–126, 1973.
- [12] Crombie, A.C., *Avicenna on medieval scientific tradition*, In: Wickens, G.M.(Ed.), *Avicenna: Scientist and Philosopher. A Millenary Symposium*. Luzac and Co., London, 1952.
- [13] Fisher, R., *The design of experiments*, Oliver and Boyd, Edinburgh, 1935.
- [14] Fisher, R.A., *The arrangement of field experiments*, Journal of the Ministry of Agriculture no. 33, pp. 503–515, 1926.
- [15] Koehler, J., Owen, A., *9 computer experiments*, Handbook of Statistics, Elsevier, vol. 13, pp. 261–308, 1996.

- [16] Chen, V. C., Tsui, K. L., Barton, R. R., Meckesheimer, M., *A review on design, modeling and applications of computer experiments*, IIE Transactions, vol. 38, no. 4, pp. 273–291, 2006.
- [17] Garud, S. S., Karimi, A. I., Kraft, M., *Design of computer experiments: A review*, Computers and Chemical Engineering, vol. 160, pp. 71–95, 2017.
- [18] Blau, G., Lasinski, M., Orcun, S., Hsu, S.-H., Caruthers, J., Delgass, N., Venkatasubramanian, V., *High fidelity mathematical model building with experimental data: A Bayesian approach*, Computers & Chemical Engineering, vol. 32, no. 4, pp. 971–989, 2008.
- [19] Levy, S., Steinberg, D. M., *Computer experiments: a review*, ASTA Advances in Statistical Analysis, vol. 94, pp. 311–324, 2010.
- [20] Mosbach, S., Kraft, M., *Influence of experimental observations of *n*-propylbenzene kinetic parameter estimates*, Proceedings of the Combustion Institute, vol. 35, no. 1, pp. 357–365, 2015.
- [21] Mosbach, S., Braumann, A., Man, P. L., Kastner, C. A., Brownbridge, G. P., Kraft, M., *Iterative improvement of Bayesian parameter estimates for an engine model by means of experimental design*, Combustion and Flame, vol. 159, no. 3, pp. 1303–1313, 2012.
- [22] Mosbach, S., Hong, J. H., Brownbridge, G. P., Kraft, M., Gudiyella, S., Brezinsky, K., *Bayesian error propagation for a kinetic model of *n*-propylbenzene oxidation in a shock tube*, International Journal of Chemical Kinetics, vol. 46, no. 7, pp. 389–404, 2014.
- [23] Mosbach, S., Menz, W. J., Kraft, M., *Outlier analysis for a silicon nanoparticle population balance model*, Combustion and Flame, vol. 177, pp. 89–97, 2017.
- [24] Pronzato, L., Müller, W. G., *Design of computer experiments: Space filling and beyond*, Statistics and Computing, vol. 22, no. 3, pp. 681–701, 2012.
- [25] Shan, S., Wang, G. G., *Metamodeling for high-dimensional simulation-based design problems*. ASME Journal of Mechanical Design, vol. 132, no. 5, paper 051009, 2010.
- [26] Cai, X., Qiu, H., Gao, L., Yang, P., Shao, X., *An enhanced RBF-HDMR integrated with an adaptive sampling method for approximating high-dimensional problems in engineering design*, Structural and Multidisciplinary Optimization, vol. 53, no. 6, pp. 1209–1229, 2016.
- [27] Liu, H., Ong, Y.-S., Cai, J. *A survey of adaptive sampling for global metamodeling in support of simulation-based complex engineering design*, Structural and Multidisciplinary Optimization, vol. 57, pp. 393–416, 2018.
- [28] Kupresanin, A., Johannesson, G., *Comparison of sequential designs of computer experiments in high dimensions*, Technical report LLNL-TR-491692, Lawrence Livermore National Laboratory (LLNL), Livermore, 2011.
- [29] Bandler, J. W., Biernacki, R. M., Chen, S. H., Grobelny, P. A., Hemmers, R. H., *Space mapping technique for electromagnetic optimization*, IEEE Transactions on Microwave Theory and Techniques, vol. 42, no. 12, pp. 2536–2544, 1994.

- [30] Koziel, S., Bandler J. W., Madsen, K., *Quality assessment of coarse models and surrogates for space mapping optimization*, Optimization and Engineering, vol. 9, no. 4, pp. 375–391, 2008.
- [31] Bakr, M. H., Bandler, J. W., Ismail, M. A., Rayas-Sánchez, J. E., Zhang, Q. J., *Neural space mapping EM optimization of microwave structures*, Proc. IEEE MTT-S International Microwave Symposium, June 11–16, 2000, Boston, MA, USA, pp. 879–882.
- [32] Amrhein, M., Krein, P. T., *3-D magnetic equivalent circuit framework for modeling electromagnetical devices*, IEEE Transactions on Energy Conversion, vol. 24, no. 2, pp. 397–405, 2009.
- [33] Gupta, A., Bhakta, S., *An integrated surrogate model for screening of drugs against Mycobacterium tuberculosis*, Journal of Antimicrobial Chemotherapy, vol. 67, no. 6, pp. 1380–1391, 2012.
- [34] Loshchilov, I., Schoenauer, M., Sebag, M., *Intensive surrogate model exploitation in self-adaptive surrogate-assisted cma-es (saacm-es)*, Proc. 15th Annual Conference on Genetic and Evolutionary Computation (GECCO '13), July 6–10, 2013, Amsterdam, The Netherlands, pp. 439–446.
- [35] Liu, B., Zhang, Q., Gielen, G. E. G., *A Gaussian process surrogate model assisted evolutionary algorithm for medium scale expensive optimization problems*, IEEE Transactions on Evolutionary Computation, vol. 18, no. 2, pp. 180–192, 2014.
- [36] Garud, S. S., Karimi, I. A., Kraft, M., *Smart sampling algorithm for surrogate model development*, Computers & Chemical Engineering, vol. 96, pp. 130–114, 2017.
- [37] Wistuba, M., Schilling, N., Schmidt-Thieme, L., *Two-stage transfer surrogate model for automatic hyperparameter optimization*, Machine Learning and Knowledge Discovery in Databases. Lecture Notes in Computer Science, vol. 9851, Springer, Cham, pp. 199–214, 2016.
- [38] Blackman, J., Field, S. E., Scheel, M. A., Galley, Ch. R., Ott, Ch. D., Boyle, M., Kidder, L. E., Pfeiffer, H. P., Szilágyi, B., *Numerical relativity waveform surrogate model for generically precessing binary black hole mergers*, Physical Review D, vol. 96, no. 2, 2017.
- [39] Varma, V., Field, S. E., Scheel, M. A., Blackman, J., Kidder, L. E., Pfeiffer, H. P., *Surrogate model of hybridized numerical relativity binary black hole waveforms*, Physical Review D, vol. 99, no. 6, 2019.
- [40] Straus, J., Skogestad, S., *Surrogate model generation using self-optimizing variables*, Computers & Chemical Engineering, vol. 119, pp. 143–151, 2018.
- [41] Song, J., Yang, Y., Wu, J., Wu, J., Sun, X., Lin, J., *Adaptive surrogate model based multiobjective optimization for coastal aquifer management*, Journal of Hydrology, vol. 561, pp. 98–111, 2018.
- [42] McCulloch, W. S., Pitts, W., *A logical calculus of the ideas immanent in nervous activity*, The Bulletin of Mathematical Biophysics, vol. 5, no. 4, pp. 115–133, 1943.
- [43] Hebb, D. O., *The organization of behavior: A neuropsychological theory*, New Delhi: Psychology Press, 2005.

- [44] Rosenblatt, F., *The perceptron: A probabilistic model for information storage and organization in the brain*, Psychological Review, vol. 65, no. 6, pp. 386, 1958.
- [45] Babtunde, D. E., Anozie, A., Omoleye, J., *Artificial neural network and its applications in the energy sector—An overview*, International Journal of Energy Economics and Policy, vol. 10, no. 2, pp. 250–264, 2020.
- [46] Rumelhart, E. D., Hinton, E. G., Williams, J. R., *Learning internal representations by error propagation*, Technical report, University of California San Diego, La Jolla Institute for Cognitive Science, 1985.
- [47] Rumelhart, E. D., Hinton, E. G., Williams, J. R., *Learning representations by back-propagating errors*, Nature, vol. 323, no. 6088, pp. 533–536, 1986.
- [48] Hinton, E. G., Osindero, S., Teh, Y.-W., *A fast learning algorithm for deep belief nets*, Neural Computation, vol. 18, no. 7, pp. 1527–1554, 2006.
- [49] Glorot, X., Bengio, Y., *Understanding the difficulty of training deep feedforward neural networks*, In: Proc. 13th International Conference on Artificial Intelligence and Statistics, 13–15 May, 2010, Sardinia, Italy, pp. 249–256.
- [50] Langkvist, M., Karlsson, L., Loutfi, A., *A review of unsupervised feature learning and deep learning for time-series modeling*, Pattern Recognition Letters, vol. 42, pp. 11–24, 2014.
- [51] Glorot, X., Bordes, A., Bengio, Y., *Deep sparse rectifier neural networks*, In: Proc. 14th International Conference on Artificial Intelligence and Statistics, 11–13 April, 2011, Fort Lauderdale, FL, USA, pp. 315–323.
- [52] Nair, V., Hinton, G. E., *Rectified linear units improve restricted boltzmann machines*, In: Proc. 27th International Conference on Machine Learning, 21–24 June, 2010, Haifa, Israel, pp. 807–814.
- [53] Jarrett, K., Kavukcuoglu, K., Ranzato, M. A., LeCun, Y., *What is the best multi-stage architecture for object recognition?*, In: Proc. IEEE 12th International Conference on Computer Vision, 23–25 Sept., 2019, Thessaloniki, Greece, pp. 2146–2153.
- [54] Zhang, G., Patuwo, E. B., Hu, M. Y., *Forecasting with artificial neural networks: The state of the art*, International Journal of Forecasting, vol. 14, no. 1, pp. 35–62, 1998.
- [55] Makridakis, S., Sipilotis, E., Assimakopoulos, V., *The M4 competition: Results, findings, conclusion and way forward*, International Journal of Forecasting, vol. 34, no. 4, pp. 802–808, 2018.
- [56] Martinez-Cantin, R., *BayesOpt: A Bayesian optimization library for nonlinear optimization: Experimental design and bandits*. Journal of Machine Learning Research, vol. 15, pp. 3915–3919, 2014.
- [57] Cully, A., Chatzilygeroudis, K., Allocati, F., Mouret J.-B., *Limbo: A flexible high-performance library for Gaussian processes modeling and data-efficient optimization*, The Journal of Open Source Software, vol. 3, no. 26, pp. 545, 2018.
- [58] Johnson, S. G., *The NLOpt nonlinear-optimization package*, Dostupné z: <http://ab-initio.mit.edu/nlopt>.

- [59] Ouaraab, A., Ahiod, B., Yang, X. S., *Discrete cuckoo search algorithm for the travelling salesman problem*, Neural Computing and Applications, vol. 24, no. 7–8, pp. 1659–1669, 2014.
- [60] Qin, A. K., Huang, V. L., Suganthan, P. N., *Differential evolution algorithm with strategy adaptation for global numerical optimization*, IEEE Transactions on Evolutionary Computation, vol. 13, no. 2, pp. 398–417, 2009.
- [61] Thangaraj, R., Pant, M., Abraham, A., Bouvrz, P., *Particle swarm optimization: Hybridization perspectives and experimental illustrations*, Applied Mathematics and Computation, vol. 217, no. 12, pp. 5208–5226, 2011.
- [62] Karaboga, D., *An idea based on honey bee swarm for numerical optimization*, Technical report, Computer Engineering Department, Engineering Faculty, Erciyes University, Turkey, 2005.
- [63] Meng, X., Liu, Y., Gao, X., Zhang, H., *A new bio-inspired algorithm: Chicken swarm optimization*, Proc. 5th International Conference ICSI, Oct. 17–20, 2014, Hefei, China, pp. 86–94.
- [64] Mo, H., Liu, L., Geng, M., *A new magnetotactic bacteria optimization algorithm based on moment migration*, Proc. 5th International Conference ICSI, Oct. 17–20, 2014, Hefei, China, pp. 103–114.
- [65] Mo, H., Liu, L., Geng, M., *A magnetotactic bacteria algorithm based on power spectrum for optimization*, Proc. 5th International Conference ICSI, Oct. 17–20, 2014, Hefei, China, pp. 115–125.
- [66] Schmit, L. A., *Structural design by systematic synthesis*, Proc. Second Conference on Electronic Computation ASCE, Pittsburgh, USA, pp. 105–132, 1960.
- [67] Venkayya, V. B., *Structural optimization: A review and some recommendations*, International Journal for Numerical Methods in Engineering, vol. 13, no. 2, pp. 203–228, 1978.
- [68] Sobieszcanski-Sobieski, J., *Structural optimization: Challenges and opportunities*, NASA (National Aeronautics and Space Administration), 1984.
- [69] Hsu, Y.-L., *A review of structural shape optimization*, Computers in Industry, vol. 25, no. 1, pp. 3–13, 1994.
- [70] Blum, C., Roli, A., *Metaheuristics in combinatorial optimization: Overview and conceptual comparison*. ACM Computing Surveys, vol. 35, no. 3, pp. 268–308, 2003.
- [71] Hare, W., Nutini, J., Tesfamariam, S., *A survey of non-gradient optimization methods in structural engineering*, Advances in Engineering Software, vol. 59, pp. 19–28, 2013.
- [72] Schmit, L. A. B., Farshi, B., *Some approximation concepts for structural synthesis*, Proc. 14th Structures, Structural Dynamics, and Materials Conference, 20–22 March, 1973, Williamsburg, VA, USA.
- [73] Crombecq, K., Gorissen, D., et al., *A novel hybrid sequential design strategy for global surrogate modeling of computer experiments*, SIAM Journal on Scientific Computing, vol. 33, no. 4, pp. 1948–1974, 2011.

- [74] Myers, R. H., Montgomery, D. C., Anderson-Cook, C. M., *Response surface methodology: Process and product optimization using designed experiments*, John Wiley & Sons, USA, 2016.
- [75] Moen, R. D., Nolan, T. W., Provost, L. P., *Improved quality through planned experimentation, 3rd ed.*, McGraw-Hill Education, USA, 2012.
- [76] ČSN ISO 3534-3, *Slovník a značky - část: Navrhování experimentů*, Praha, Český normalizační institut, 2001.
- [77] Iman, L. R., *Latin Hypercube Sampling*, Wiley Online Library, 1999.
- [78] Svoboda, T., Hilar, M., *Pravděpodobnostní analýza metodou Latin Hypercube Sampling*, Tunel, vol. 21, no. 4, pp. 80–89, 2012.
- [79] Liu, H., Ong, Y.-S., Cai, J., *A survey of adaptive sampling for global metamodeling in support of simulation-based complex engineering design*, Structural and Multidisciplinary Optimization, vol. 57, no. 1, pp. 393–416, 2017.
- [80] Eason, J., Cremaschi, S., *Adaptive sequential sampling for surrogate model generation with artificial neural networks*, Computers & Chemical Engineering, vol. 68, pp. 220–232, 2014.
- [81] Johnston, K., Hoef Ver, J. M., Krivoruchko, K., Lucas, N., *ArcGIS9: Using ArcGIS Geostatistical Analyst*, USA, 2001.
- [82] Hochreiter, S., Schmidhuber, J., *Long short-term memory*, Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.
- [83] Mach, F., *Pokročilé metody a algoritmy pro analýzu sdružených úloh v elektromagnetismu*, Plzeň, 2015. Dizertační práce. Fakulta elektrotechnická, Západočeská univerzita v Plzni, vedoucí práce prof. Ing. Pavel Karban, Ph.D.
- [84] Antoniou, A., Lu, W.-S., *Practical optimization, algorithms and engineering applications*, Springer, USA, 2007.
- [85] Antoniou, A., Murray W., a Wright, M. H. *Practical optimization: algorithms and engineering applications*. Springer, New York, N.J., 2007.
- [86] Azizi, D., Gholami, A., *Multiobjective optimization of stator slot insulation of high-voltage generator based on coupled SNOPT-finite element method Analysis*, IEEE Electrical Insulation Magazine, vol. 29, no. 2, pp. 69–76, 2013.
- [87] Nelder, J. A., Mead, R., *A simplex method for function minimization*, The Computer Journal, vol. 7, no. 4, pp. 308–313, 1965.
- [88] Powell, M. J. D., *The BOBYQA algorithm for bound constrained optimization without derivatives*. DAMTP [online], 39 [cit. 2019-01-17], 2009. Dostupné z <http://www.damtp.cam.ac.uk>.
- [89] Karban, P., Kropík, P., Kotlan, V., Doležel, I., *Bayes approach to solving T.E.A.M. benchmark problems 22 and 25 and its comparison with other optimization techniques*, Applied Mathematics and Computation, vol. 319, pp. 681–692, 2018.

- [90] Serkan, K., *Multidimensional particle swarm optimization for machine learning and pattern recognition*, Springer-Verlag Berlin and Heidelberg GmbH & Co., 2013.
- [91] Solin, P., Andrs, D., Cervený, J., Simko, M., *PDE-independent adaptive hp-FEM based on hierarchic extension of finite element spaces*, Journal of Computational and Applied Mathematics, vol. 233, no. 12, pp. 3086–3094, 2010.
- [92] Arndt, D., Bangerth, W., Blais, B., Fehling, M., Gassmöller, R., Heister, T., Heltai, L., Köcher, U., Kronbichler, M., Maier, M., Munch, P., Pelteret, J.-P., Proell, S., Simon, K., Turcksin, B., Wells, D., Zhang, J., *The deal.II library, Version 9.3*, Journal of Numerical Mathematics, vol. 29, no. 3., pp. 1–18, 2021.
- [93] Karban, P., Mach, F., Kůs, P., Pánek, D., Doležal, I., *Numerical solution of coupled problems using code Agros2D*, Computing, vol. 95, pp. 381–408, 2013.
- [94] Solin, P., Andrs, D., Cervený, J., Simko, M., *PDE-independent adaptive hp-FEM based on hierarchic extension of finite element spaces*, Journal of Computational and Applied Mathematics, vol. 233, no. 12, pp. 3086–3094, 2010.
- [95] Tran, T.-V., Brisset, S., Brochet, P., *Combinatorial and multi-level optimizations of a safety isolating transformer*, International Journal of Applied Electromagnetics and Mechanics, vol. 26, no. 3, pp. 201–208, 2007.
- [96] Brisset, S., Brochet, P., *Analytical model for the optimal design of a brushless DC wheel motor*, COMPEL - The International Journal for Computation and Mathematics in Electrical and Electronic Engineering, vol. 24, no. 3, pp. 829–848, 2005.
- [97] K&V Elektro, *Transformátor JOC E4070 230V/230V 500VA* [online], [cit. 21.11.2021]. Dostupné: <https://www.kvelektro.cz>.
- [98] Rambabu, R., Vadakkepat, P., Chen Tan, K., Jiang, M., *A Mixture-of-Experts prediction framework for evolutionary dynamic multiobjective optimization*, IEEE Transactions on Cybernetics, vol. 50, no. 12, pp. 5099–5112, 2020.
- [99] Hwang, J., *Mixture of experts (MOE)*, in SMT: Surrogate Modeling Toolbox, 2017 [online], [cit. 29.11.2021]. DOI: https://smt.readthedocs.io/en/latest/_src_docs/applications/moe.html.
- [100] Sehovac, L., Grolinger, K., *Deep learning for load forecasting: Sequence to sequence recurrent neural networks with attention*, IEEE Access, vol. 8, pp. 36411–36426, 2020.
- [101] Lin, L., Xu, Y., Liang, X., Lai, J., *Complex background subtraction by pursuing dynamic spatio-temporal models*, IEEE Transactions on Image Processing, vol. 23, no. 7, pp. 3191–3202, 2014.
- [102] Mackwood A.P. Crafer, R. C. *Thermal modelling of laser welding and related processes: A literature review*, Optics & Laser Technology, vol. 37, no. 2, pp. 99–115, 2005.
- [103] Courtois, M., Carin, M., Le Mason, P., Gaied, S., Balabane, M., *A complete model of keyhole and melt pool dynamics to analyze instabilities and collapse during laser welding*, Journal of Laser Applications, vol. 26, no. 4, paper 042001, 2021.

- [104] Panek, D., Kotlan, V., Hamar, R., Dolezel, I., *Novel algorithm for modeling combined laser and induction welding respecting keyhole effect*, Applied Mathematics and Computation, vol. 319, pp. 254–263, 2018.
- [105] Collie, G., Higgins, R., J., Black, I., *Modelling and predicting the deformed geometry of thick-walled pipes subjected to induction bending*, Proceedings of the Institution of Mechanical Engineers, Part L: Journal of Materials Design and Applications, vol. 224, pp. 177–189, 2010.
- [106] Fernandes, F. A. O., Oliveira, D. F., Pereira, A. B., *Optimal parameters for laser welding of advanced high-strength steels used in the automotive industry*, Procedia Manufacturing, vol. 13, pp. 219–226, 2017.
- [107] Dlouhy, M., Kotlan, V., Dolezel, I., *Indirect induction hardening of thin electrically non-conductive layers*, Proc. conference ELEKTRO online, 25–28 May, 2020, Taormina, Italy, pp. 1–4.
- [108] Rudnev, V., Loveless, D., Cook, R. L., *Handbook of induction heating*, CRC Press, Boca Raton, FL, USA, 2017.

SEZNAM OBRÁZKŮ

Obrázek 1	Lamelový přívod frekvenčního měniče	4
Obrázek 2	Příklad laserového svařování	5
Obrázek 3	Řešený problém ohřevu vypalování	6
Obrázek 4	Princip optimalizace s využitím náhradního modelu	7
Obrázek 5	Princip náhradních modelů	10
Obrázek 6	Princip optimalizačního procesu	10
Obrázek 7	Principiální schéma uspořádání	11
Obrázek 8	Potenciál a intenzita elektrického pole v okolí nekonečně dlouhého vodiče	12
Obrázek 9	Použité metody návrhu experimentu, počet vzorových bodů $N = 15$	13
Obrázek 10	Aproximace navzorkovaného prostoru	14
Obrázek 11	Aproximace navzorkovaného prostoru - odchylka od původního řešení	15
Obrázek 12	Cílová funkce	15
Obrázek 13	Optimalizace rovnoměrně navzorkovaného prostoru získané z analytického řešení a Gaussova procesu	16
Obrázek 14	Optimalizace náhodně navzorkovaného prostoru získané z analytického řešení a Gaussova procesu	16
Obrázek 15	Rovnoměrné rozdělení - aplikace na ilustrativní příklad 2 (vlevo) a obecné zobrazení (vpravo)	27
Obrázek 16	Náhodné rozdělení - aplikace na ilustrativní příklad 2 (vlevo) a obecné zobrazení (vpravo)	28
Obrázek 17	Faktoriální návrh - aplikace na ilustrativní příklad 2 (vlevo) a obecné zobrazení (vpravo)	29
Obrázek 18	Mřížka Latin Square	29
Obrázek 19	LHS - aplikace na ilustrativní příklad 2 (vlevo) a obecné zobrazení (vpravo)	30
Obrázek 20	Princip Haltonovy sekvence	31
Obrázek 21	Haltonova sekvence - aplikace na ilustrativní příklad 2 (vlevo) a obecné zobrazení (vpravo)	31
Obrázek 22	Rizika pro extrémní průzkum nebo extrémní využití dat [80]	33
Obrázek 23	Proces adaptivního vzorkování - aplikace na ilustrativní příklad 2 . .	35
Obrázek 24	Lineární regrese s rovnoměrným dělením	38
Obrázek 25	Regrese polynomem druhého řádu s rovnoměrným dělením	38
Obrázek 26	Toleranční pásmo vymezené gaussovským procesem	39
Obrázek 27	Gaussov proces - aproximace rovnoměrně navzorkovaného prostoru s počtem vzorových bodů $N = 15$ (vlevo) a odchylka aproximace (vpravo)	39
Obrázek 28	Základní princip krigování	40
Obrázek 29	Struktura neuronové sítě	41
Obrázek 30	Nejčastěji používané aktivační funkce	41
Obrázek 31	MLP - aproximace rovnoměrně navzorkovaného prostoru s počtem vzorových bodů $N = 15$ (vlevo) a odchylka aproximace (vpravo) . . .	42

Obrázek 32	Princip rekurentní a klasické neuronové sítě	42
Obrázek 33	Principiální schéma LSTM algoritmu	43
Obrázek 34	Princip klasifikační metody rozhodovacích stromů	44
Obrázek 35	Rozhodovací stromy - aproximace rovnoměrně navzorkovaného prostoru s počtem vzorových bodů $N = 15$ (vlevo) a odchylka aproximace (vpravo)	45
Obrázek 36	Optimum pro funkci a) bez omezení a za b) s omezením	48
Obrázek 37	Příklad Paretovy fronty	50
Obrázek 38	Ilustrativní příklady možné prezentace řešení pro jednotlivé optimalizační problémy	51
Obrázek 39	Možnosti optimalizace geometrie zařízení [83]	52
Obrázek 40	Principiální schéma uspořádání (jeden pásový vodič)	53
Obrázek 41	Principiální schéma uspořádání (tři pásové vodiče)	54
Obrázek 42	Rozložení vektorového potenciálu s vyznačeným optimem vlevo a rozložení modulu magnetické indukce vpravo	55
Obrázek 43	Gradientní sestup – nevhodně zvolený krok	58
Obrázek 44	Gradientní sestup – adaptivní krok	58
Obrázek 45	Gradientní sestup – nevhodně zvolená počáteční poloha	58
Obrázek 46	Gradientní sestup s vazbou s adaptivním krokem (vlevo) a s velkým krokem (vpravo)	59
Obrázek 47	Nelder–Mead – určení těžiště simplexu (vlevo) a operace zrcadlení (vpravo)	60
Obrázek 48	Nelder–Mead – operace protažení (vlevo) a kontrakce (vpravo)	61
Obrázek 49	Nelder–Mead – operace zmenšení (vlevo) a zobrazení základních bodů všech operací (vpravo)	61
Obrázek 50	Princip metody Nelder–Mead	62
Obrázek 51	Vývojový diagram genetického algoritmu	64
Obrázek 52	Ukázka algoritmu NSGA-II	65
Obrázek 53	Toleranční pásmo vymezené gaussovským procesem	65
Obrázek 54	Ilustrativní cílová funkce	66
Obrázek 55	Bayesovská optimalizace	68
Obrázek 56	Ukázka Bayesovské optimalizace	69
Obrázek 57	Využívané softwary pro tvorbu počítačových simulací a řešení optimalizační problémy	71
Obrázek 58	Třívrstvá struktura Artapu a hlavní třídy aplikační vrstvy	72
Obrázek 59	Hlavní okno – postprocessor	74
Obrázek 60	Hlavní okno – postprocessor	76
Obrázek 61	Transformátor JBC 230 V/24 V 100 VA [97]	80
Obrázek 62	Schéma tepelného modelu	81
Obrázek 63	Zjednodušené obvodové schéma řešeného elektromagnetického obvodu	81
Obrázek 64	Rozměry řešeného transformátoru	82
Obrázek 65	Magnetizační charakteristika transformátorových plechů	82
Obrázek 66	Porovnání několika metod při monokriteriální optimalizaci transformátoru	85
Obrázek 67	Porovnání Paretovy fronty s řešením testovacího problému	86
Obrázek 68	Rozložení magnetické indukce v řezu transformátorem (ilustrativní řešení)	86
Obrázek 69	Elektromotorické síly, proudy a výkony pro tři fáze a výsledný moment	88

Obrázek 70	Rozměry vinutí, permanentních magnetů a vnějšího rotoru v radiálním řezu	89
Obrázek 71	Povolený rozsah vstupních parametrů po aplikaci vazby	92
Obrázek 72	Přesné řešení a predikce ve vybraných hladinách pro jedno kritérium	95
Obrázek 73	Porovnání několika metod při jednokritériální optimalizaci motoru	96
Obrázek 74	Přesné řešení a predikce ve vybraných hladinách pro dvě kritéria	97
Obrázek 75	Porovnání Paretovy fronty s řešením testovacího problému	98
Obrázek 76	Řešení numerického modelu v programu COMSOL	98
Obrázek 77	Schéma simulovaného obvodu	99
Obrázek 78	Porovnání analytického řešení a trénovací funkce pomocí LSTM	100
Obrázek 79	Princip trénování neuronové sítě s využitím algoritmu RNN	101
Obrázek 80	Princip testování natrénované sítě	102
Obrázek 81	Trénovací a validační ztrátová funkce	102
Obrázek 82	Porovnání analytického řešení a validační funkce pomocí LSTM	102
Obrázek 83	Základní princip predikce proudu	103
Obrázek 84	Porovnání analytického řešení s predikcí pomocí LSTM (testovací příklad 1)	103
Obrázek 85	Porovnání analytického řešení s predikcí pomocí LSTM (testovací příklad 2)	104
Obrázek 86	Obecné schéma lamely frekvenčního měniče s parametry a rozměry v mm	106
Obrázek 87	Základní uspořádání a matematický model	107
Obrázek 88	Rozložení proudových hustot na vyrobené lamele	107
Obrázek 89	Závislost cílových funkcí na hledaných parametrech Δx a Δl	108
Obrázek 90	Monokritériální optimalizace plného modelu	109
Obrázek 91	Vícekritériální optimalizace plného modelu	109
Obrázek 92	Statická aproximace navzorkovaného prostoru - gaussovský proces	110
Obrázek 93	Statická aproximace navzorkovaného prostoru - rozhodovací stromy	111
Obrázek 94	Statická aproximace navzorkovaného prostoru - Kriging	111
Obrázek 95	Statická aproximace navzorkovaného prostoru - KPLSK	112
Obrázek 96	Porovnání výsledků optimalizačního procesu s a bez náhradního modelu	113
Obrázek 97	Adaptivní tvorba náhradního modelu (10, 15, 20, 100 a 150 bodů)	114
Obrázek 98	Příklad a) laserového svařování a b) laserové depozice kovů	115
Obrázek 99	Princip optimalizace funkce aproximující měrné teplo	118
Obrázek 100	Principiální schéma laserového svařování	118
Obrázek 101	Profil svaru před optimalizací - naměřené hodnoty (vlevo) a modelované hodnoty (vpravo)	119
Obrázek 102	Výpočetní síť profilu a) se zanedbáním „keyhole“ efektu a b) s vlivem „keyhole“ efektu	119
Obrázek 103	Principiální schéma laserové depozice	120
Obrázek 104	Numerický model laserové depozice	120
Obrázek 105	Princip optimalizace funkce aproximující měrné teplo	121
Obrázek 106	Aproximační funkce materiálové charakteristiky měrného tepla ρc_p pro ocel AM-355	122
Obrázek 107	Průběh optimalizačního procesu	122
Obrázek 108	Výsledné průběhy materiálové charakteristiky pro měrné teplo ρc_p	123
Obrázek 109	Porovnání profilu svaru z experimentu (vlevo) a modelovaného profilu s optimalizovanou charakteristikou (vpravo)	123

Obrázek 110	Výsledky optimalizačního procesu pomocí algoritmu NSGA-II	124
Obrázek 111	Výsledky optimalizačního procesu pomocí Bayesovské optimalizace	124
Obrázek 112	Výsledky optimalizačního procesu pomocí algoritmu BOBYQA	125
Obrázek 113	Porovnání výsledků optimalizace NSGA-II s původními charakteristikami materiálu na časovém průběhu teploty ve vybraném bodě	125
Obrázek 114	Srovnání výsledků Bayesovy optimalizace s původními charakteristikami materiálu na časovém průběhu teploty ve vybraném bodě	126
Obrázek 115	Srovnání výsledků optimalizace metodou BOBYQA s původními charakteristikami materiálu na časovém průběhu teploty ve vybraném bodě	126
Obrázek 116	Výsledky optimalizovaných variant zobrazené v detailu posledního návaru	126
Obrázek 117	Systém pro vypékání nevodivých vrstev (induktor a systém termočlánků)	127
Obrázek 118	Časový vývoj teploty v určitém bodě tenké vrstvy (nahore) a proudu dodávaného do induktoru (dole)	128
Obrázek 119	Ilustrativní řešení numerického modelu se zobrazenou hustotou objemových ztrát	129
Obrázek 120	Porovnání teplotního průběhu získaného z měření a z řešení numerického modelu	130
Obrázek 121	Trénovací a validační ztrátová funkce	130
Obrázek 122	Porovnání řešení z COMSOLu a validační funkce pomocí LSTM	131
Obrázek 123	Porovnání řešení z COMSOLu s predikcí (testovací příklad 1)	131
Obrázek 124	Porovnání z COMSOLu s predikcí (testovací příklad 2)	132
Obrázek 125	Základní princip PID regulátoru	133
Obrázek 126	Časová závislost teploty a výkonu s PID regulátorem (testovací příklad 1)	133
Obrázek 127	Časová závislost teploty a výkonu s PID regulátorem (testovací příklad 2)	134

SEZNAM TABULEK

Tabulka 1	Odchylka intenzity pole náhradního modelu od analytického řešení .	14
Tabulka 2	Odchylka intenzity pole náhradního modelu od analytického řešení pro počet vzorových bodů $N = 10, 15, 20$	46
Tabulka 3	Proudy v jednotlivých vodičích	54
Tabulka 4	Dosažené minimum magnetického vektorového potenciálu pomocí různých algoritmů	67
Tabulka 5	Podporovaná fyzikální pole	75
Tabulka 6	Vstupní parametry pro optimalizaci	83
Tabulka 7	Hodnoty limitujících požadavků pro optimum úlohy (benchmark) . .	83
Tabulka 8	Optimální hodnoty vstupních parametrů pro jednotlivé algoritmy . .	84
Tabulka 9	Vstupní parametry pro optimalizaci	90
Tabulka 10	Vstupní parametry a odchylka od cílové funkce	91
Tabulka 11	Výstupní veličiny	93
Tabulka 12	Použitá neuronová síť	101
Tabulka 13	Porovnání aproximačních algoritmů pro rovnoměrně rozdělený návr- hový prostor	112
Tabulka 14	Porovnání optimalizačního procesu s náhradním modelem a bez něho	113
Tabulka 15	Souřadnice bodů křivky aproximující materiálovou charakteristiku měrného tepla ρc_p oceli AM-355	122
Tabulka 16	Výsledné hodnoty cílové funkce po ukončení optimalizačního procesu	123
Tabulka 17	Výsledné hodnoty cílové funkce po ukončení optimalizačního procesu	124
Tabulka 18	Odchylky řešení pro jednotlivé algoritmy	131