Bakalářská práce

České
vysoké
učení technické
v Praze

F3

Fakulta elektrotechnická
Katedra řídící techniky

Model vytápění budovy

Tran Anh Duc

Vedoucí: Ing. Martin Macaš, Ph.D.
Obor: Systémy a řízení
Leden 2016

Poděkování

Chtěl bych poděkovat svému vedoucímu bakalářské práce Ing. Martinu Macašovi, Ph.D. za odborné vedení, za pomoc a rady při zpracování této práce. Dále děkuji svým rodičům, kteří mi jsou oporou při studiu.

Prohlášení

Prohlašuji, že jsem předloženou práci vypracoval samostatně a že jsem uvedl veškeré použité informační zdroje v souladu s Metodickým pokynem o dodržování etických principů při přípravě vysokoškolských závěrečných prací.

V Praze, 12. ledna 2016

Abstrakt

Tato bakalářská práce se zabývá black box modelováním dynamiky vytápění budovy, tedy vytvářením modelů systému bez znalosti jejích fyzikálních veličin a to pouze pomocí naměřených nebo simulovaných dat. Teoretická část práce poskytuje dostupné informace o modelované budově, matematické základy pro tvorbu modelů a popisuje jejich rozdíly. Dále popisuje, jak při modelování systému postupovat a ukazuje různé přístupy k problému. Praktická část popisuje samotný postup při identifikaci a modelování systému. Cílem této práce je porovnání black box modelů budovy a zhodnocení jejich vhodnosti pro další použití.

Klíčová slova: identifikace systému, black box, model vytápění budovy, ARX, ARMAX, Output-Error, stavový model, NARX, Hammerstein-Wiener

Vedoucí: Ing. Martin Macaš, Ph.D.

Abstract

This bachelor thesis is about black box modeling of building heating dynamics without any prior knowledge of its physical parameters using only measured or simulated date. The theoretical part of this thesis provides available information about the building, mathematical background for models and describes their differences. The practical part shows the procedure of model identification. The main goal of this thesis is to compare buildings black box models and evaluate they are suitable for future work.

Keywords: system identification, black box, HVAC, ARX, ARMAX, Output-Error, State-Space model, NARX, Hammerstein-Wiener

Obsah

1 Úvod **1**

2 Modelování a identifikace systému **3**

2.1 Základní kroky při identifikaci systému 3

2.2 Validační metriky 4

3 Popis modelovaného systému **5**

3.1 Popis budovy 5

3.2 Popis simulátoru 6

4 Black box modely **7**

4.1 Lineární modely 7

4.1.1 Polynomiální modely 7

4.1.2 Stavový model 9

4.2 Nelineární modely 10

4.2.1 Nonlinear ARX 10

4.2.2 Hammerstein-Wiener 11

5 Experimenty a výsledky **13**

5.1 Identifikace ze simulovaných dat 13

5.1.1 Vstupní signály simulátoru . . 13

5.1.2 Simulovaná data pro identifikaci 14

5.1.3 Identifikované modely 14

5.2 Identifikace na reálných datech . 18

6 Diskuze **23**

7 Závěr **25**

A Literatura **27**

B Grafy a obrázky **29**

C Tabulky **37**

D Zadání práce **39**

Obrázky

3.1 Vnější pohled na budovu F40 5
3.2 Rozdělení zón budovy F40 6

4.1 Struktura ARX modelu 8
4.2 Struktura ARMAX modelu 8
4.3 Struktura Box-Jenkins modelu . . 9
4.4 Struktura Output-Error modelu . 9
4.5 Struktura stavového modelu 10
4.6 Struktura NARX modelu 10
4.7 Struktura Hammerstein-Wiener modelu 11

B.1 **Příklad simulovaných dat pro identifikaci** Hodinová spotřeba znázorňuje výstup a ostatní signály znamenají vstup. 29
B.2 **ARX model** Příklad porovnání skutečné a predikované spotřeby modelem $arx110s$ 30
B.3 **ARMAX model** Příklad porovnání skutečné a predikované spotřeby modelem $armax2210s$ . . . 30
B.4 **State-Space model** Příklad porovnání skutečné a predikované spotřeby modelem $statespace1s$ . . . 31
B.5 **NARX model** Příklad porovnání skutečné a predikované spotřeby modelem $narx110ss$ 31
B.6 Příklad reálných dat pro identifikaci 32
B.7 **ARX model** Příklad porovnání skutečné a predikované spotřeby modelem $arx120r$ 33
B.8 **ARMAX model** Příklad porovnání skutečné a predikované spotřeby modelem $armax1210r$ . . . 33
B.9 **State-Space model** Příklad porovnání skutečné a predikované spotřeby modelem $statespace1r$ . . 34
B.10 **NARX model** Příklad porovnání skutečné a predikované spotřeby modelem $narx220tr$ 34
B.11 **Output-Error model** Příklad porovnání skutečné a predikované spotřeby modelem $oe120r$ 35
B.12 **Hammerstein-Wiener model** Příklad porovnání skutečné a predikované spotřeby modelem $hww120pr$ 35
B.13 Simulinkový model 36

Tabulky

3.1 Vstupní signály simulátoru (fyzikálního modelu) budovy 6

5.1 Simulovaná data použitá pro identifikaci 14

5.2 **Porovnání ARX modelů různých řádů** Modely byly vytvořeny ze simulovaných dat. n_a, n_b, n_k jsou pořadě řádu výstupu, řády vstupů a dopravního zpoždění vstupů 15

5.3 **Porovnání ARMAX modelů různých řádů** Modely byly vytvořeny ze simulovaných dat. n_a, n_b, n_c, n_k jsou pořadě řády výstupu, vstupů, řád klouzavého průměru šumu a dopravního zpoždění vstupů. 16

5.4 **Porovnání Stavových modelů různých řádů** Modely byly vytvořeny ze simulovaných dat. . . . 16

5.5 **Porovnání NARX modelů** Modely byly vytvořeny ze simulovaných dat. Parametry nelineární části byly určeny algoritmy *sigmoid network nonlinear estimator*, *wavelet network nonlinear estimator* a *binary tree nonlinear estimator* [1]. Každá síť obsahovala 10 neuronů. Lineární část modelu tvoří jeden ARX v 5.2. 17

5.6 **Porovnání modelů s různou strukturou** Modely byly vytvořeny ze simulovaných dat (přípona s v označení modelu). Označení modelů jsou mým označením a vychází ze jména struktury modelu a jeho řádech. V případě NARX modelu přípona s značí, že jeho nelineární část tvoří sigmoid network. 17

5.7 Reálná data použitá pro identifikaci 18

5.8 **Porovnání ARX modelů různých řádů** Tyto modely byly vytvořeny z reálných dat. n_a, n_b, n_k jsou pořadě řády výstupu, vstupů a dopravního zpoždění vstupů 19

5.9 **Porovnání ARMAX modelů různých řádů** Tyto modely byly vytvořeny z reálných dat. n_a, n_b, n_c, n_k jsou pořadě řády výstupu, vstupů, chyby a dopravního zpoždění vstupů. 19

5.10 **Porovnání State-Space modelů různých řádů** Tyto modely byly vytvořeny z reálných dat. 20

5.11 **Porovnání NARX modelů** Tyto modely byly vytvořeny z reálných dat. Parametry nelineární části byly určeny algoritmy *sigmoid network nonlinear estimator*, *wavelet network nonlinear estimator* a *binary tree nonlinear estimator* [1]. Každá síť obsahovala 10 neuronů. Lineární část modelu tvoří vybrané ARX modely z 5.2 20

5.12 **Porovnání Output-Error modelů různých řádů** Tyto modely byly vytvořeny z reálných dat. n_b, nf jsou pořadě řády polynomů B,F 4 a n_k je řád dopravního zpoždění vstupů 21

5.13 **Porovnání Hammerstein-Wiener modelů** Tyto modely byly vytvořeny z reálných dat. Parametry nelineárních částí (NL) byly vypočteny algoritmy *piecewise-linear nonlinear estimator*, *wavelet network nonlinear estimator* a *sigmoid network nonlinear estimator* [1]. Každá síť obsahovala 10 neuronů. Lineární část modelu tvoří model $oe120r$ 5.12 21

C.1 Struktura reálných dat 37

Kapitola 1

Úvod

V Evropě je pro vytápění budov použita téměř polovina veškeré energie, proto je trendem vymýšlet a stavět budovy, které se tuto energii snaží efektivně využít[2]. Aby toto bylo možné, je potřeba znát chování vytápěcího systém budovy po působení vnějších nebo vnitřních vstupů. Provádět dlouhodobé experimenty na obydlené budově je ale nepraktické a stál by mnoho peněz a času. Z tohoto důvodu se vytváří matematické modely, na kterých můžou být pomocí výpočetní techniky prováděny různé experimenty.

Cílem této bakalářské práce je vytvořit a porovnat black box modely, které dokáží předpovídat spotřebu energie pro vytápění kancelářské budovy během následující hodiny. Tyto modely mají být vytvořeny na základě vstupních a výstupních dat a bez znalosti fyzikálních vlastností budovy. Důvodem ke zkoumání tohoto přístupu modelování dynamického systému je, že fyzikální modely sice dokáží chování přesněji predikovat, ale obsahují mnoho parametrů a jsou tím pádem výpočetně velmi náročné a drahé pro řízení[3]. Takto vytvořený model také nedokáže popsat mnoho neměřitelných faktorů jako jsou například obydlenost budovy, otevřená okna, závady v těsnění a jiné. Všechny tyto faktory není pro black box modelování potřeba znát, za to je ale potřeba mít pro vytvoření modelu dostatečné množství dat. Mimo to jsou black box modely mnohem jednodušší a mnohdy dostačující.

Zkoumanou budovou se pro mě stal kancelářský dům s označením F40, který se nachází v Casaccia Research Centre, Itálie. Pro tuto budovu byly již vytvořeny white box a gray box modely [4]. Přínosem mé práce je tedy průzkum black box metod modelování pro tuto budovu a zhodnocení jejich vhodnosti pro budoucí zkoumání.

Jako zdroj dat pro identifikaci mi byla poskytnuta reálná naměřená data během zimy 2014/2015 z budovy F40. Dále jsem obdržel simulinkový model této budovy, který byl vytvořen v prostředí HAMBASE [5], jehož výstupní data jsem také měl použít pro identifikaci.

V kapitole 2 jsem uvedl obecné metody a kroky při identifikaci systému. Dále jsem zde popsal srovnávací metriky, které byly použity pro srovnání modelů. Informace o modelovaném objektu a také původ dat, na kterých identifikace proběhla, jsem poskytnul v kapitole 3. V kapitole 4 jsem předložil přehled black box modelů, které jsem v kapitole 5 implementoval. Před samotným závěrem této práce jsem své výsledky diskutoval v kapitole 6.

Kapitola 2

Modelování a identifikace systému

Přístupy modelování systémů se rozdělují do tří směrů podle množství informací, které o daném systému máme k dispozici. V prvním přístupu známe fyzikální vazby systému a všechny jeho parametry (tzv. white box / physical model). U druhého přístupu známe fyzikální vazby, ale některé jeho parametry neznáme (tzv. gray box model / semi-physical). Poslední přístup je založen jen na naměřených datech (tzv. black box model / data-driven model).

White box model Tento přístup vyžaduje mít veškeré znalosti o systému. Je potřeba znát všechny jeho vnitřní fyzikální vazby a všechny jeho parametry. Tyto modely jsou zpravidla přesnější než ostatní modely, ale jsou velmi složité, tudíž jsou i výpočetní nároky vyšší a pro řízení je to velmi nákladné [3]. Pro validaci těchto modelů je však také zapotřebí mít naměřená data. Na základě tohoto přístupu byl vytvořen simulátor, se kterým jsem v této práci pracoval.

Gray box model Pokud známe alespoň vnitřní strukturu systému, je vhodné použít gray box přístup. O proti předchozímu přístupu neznáme všechny parametry tohoto modelu. Tyto parametry se pak snažíme identifikovat z naměřených dat. Model budovy F40 vytvořil tímto přístupem Fabio Moretti [4].

Black box model Pokud neznáme vnitřní strukturu systému a ani jeho parametry, používá se black box přístup. Přestože o modelu nic nevíme, lze jeji určit z naměřených dat a jelikož chybí znalost tohoto systému, je potřeba mít co nejvíce naměřených vstupních a výstupních dat. Výhodou tohohle modelu oproti předchozím je, že bývají jednodušší a pro řízení mohou být postačující. Podrobně jsem tyto modely popsal v kapitole 4.

2.1 Základní kroky při identifikaci systému

Proces identifikace systému se zabývá určením modelu na základě pozorování vstupních a výstupních dat. Pro stanovení modelu systému jsou zapotřebí následující :

- Vstupní a výstupní data
- Zvolené struktury modelů, které budu testovat
- Kritéria podle kterých budou nejlepší modely dané struktury vybírány

Při samotné identifikaci se pak opakují kroky, kdy se vybírá struktura modelu s jinými řády, proběhne na nich identifikace a validace a poté se vybere model podle kritérií, které jsou předem stanoveny. Tento cyklus se dá popsat následovně:

1. Návrh experimentu pro získání dat, na kterých identifikace bude probíhat
2. Prozkoumání dat, předzpracování dat například odstraněním trendů a extrémů, případně použití filtrů pro zdůraznění některých frekvenčních pásem.
3. Volba struktury modelu a jeho řádů
4. Určení parametrů modelu, které dávají nejlepší výsledky pro danou strukturu daného řádu
5. Validace modelu na jiném souboru dat podle stanovených validačních metrik 2.2. Určení, zda model postačuje našim potřebám podle našich kritérií Pokud ano, volíme tento model, pokud ne, vracíme se ke kroku 3

2.2 Validační metriky

Pro experimentální srovnání modelů byly použity následující metriky:

Mean Squared Error (MSE)

Tato metrika znázorňuje střední kvadratický rozdíl měřené a skutečné hodnoty a její jednotkou je kvadrát jednotky měřené veličiny. Matematicky je vyjádřena rovnicí 2.1. Nižší hodnota znázorňuje lepší výsledek.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(\hat{y}(i) - y(i))^2 \tag{2.1}$$

kde $y(i)$ znázorňuje skutečnou hodnotu, $\hat{y}$ predikovanou a N počet vzorků.

Mean Absolute Percentage Error (MAPE)

Oproti předchozí metrice je tato vyjádřena v procentuálně. Opět platí, že nižší hodnota znamená lepší výsledek [6]. Vypočítá se pomocí rovnice 2.2.

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{\hat{y}(i) - y(i)}{y(i)}\right| \tag{2.2}$$

.

Goodness of fit

Tato metrika se používá pro procentuální určení, jak s predikovaná hodnota blíží skutečnou. V tomto případě znamená nižší hodnota horší výsledek a vychází z rovnice 2.3

$$FIT = 100 \cdot \left(1 - \frac{\sqrt{\sum_{i=1}^{N}(\hat{y}(i) - y(i))^2}}{y_{max} - y_{min}}\right) = 100 \cdot (1 - NRMSE) \tag{2.3}$$

kde y_{min} znázorňuje minimální hodnotu, y_{max} maximální hodnotu a NRMSE je Normalized Root Mean Squared Error [7].

Kapitola 3

Popis modelovaného systému

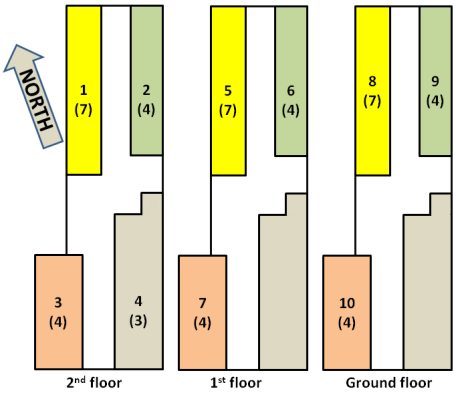
3.1 Popis budovy

Reálná budova s názvem F40 (viz 3.1) se nachází v ENEA (Casaccia Research Centre, Řím, Itálie). Tato třípodlažní budova v sobě obsahuje 41 kancelářských místností s plochami mezi 14 a 36 m^2, 2 místnosti pro zpracování elektronických dat s plochami 20 m^2, čtyry laboratoře, jednu kontrolní místnost a dvě meetingové místnosti. Vytápěcí místnost se nachází ve sklepním prostoru. Každá místnost obsahuje klimatizaci řízenou termostatem s hysterezí 1^oC a pracuje v ní jeden až dva lidé.

V zimním období je vytápění budovy prováděno plynovým kotlem. V letním období je klimatizace zajištěna třemi elektrickými kompresorovými chladícími zařízeními. V této práci se budu zabývat jen vytápěním v zimních sezónách. Cirkulaci kapaliny zajišťuje trojice odstředivých čerpadel.

Celá budova je vybavena pokročilým monitorovacím systémem, který sbírá data (klimatické podmínky, spotřeba elektrické a tepelné energie) ze senzorů.

Obrázek 3.1: Vnější pohled na budovu F40



Obrázek 3.2: Rozdělení zón budovy F40

Označení	Popis	Jednotka
T_{s1} - T_{s9}	Setpoint teploty vzduchu v zóně 1-9	$[°C]$
T_{sw}	Setpoint teploty vody ve vytápění	$[°C]$
Q_{sw}	Setpoint toku vody ve vytápění	$[kg/h]$
W_1	Intenzita nepřímého sluneční záření	$[W/m^2]$
W_2	Venkovní teplota vzduchu	$[°C]$
W_3	Intenzita přímého sluneční záření	$[W/m^2]$
W_4	Zakrytí mraky	(1...8)
W_5	Venkovní relativní vlhkost	[%]
W_6	Rychlost větru	$[ms^{-1}]$
W_7	Směr větru	[° od severu]

Tabulka 3.1: Vstupní signály simulátoru (fyzikálního modelu) budovy

3.2 Popis simulátoru

Jako druhý zdroj dat mi byl poskytnut simulinkový model, který byl vytvořen pomocí whitebox metod na základě reálných dat naměřených na budově popsané výše. Tento model byl vytvořen v prostředí HAMBASE [5] v Universitá Politecnica delle Marche, Itálie. Schéma modelu se nachází v přílohách B.13. Vstupem do tohoto modelu jsou setpointy teplot v zónách, setpoint teploty topící vody a naměřené veličiny počasí. Tyto vstupy jsou popsané v tabulce 3.1. Naměřené veličiny počasí jsem obdržel spolu s modelem. Tato data byla naměřena na reálné budově F40, tudíž jsem nad vytváření těchto signálů nemusel uvažovat. Z výstupních dat simulátoru mě pro identifikaci zajímala hlavně hodinová spotřeba energie a teploty vzduchu v zónách.

Kapitola 4

Black box modely

4.1 Lineární modely

4.1.1 Polynomiální modely

Polynomiální model využívá obecný zápis přenosových funkcí pro vyjádření vztahu mezi vstupem $u(k)$, výstupem $y(k)$ a bílým šumem $e(k)$. Tento vztah je dán následujícím tvarem:

$$A(z)y(k) = \sum_{i=1}^{n_u} \frac{B_i(z)}{F_i(z)} u_i(k - n_{ki}) + \frac{C(z)}{D(z)} \varepsilon(k) \tag{4.1}$$

kde u_i vyjadřuje i-tý vstup, n_u počet vstupů, n_{ki} dopravní zpoždění i-tého vstupního signálu. $\varepsilon(k)$ znázorňuje bílý šum. Proměnné $A(z), B(z), C(z), D(z)$ a $F(z)$ jsou polynomy přenosových funkcí vyjádřených v z notaci, kde z^{-1} se využívá k vyjádřední zpožděného signálu [1]. Řády těchto polynomů dále značím jako n_a, n_b, n_c, n_d,n_f a n_k. Příklad polynomu $A(z)$ řádu $na = 2$: $A(z) = 1 + a_1 z^-1 + a_2 z^-2$.

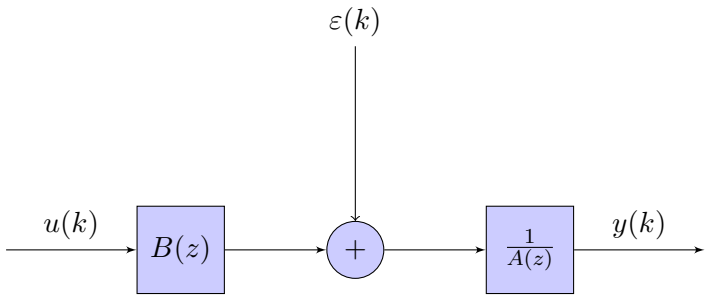
Z obecné rovnice 4.1 vychází ostatní modely a liší se tím, jaké přenosové funkce v sobě obsahují. Každý model nám dává více či méně možností popsat dynamiku systému a charakteristiky šumu. Tyto struktury modelů popisuji níže.

ARX model

ARX (autoregressive with external inputs) struktura modelu patří k nejjednodušším. Jedná se o autoregresní model s vnějšími vstupy. Tato struktura popisuje dynamiku systému a vstupní šum je touto dynamikou také ovlivněn. Struktura tohoto modelu je popsána vztahem:

$$A(z)y(k) = \sum_{i=1}^{n_u} B_i(z) u_i(k - n_{ki}) + \varepsilon(k) \tag{4.2}$$

Tento model nám nedovolí popsat chování šumu zvlášť, ale je jednoduchý a pro základní model postačuje. Struktura je znázorněna na obrázku 4.1.



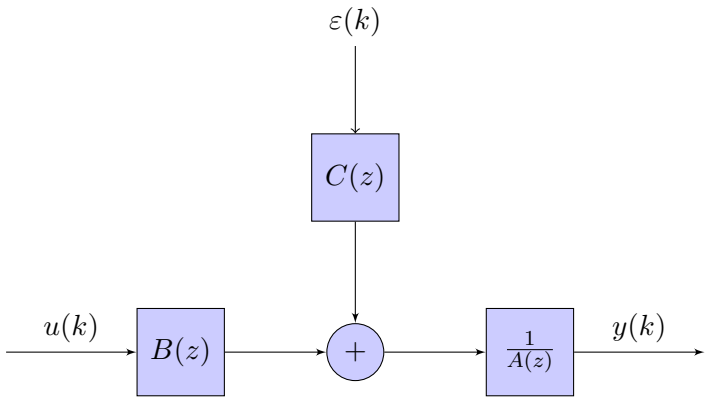
Obrázek 4.1: Struktura ARX modelu

ARMAX model

ARMAX (autoregressive moving average with external inputs) je struktura, která rozšiřuje výše zmíněnou ARX o možnost ovlivnit výstupní šum pomocí polynomu C. Tento model je vhodné použít, pokud se dominantní rušení objevuje na vstupu. Vztah

$$A(z)y(k) = \sum_{i=1}^{n_u} B_i(z)u_i(k - nk_i) + C(z)\varepsilon(k) \tag{4.3}$$

popisuje strukturu tohoto modelu a vizuálně si jí lze představit podle obrázku 4.2.



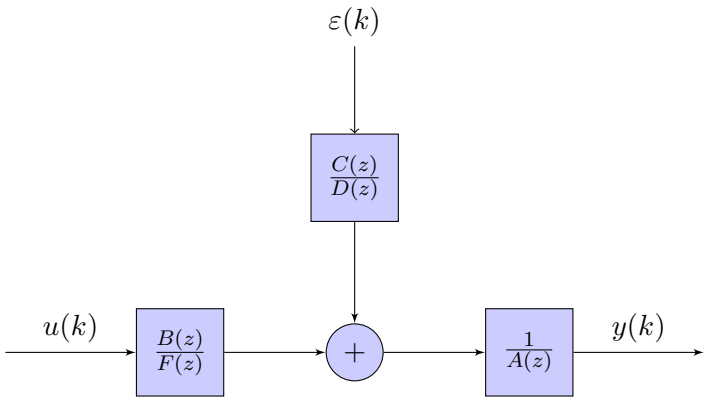
Obrázek 4.2: Struktura ARMAX modelu

Box-Jenkins (BJ)

Tato struktura je nejflexibilnější, avšak nejsložitější. Dovoluje nám úplně rozdělit popis dynamiky systému a chování šumu. Box-Jenkins strukturu je vhodné použít, pokud šum není na vstupu, ale až na výstupu. Matematicky se vyjadřuje rovnicí

$$y(k) = \sum_{i=1}^{n_u} \frac{B_i(z)}{F_i(z)} u_i(k - nk_i) + \frac{C(z)}{D(z)}\varepsilon(k) \tag{4.4}$$

a vyobrazena je na obrázku 4.3



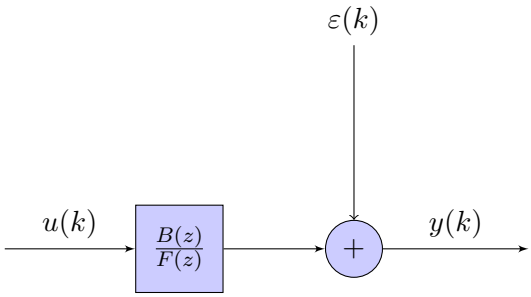
Obrázek 4.3: Struktura Box-Jenkins modelu

Output-Error (OE)

OE struktura popisuje pouze dynamiku systému, neboli je vhodné ji použít, pokud chceme modelovat právě jen dynamiku systému a dynamiku šumu nikoliv. Matematický popis vypadá následovně

$$y(k) = \sum_{i=1}^{n_u} \frac{B_i(z)}{F_i(z)} u_i(k - nk_i) + \varepsilon(k) \tag{4.5}$$

. Vyobrazení se nachází na obrázku 4.4.



Obrázek 4.4: Struktura Output-Error modelu

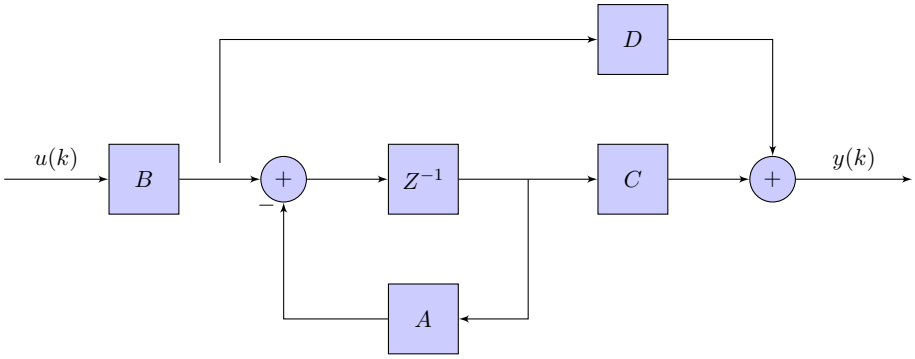
4.1.2 Stavový model

Tento model využívá k popisu systému soustavu differenciálních (pro spojité systémy), nebo differenčních (pro nespojité systémy) rovnic prvního řádu [8]. Rovnice 4.6 tento model popisuje.

$$\begin{aligned} x(k+1) &= \mathbf{A}x(k) + \mathbf{B}u(k), \\ y(k) &= \mathbf{C}x(k) + \mathbf{D}u(k), \\ x(0) &= x_0 \end{aligned} \tag{4.6}$$

kde proměnné $x(k)$, $y(k)$, $u(k)$ představují pořadě stavový vektor,výstup systému a vstup systému. Dále matice $\mathbf{A}$ popisuje vztah mezi stavy systému

v aktuálním a následujícím kroku, matice **B** popisuje vztah mezi vstupy do systému a stavy systému v následujícím kroku, matice **C** popisuje vztah mezi stavy systému a výstupy systému a matice **D** popisuje vztah mezi vstupy do systému a výstupy systému (velice často bývá nulová).



Obrázek 4.5: Struktura stavového modelu

4.2 Nelineární modely

Obecný tvar nelineárních modelů vypadá následovně

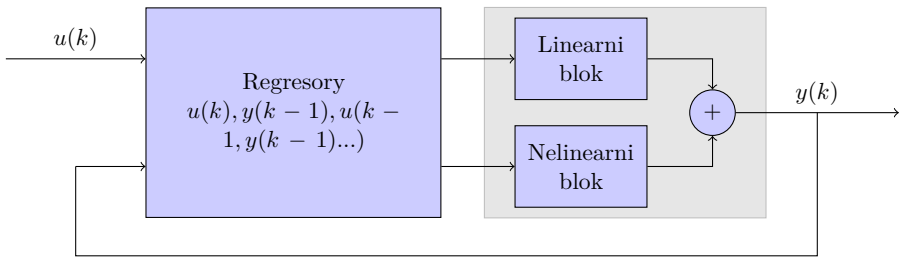
$$y(k) = f(y(k-1), y(k-2), ..., y(k-n_y), u(k), u(k-1), u(k-2), ..., u(k-n_u)) + \varepsilon(k) \tag{4.7}$$

kde y je výstup u vstup, ε je bílý šum , n_y řád výstupu a n_u řád vstupu. Model v tomto tvaru je nelineární, pokud funkce f je nelineární. f může vyjadřovat libovolné nelinearity, jako například saturace nebo přepínače.

V praxi jsou všechny systémy nelineární. Většina lze ale popsat jednoduššími lineárními modely, které mohou být dostačující. Nelineární modely by se měly brát v úvahu až když nelze systém popsat lineárními modely, aby byl dostačující, nebo o systému víme, že je nelineární a nebo je systém popsaný lineární a nelineární částí.

4.2.1 Nonlinear ARX

Nelineární ARX model je rozšířený ARX model o nelineární část. Matematicky je popsán rovnicí 4.7. Blokové schéma nelineárního ARX systému je vyobrazeno na obrázku 4.6

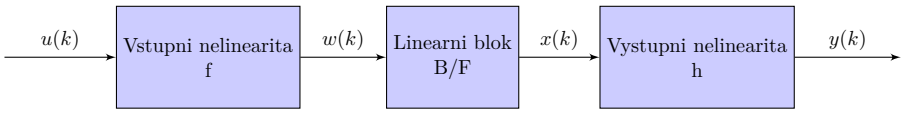


Obrázek 4.6: Struktura NARX modelu

Lineární blok tvoří ARX model a nelineární část tvoří jeden ze tří následujících modelů: *dopředná neuronová síť (sigmoid network)*, *waveletová neuronová síť (wavelet network)* a *binární strom (binary tree)* [1]. Pro odhad parametrů těchto modelů byly použity následující metody: Levenberg-Marquardt algoritmus, Gauss-Newton algoritmus

4.2.2 Hammerstein-Wiener

Hammersteinův model se skládá z nelineárního bloku po němž následuje lineární blok. Wienerův model využívá tytéž bloky, ovšem v obráceném pořadí, neboli nejdříve má blok lineární a poté následuje nelineární. Kombinace těchto dvou modelů nám dává Hammerstein-Wiener model, jehož schéma popisuje obrázek 4.7.



Obrázek 4.7: Struktura Hammerstein-Wiener modelu

kde $w(k) = f(u(k))$ je nelineární funkce transformující vstupní data $u(t)$, pričemž dimenze je zachována.

$x(t) = \frac{B}{F}w(k)$ je lineární funkce a B a F jsou polynomy podobné modelu Output-Error modelu popsaného výše.

$y(t) = h(x(k))$ je nelineární funkce tranformující výstupní data lineárního bloku na výstupní data systému. Lineární blok tvoří OE model a nelineární část tvoří jeden ze tří následujících modelů: *po částech lineární nelinearita (piecewise-linear)*, *dopředná neuronová síť (sigmoid network)*, *waveletová neuronová síť (wavelet network)* [1]. Pro odhad parametrů těchto modelů byly použity následující metody: Levenberg-Marquardt algoritmus a Gauss-Newton algoritmus

Pro určení parametrů nelineárních regresorů lze použít více metod odhadu: *piecewise-linear nonlinear estimator*, *sigmoid network nonlinear estimator* a *wavelet network nonlinear estimator* [1].

Kapitola 5

Experimenty a výsledky

V praktické části mé práce jsem se zabýval identifikací systémů z dvou různých souborů dat. První soubor dat pocházel ze simulátoru a druhý soubor byl získán reálným měřením na budově F40. Všechna tyto data byla vzorkována s periodou 1 hodina a znázorňovala průměrnou hodnotu za uplynulou hodinu. Pro získání modelu jsem se držel obecnými kroky popsaný v kapitole 2. Tyto kroky dále rozeberu zvlášť pro jednotlivé soubory dat. Pro určení parametrů modelu jsem použil software Matlab® System Identification Toolbox.

Označení modelů v tabulkách jsou mnou vymyšlené, abych se na ně v textu mohl odkazovat. U lineárních modelů název tvoří zkratka modelu, jeho řády a koncovka s, nebo r, jestli se jedná o simulovaná, či reálná data. Například ARX model s řády $n_a = 1$, $n_b = 2$, $n_k = 0$ vytvořený pro simulovaná data má název $arx120s$. Význam použitých řádů je popsán v kapitole 4. Má označení nelineárních modelů vznikly složitějším procesem, proto je popisuji dále v textu u každého modelu zvlášť.

5.1 Identifikace ze simulovaných dat

První soubor dat pochází ze simulace simulinkového modelu popsaného v sekci 3.2. Vstupem do tohoto modelu jsou setpointy teplot v zónách, setpoint teploty topící vody, tok vody a naměřené veličiny počasí. Tyto vstupy jsou popsané v tabulce 3.1. Naměřené veličiny počasí jsem obdržel spolu s modelem. Tato data byla naměřena na reálné budově F40, tudíž tyto vstupy jsem nemusel předzpracovávat.

Výhodou modelu oproti měření reálných dat, je možnost si systém vybudit jakýmkoliv signálem a data budu mít k dispozici ihned a pokud jsou data neuspokojivá, mohu provádět experimenty dokud nedosáhnu požadovaných dat pro identifikaci. Na reálném systému by byly tyto experimenty nákladné jak finančně, tak časově náročné.

5.1.1 Vstupní signály simulátoru

Pro vybuzení systému jsem měl možnost nastavit si setpointy teplot v jednotlivých zónách a setpoint teploty vyhřívací vody. Všechny požadované teploty byly nastavovány s ohledem na denní dobu (den/noc). Setpointy

Označení	Popis	Jednotka
T_{s1} - T_{s9}	Setpoint teploty vzduchu v zóně 1-9	$[°C]$
T_1 - T_9	Teplota vzduchu v zóně 1-9	$[°C]$
T_w	Teplota vytápěcí vody	$[°C]$
W_1	Venkovní teplota vzduchu	$[°C]$
W_2	Intenzita slunečního záření	$[W/m^2]$
Q	Hodinová spotřeba energie pro vytápění	$[kWh]$

Tabulka 5.1: Simulovaná data použitá pro identifikaci

teplot vzduchu byly pro každý den náhodné. Pohybovaly se od 17^oC do 23^oC a přes noc byla požadovaná teplota vzduchu vždy nastavena na 17^oC (tzv. setback). Požadované hodnoty vyhřívací vody byly také nastaveny náhodně od 35^oC do 70^oC a měnily se s periodou 24 hodin. Skoky teplot jsem volil náhodně, abych tak lépe zachytil dynamiku systému. Další možností bylo také použít takzvaný pseudonáhodný binární vstupní signál (PRBS). Jedná se o signál, který nabývá hodnot 0, nebo 1 a to v pseudonáhodném pořadí. Tento signál je vhodný pro identifikaci lineárních systémů, protože deterministickým signálem napodobuje bílý šum v diskrétním čase a vybuzuje všechny frekvence rovnoměrně. Ještě vhodnějším signálem by mohl být aplitudově modulovaný pseudonáhodný binární signál (APRBS), protože dokáže měnit i aplitudu [9]. Tyto signály jsem však nepoužil, protože jsem se chtěl alespoň trochu přiblížit možnému reálnému nastavování teploty, hlavně aby souhlasilo nastavení teplot dne oproti noci.

5.1.2 Simulovaná data pro identifikaci

Pro identifikaci parametrů jsem jako vstupní data zvolil teploty v zónách, setpointy teploty v zónách a teplotu vytápěcí vody. Dále jsem z počasí zvolil venkovní teplotu a okolní sluneční záření, protože dokáží vnitřní teplotu budovy ovlivnit až o jednotky ^oC, což má také nepřímý dopad na spotřebu energie. Jako výstupní data jsem použil hodinovou spotřebu energie pro vytápění. Celkovou strukturu dat použitých pro identifikaci popisuje tabulka 5.1 a příklad vstupů znázorňuje graf B.1 v přílohách. Tato simulovaná data byla dále rozdělena do 3 sad. První sada (1. 1. 2005 - 8. 2. 2005) sloužila k identifikaci parametrů, druhá sada (9. 2. 2005 - 17. 3. 2006) k ladění a poslední sada (spojená z dat z období 1. 1. 2006 - 17. 3. 2006 a 1. 1. 2007 - 17. 3. 2007) k validaci.

5.1.3 Identifikované modely

Ze simulovaných dat jsem vytvořil modely **ARX**, **ARMAX**, **stavový model** a **NARX**. Identifikace jednotlivých modelů jsem popsal níže.

Název modelu	na	nb	nk	MSE $[(kWh)^2]$	MAPE [%]	FIT [%]
arx110s	1	1	0	0.00031078	22.71	89.20
arx120s	1	2	0	0.00030889	22.01	89.23
arx210s	2	1	0	0.00030529	22.22	89.29
arx220s	2	2	0	0.00029403	21.00	89.49
arx111s	1	1	1	0.00054831	35.43	85.65
arx121s	1	2	1	0.00054832	35.83	85.65
arx211s	2	1	1	0.00052492	33.82	85.96
arx221s	2	2	1	0.00053863	35.28	85.78
arx112s	1	1	2	0.00058011	37.18	85.24
arx122s	1	2	2	0.00057010	35.65	85.37
arx212s	2	1	2	0.00057452	36.73	85.31
arx222s	2	2	2	0.00057011	35.49	85.37

Tabulka 5.2: Porovnání ARX modelů různých řádů
Modely byly vytvořeny ze simulovaných dat.
n_a, n_b, n_k jsou pořadě řádu výstupu, řády vstupů a dopravního zpoždění vstupů

ARX

Tento model slouží hlavně k základnímu určení řádů vstupů a výstupů. Je zásadou, že je lepší zvolit nižší řád, pokud nám vyšší řád výrazně nezlepší výsledek. Model s vyššími řády může estimační data lépe napodobit, ovšem za cenu větší nejistoty parametrů [1]. Z tohoto důvodu se může stát, že systém s vyšším řádem může na jiné sadě dat dosáhnout horších výsledků.

Řád výstupu (n_a), vstupů (n_b) a dopravního zpoždění (n_k) pro model ARX jsem zkoumal v rozsahu 1 až 2, 1 až 2 a 0 až 2. Tyto řády jsem zvolil proto, abych předešel problému posaného výše. Určení parametrů samotných modelů pomocí metody QR-factorization [10]. Výsledky všech ARX modelů se nachází v tabulce 5.2, která popisuje modely ARX různých řádů a jejich výsledky na ladících datech. Graf B.2 ukazuje odezvu zvoleného modelu $arx110s$ na vstupní signál z ladících dat a porovnává jí k reálné hodnotě.

ARMAX

Při určování řádů tohoto modelu jsem postupoval obdobně jako u modelu ARX. Při volbě řádu dopravního zpoždění jsem využil předchozích výsledků porovnání ARX modelů a zvolil jsem řád $n_k = 0$, neboť ARX model dosahuje výrazně lepších výsledků oproti vyšším řádům a ARMAX se od ARX modelu liší jen řádem klouzavého průměru bílého šumu (n_c). Řád výstupu (n_a), vstupů (n_b) a řád chyby n_c jsem zvolil znovu jen nízké: n_a v rozsahu 1-2, n_b v rozsahu 1-3 a n_c v rozsahu 1-2. Pro určení parametrů samotných modelů byla použita Gauss-Newtonova prohledávací metoda [10]. Porovnání mnou vytvořených ARMAX modelů se nachází v tabulce 5.3, která popisuje modely ARMAX různých řádů a jejich výsledky na ladících datech. Vizuálně model $arma2210s$ porovnat v grafu B.3.

Název modelu	na	nb	nc	nk	MSE $[(kWh)^2]$	MAPE [%]	FIT [%]
armax1110s	1	1	1	0	0.00030591	22.05	89.28
armax1120s	1	1	2	0	0.00030238	21.99	89.34
armax1210s	1	2	1	0	0.00029581	21.54	89.46
armax1220s	1	2	2	0	0.00029297	21.32	89.51
armax2110s	2	1	1	0	0.00030346	21.95	89.33
armax2120s	2	1	2	0	0.00030285	21.93	89.34
armax2210s	2	2	1	0	0.00029165	20.92	89.54
armax2220s	2	2	2	0	0.00029246	21.11	89.52

Tabulka 5.3: Porovnání ARMAX modelů různých řádů
Modely byly vytvořeny ze simulovaných dat.
n_a, n_b, n_c, n_k jsou pořadě řády výstupu, vstupů, řád klouzavého průměru šumu a dopravního zpoždění vstupů.

Název modelu	Řád modelu	MSE $[(kWh)^2]$	MAPE [%]	FIT [%]
statespace1s	1	0.00049362	32.90	86.39
statespace2s	2	0.00048680	33.67	86.48
statespace3s	3	0.00081837	45.69	82.47
statespace4s	4	0.00068562	39.73	83.96
statespace5s	5	0.00070240	40.15	83.76
statespace6s	6	0.00045036	30.90	87.00
statespace7s	7	0.00048879	32.10	86.45
statespace8s	8	0.00085525	43.96	82.08
statespace9s	9	0.00050029	32.77	86.30
statespace10s	10	0.00107830	51.68	79.88

Tabulka 5.4: Porovnání Stavových modelů různých řádů
Modely byly vytvořeny ze simulovaných dat.

Stavový model

U tohoto modelu se volí řád stavů n, který odpovídá počtu stavů stavového vektoru x. Můj stav je jen jeden - spotřeba energie. Zvolím-li vyšší řád než je reálný počet stavů ($n > 1$), další stavy budou vyjadřovat opožděný původní stav. Neboli čím vyšší řád modelu, tím více minulých hodnot výstupů se započítá do výstupu. Jedná se o obdobu řádu n_a u ARX a ARMAX modelů. Jelikož se volí jen tento řád, mohl jsem prozkoumat více modelů, tudíž jsem zvolil řády 1 až 10. Pro identifikaci parametrů byl použit prohledávací algoritmus n4sid [11] se zaměřením na predikci. Výsledky všech stavových modelů se nachází v tabulce 5.4. Označení modelu vychází z anglického překladu State Space model, číslo znázorňuje řád modelu s přípona s značí, že jde o simulovaná data. Porovnání odezvy modelu *statespace*1 a skutečné hodnoty lze vidět v grafu B.4.

NARX

Jelikož se NARX model skládá z lineární a nelineární části, zvolil jsem za lineární část modely $arx110s$ a $arx120s$ z předchozí identifikace ARX modelů. Z nelineárních algoritmů jsem mohl volit mezi *sigmoid network nonlinear estimator*, *wavelet network nonlinear estimator* a *binary tree nonlinear estimator*. Tyto algoritmy jsou blíže popsány v dokumentaci Matlab® System

Název	Nelineární blok	Lineární blok	MSE $[(kWh)^2]$	MAPE [%]	FIT [%]
narx110ss	sigmoid network	arx110s	0.00028733	19.54	89.61
narx120ss	sigmoid network	arx120s	0.00029307	20.84	89.51
narx110ws	wavelet network	arx110s	0.00031076	22.67	89.20
narx120ws	wavelet network	arx120s	0.00030900	21.98	89.23
narx110bs	binary tree	arx110s	0.00028442	21.13	89.67
narx120bs	binary tree	arx120s	0.00028463	21.27	89.66

Tabulka 5.5: Porovnání NARX modelů
Modely byly vytvořeny ze simulovaných dat. Parametry nelineární části byly určeny algoritmy *sigmoid network nonlinear estimator*, *wavelet network nonlinear estimator* a *binary tree nonlinear estimator* [1]. Každá síť obsahovala 10 neuronů. Lineární část modelu tvoří jeden ARX v 5.2.

Identification Toolbox[1]. Vytvořil jsem kombinace zvolených ARX modelů s temito nelineárními modely. Všechny tyto modely využívaly při výpočtech 10 neuronů. Porovnání proběhlo na ladících datech a výsledky jsou v tabulce 5.5. Označení modelů vychází z názvu modelu - NARX, řádů použitého ARX modelu a příponu tvoří první písmeno použité nelineární části a s znamená, že identifikace proběhla na simulovaných datech. Porovnání odezvy modelu vybraného $narx110ss$ a skutečné hodnoty lze vidět v grafu B.4.

Výsledné srovnání všech modelů

Pro výsledné srovnání jsem z každé předchozí části vybral jeden reprezentativní model. Pokud bych porovnával tyto modely na datech, na kterých jsem je vybíral, bylo by výsledné srovnání nevěrohodné. Z toho důvodu jsem tyto modely porovnal na validační sadě dat a výsledné porovnání vybraných modelů je znázorněno v tabulce 5.6.

Název modelu	MSE $[(kWh)^2]$	MAPE [%]	FIT [%]
arx110s	0.00031248	27.21	90.42
armax2210s	0.00031062	25.73	90.61
narx110ss	0.00028094	23.70	90.91
statespace1s	0.00047395	34.61	88.20

Tabulka 5.6: Porovnání modelů s různou strukturou
Modely byly vytvořeny ze simulovaných dat (přípona s v označení modelu). Označení modelů jsou mým označením a vychází ze jména struktury modelu a jeho řádech. V případě NARX modelu přípona s značí, že jeho nelineární část tvoří sigmoid network.

5.2 Identifikace na reálných datech

Druhý soubor dat pochází z měření na reálné budově F40. Jedná se o měření ze zimy 2015. Celkově byly měřeny teploty vzduchu a jejich setpointy ve 43 místnostech, teplota vyhřívací vody a spotřeba energie pro vytápění. Z vlivů počasí byly měřeny veličiny venkovní teplota a vlhkost vzduchu, tlak, rychlost vzduchu s jeho směrem, dešťové srážky a intenzita slunečního záření.

Aby identifikace nebyla vzhledem k počtu vstupů příliš složitá, byly místnosti sloučené do zón 3.2. Poté byla pro každou zónu vypočtena průměrná

Označení	Popis	Jednotka
T_{s1} - T_{s10}	Setpoint teploty vzduchu v zóně 1-10	$[°C]$
T_1 - T_{10}	Teplota vzduchu v zóně 1-10	$[°C]$
T_w	Teplota vytápěcí vody	$[°C]$
W_1	Venkovní teplota vzduchu	$[°C]$
W_2	Intenzita slunečního záření	$[W/m^2]$
Q	Hodinová spotřeba energie pro vytápění	$[kWh]$

Tabulka 5.7: Reálná data použitá pro identifikaci

teplota vzduchu. Naměřená data vykazovala mnoho chyb, jejichž původ může být různý (např. krátkodobá porucha senzoru apod.). Z dat byly tedy jen vybrané celé dny, které tyto chyby neobsahovaly. Výsledkem byla data o délce 15 dnů.

Takto předzpracovaná data jsem obržel pro identifikaci. Jejich struktura je vyobrazená v tabulce C.1. Do identifikačního procesu jsem použil jen některá data, aby nebyly modely příliš složité. Ze vstupů jsem opět zvolil setpointy a skutečné teploty vzduchu v zónách, teplotu vyhřívací vody a intenzitu slunečního záření. Vzhledem k tomu, že použitelných dat bylo málo, rozdělil jsem tyto data pouze do dvou souborů dat. Identifikační sada dat obsahovala 11 dnů a validační 4 dny. Struktura dat použitých při identifikaci ukazuje tabulka 5.7. Příklad těchto dat je vyobrazen v grafu B.6.

Postup určování řádů a parametrů modelů se od identifikace ze simulovaných dat ve většině případech nelišil. Z tohoto důvodu jsem celé postupy pro jednotlivé modely nepopsal a zmiňuji jen odlišnosti oproti předchozí identifikaci na simulovaných datech.

ARX

Oproti tvorbě modelů ze simulovaných dat jsem zkoumal jen modely s dopravním zpožděním $n_k = 0$, protože dopadly výrazně lépe než modely s vyšším řádem dopravního zpoždění. Porovnání všech ARX modelů vy tvořených z reálných dat se nachází v tabulce 5.8, která popisuje modely ARX různých řádů a jejich výsledky na validačních datech. Graf B.7 ukazuje odezvu zvoleného modelu $arx120r$ na vstupní signál z ladících dat a porovnává jí k reálné hodnotě.

ARMAX

Rozsah řádů jem vybral stejné jako při identifikaci na simulovaných datech. Zkoumal jsem tedy řády řády výstupu n_a v rozsahu 1-2, vstupu n_b v rozsahu 1-2, chyby n_c v rozsahu 1-2 a řád dopravního zpoždění jsem nechal $n_k = 0$. Porovnání mnou vytvořených ARMAX modelů z reálných dat se nachází v tabulce 5.9, která popisuje modely ARMAX různých řádů a jejich výsledky na validačních datech. Vizuálně lze model $armax1210r$ porovnat v grafu B.8.

Název modelu	na	nb	nk	MSE $[(kWh)^2]$	MAPE [%]	FIT [%]
arx110r	1	1	0	0.00040031	9.85	80.61
arx120r	1	2	0	0.00019796	6.92	86.36
arx130r	1	3	0	0.00032639	9.72	82.49
arx210r	2	1	0	0.00041834	9.63	80.17
arx220r	2	2	0	0.00019495	6.91	86.46
arx230r	2	3	0	0.00034646	9.90	81.96

Tabulka 5.8: Porovnání ARX modelů různých řádů
Tyto modely byly vytvořeny z reálných dat.
n_a, n_b, n_k jsou pořadě řády výstupu, vstupů a dopravního zpoždění vstupů

Název modelu	na	nb	nc	nk	MSE $[(kWh)^2]$	MAPE [%]	FIT [%]
armax1110r	1	1	1	0	0.00039322	9.57	80.78
armax1120r	1	1	2	0	0.00037447	9.35	81.24
armax1210r	1	2	1	0	0.00029779	8.75	83.27
armax1220r	1	2	2	0	0.00321590	39.77	45.04
armax2110r	2	1	1	0	0.00038268	9.04	81.04
armax2120r	2	1	2	0	0.00042164	9.77	80.10
armax2210r	2	2	1	0	0.00316540	38.43	45.47
armax2220r	2	2	2	0	0.00211280	30.58	55.45

Tabulka 5.9: Porovnání ARMAX modelů různých řádů
Tyto modely byly vytvořeny z reálných dat.
n_a, n_b, n_c, n_k jsou pořadě řády výstupu, vstupů, chyby a dopravního zpoždění vstupů.

Stavový model

Oproti identifikaci na simulovaných datech jsem zvolil nižší nejvyšší řád, protože vyšší řády modelu by na takhle malém souboru dat o ničem nevypovídaly. Označení modelu se liší posledním písmenkem r - reálná data. Porovnání odezvy modelu $statespace1r$ a skutečné hodnoty ukazuje graf B.9.Tabulka výsledků se stavovými modely 5.4 porovnává jejich odezvy na validačním souboru dat.

NARX

Postup se oproti identifikaci na simulovaných datech lišil jen ve volbě lineárních modelů. Pro tuto identifikaci jsem zvolil modely $arx120r$ a $arx220$, jelikož dosahovaly nejlepších výsledků. Porovnání proběhlo na validačních datech a výsledky jsou v tabulce B.5. Označení se jen liší v příponě r - reálná data. Výstup jednoho z modelů lze pozorovat v grafu B.5.

Output-Error

Pro tento model jsem zkoumal řády polynomů B, F popsaných v kapitole 4 a řád dopravního zpoždění vstupů. Řád nb jsem zkoumal v rozsahu 1-2,

Název modelu	Řád modelu	MSE $[(kWh)^2]$	MAPE [%]	FIT [%]
statespace1r	1	0.0007196	13.98	74.00
statespace2r	2	0.0005016	14.14	78.29
statespace3r	3	0.0011355	16.55	67.34
statespace4r	4	0.0015103	21.88	62.33

Tabulka 5.10: Porovnání State-Space modelů různých řádů
Tyto modely byly vytvořeny z reálných dat.

Název	Nelineární blok	Lin. blok	MSE $[(kWh)^2]$	MAPE [%]	FIT [%]
narx120sr	sigmoid network	arx120r	0.01520900	55.55	-19.51
narx220sr	sigmoid network	arx220r	0.01449500	60.29	-16.67
narx120wr	wavelet network	arx120r	0.00019794	6.92	86.36
narx220wr	wavelet network	arx220r	0.00019494	6.91	86.46
narx120br	binary tree	arx120r	0.00016097	5.95	87.70
narx220br	binary tree	arx220r	0.00014868	5.72	88.18

Tabulka 5.11: Porovnání NARX modelů
Tyto modely byly vytvořeny z reálných dat. Parametry nelineární části byly určeny algoritmy *sigmoid network nonlinear estimator*, *wavelet network nonlinear estimator* a *binary tree nonlinear estimator* [1]. Každá síť obsahovala 10 neuronů. Lineární část modelu tvoří vybrané ARX modely z 5.2

řád nf 1-3 a řád nk 0-2. Tabulka 5.12 porovnává modely, které měly MAPE nižší než 10. Porovnání odezvy modelu $oe120r$ a skutečné hodnoty lze vidět v grafu B.11.

Hammerstein-Wiener

Hammerstein-Wiener model se skládá z lineární a dvou nelineárních části. Jedna nelineární část je na vstupu a jedna na výstupu modelu. Mezi nimi je lineární model. Z tohoto důvodu jsem zvolil za lineární model Ouput-Error model (konkrétně $oe120r$), protože tento model popisuje pouze dynamiku systému. O výstupní chybu se stará nelineární model. Parametry nelineární části byly určeny algoritmy *piecewise-linear nonlinear estimator sigmoid network nonlinear estimator* a *wavelet network nonlinear estimator*.Tyto algoritmy využívaly při výpočtech 10 neuronů a jsou blíže popsány v dokumentaci Matlab® System Identification Toolbox[1]. Porovnání výstupu modelu $hww120pr$ se skutečností lze pozorovat na grafu B.12 obsaženého v příloze. Kompletní výsledky a porovnání popisuje tabulka 5.13. Název modelu je složen ze zkratky Hammerstein-Wiener - hw, prvního písmena vstupního nelineárního modelu, řády Output-Error modelu, prvního písmena výstupního nelineárního modelu a koncovka r znamená pro reálná data.

Název modelu	nb	nf	nk	MSE $[(kWh)^2]$	MAPE [%]	FIT [%]
oe120r	1	2	0	0.00011603	5.15	89.56
oe132r	1	3	2	0.00014446	5.71	88.35
oe212r	2	1	2	0.00021986	8.28	85.63
oe222r	2	2	2	0.00018161	7.17	86.94

Tabulka 5.12: Porovnání Output-Error modelů různých řádů
Tyto modely byly vytvořeny z reálných dat.
n_b, nf jsou pořadě řády polynomů B,F 4 a n_k je řád dopravního zpoždění vstupů

Název	Vstupní NL	Výstupní NL	MSE $[(kWh)^2]$	MAPE [%]	FIT [%]
hws120sr	sigmoid network	sigmoid network	0.00119700	21.38	66.47
hws120wr	sigmoid network	wavelet network	0.00102000	18.94	69.05
hws120pr	sigmoid network	piecewise linear	0.00079646	21.49	72.65
hww120sr	wavelet network	sigmoid network	0.00080862	16.39	72.44
hww120wr	wavelet network	wavelet network	0.00186670	28.75	58.13
hww120pr	wavelet network	piecewise linear	0.00054128	12.44	77.45
hwp120sr	piecewise-linear	sigmoid network	0.00158140	22.76	61.46
hwp120wr	piecewise-linear	wavelet network	0.00055757	13.35	77.11
hwp120pr	piecewise-linear	piecewise linear	0.00052176	16.81	77.86

Tabulka 5.13: Porovnání Hammerstein-Wiener modelů
Tyto modely byly vytvořeny z reálných dat. Parametry nelineárních částí (NL) byly vypočteny algoritmy *piecewise-linear nonlinear estimator*, *wavelet network nonlinear estimator* a *sigmoid network nonlinear estimator* [1]. Každá síť obsahovala 10 neuronů. Lineární část modelu tvoří model $oe120r$ 5.12

Kapitola 6

Diskuze

Modely vytvořené ze simulovaných dat

ARX modely s nižší řádem dopravního zpoždění dosahovaly nižšího MAPE, protože se v systému žádné dopravní zpoždění nevyskytuje. Modely s vyššími řády výstupů a vstupů vykazovaly lepší výsledky, ovšem pro další použítí bych je ale nezvolil, protože jejich zlepšení výsledků není tak významný. Vhodnější je zvolit model s nižšími řády, neboť tím klesá složitost a nejistota modelu [1].

ARMAX model se od ARX modelu liší jen řádem klouzavého průměru bílého šumu n_c, proto jsem mohl použít předešlé výsledky a zkoumat jen modely s dopravním zpožděním řádu 0. Tyto modely se výsledky také mnoho od sebe nelišily. Může to být dáno tím, že vyšší řády vstupů a výstupů nejsou pro aktuální výstup tolik důležité a tudíž jim prohledávací algoritmus mohl určit nižší váhu.

Stavový model dopadl ze všech srovnávaných modelů nejhůře. Je to patrně dáno tím, že ze všech modelů je nejméně flexibilní. Jediný parametr, který mohl být u tohoto modelu zvolen byl v podstatě jen řád výstupu.

Nejlépe ze všech modelů dopadl nelineární model NARX. Jeden z možných důvodů může být ten, že nelineární část tohoto modelu při určování parametrů regresorů mnohem flexibilnější než u lineárních modelů. Ovšem tato složitost by se mohla být nežádoucí, neboť výpočetní složitost při reálném použití by byla vyšší a přitom by model nedosahoval o mnoho lepších výsledků.

Co se validačních hodnot v porovnávací tabulce 5.6 týče, jsou jen orientační. Reálné chyby mohou být ve skutečnosti vyšší, ale jelikož nemáme nekonečně velký vzorek dat, tak přesnou chybu ani nemůžeme určit.

Mnou vytvořené modely slouží pouze jako orientační možnosti použití black box modelovacích metod na budově F40 a pro reálné použití například pro řízení, či optimalizaci spotřeby, nejsou vhodné.

Modely vytvořené z naměřených dat

Modely vytvořené z reálných dat dosahovaly nižších MAPE, než modelů identifikovaných ze simulovaných dat. Příčina může být ta, že v reálné budově nebyly teploty vyhřívací vody tak často a tak razantně měněny, jako tomu bylo u simulovaných dat, přičemž tento vstup je nejvíce spjatý se spotřebou. Další přičinu může tvořit množství validačních dat. Na takto malém vzorku

dat (96 hodin) jsou výsledky velmi nepřesné. Z tohoto důvodu jsem ani různé struktury modelů mezi sebou neporovnával. Důvodem vyšší chyby MAPE u modelů vytvořených ze simulovaných dat může být velká variabilita simulovaného signálu pozorovaná na obrázcích 4.1 až 4.6. U predikce reálných signálů se tato variabilita neobjevuje a může být tedy chybou simulátoru.

Cílem vytvoření modelů nebylo získat použitelný model, ale zjistit, zda mají black box metody pro modelování budovy F40 smysl. Z tohoto pohledu si myslím, že smysl mají, protože lze relativně (ke složitosti gray box a white box modelů) jednoduchými modely popsat dynamiku vytápění budovy, která bude dostatečná pro řízení.

Kapitola 7

Závěr

Cílem této bakalářské práce bylo vytvořit black box modely na základě dat získaných poskytnutého simulátoru budovy F40 a reálných naměřených dat. Dále je cíl zhodnotit vhodnost použití black box modelovacích technik pro vytvoření modelu této budovy.

V úvodu jsem popsal o jaký modelovaný systém se jedná, kde se nachází a za jakých podmínek byla získávána data z tohoto systému. Jednalo se o 3 patrový kancelářský dům, který byl rozdělen do určitých zón, ve kterých se měřila teplota vzduchu.

V teoretické části jsem se zabýval obecně problémem identifikace a black box modely. Shrnul jsem a matematicky popsal používané struktury modelů. Dále jsem popsal v čem se liší, jejich výhody a nevýhody jejich použití.

V praktické části jsem získával data z poskytnutého simulátoru. Tento model jsem budil náhodnými skokovými signály s konstantní periodou, abych získal odezvy, které mi poté usnadnily identifikaci parametrů modelu. Data jsem poté rozdělil na identifikační, ladící a validační. Identifikační sada dat sloužila pro identifikaci parametrů modelů, ladící pro výběr řádů modelů a validační pro srovnání identifikovaných modelů mezi sebou. Po získání těchto dat jsem vytvořil black box modely a identifikoval jejich parametry. Vybrané modely odlišných struktur jsem poté mezi sebou porovnal na validační sadě dat. Z porovnávaných modelů dosahoval model $narx110ss$ nejlepšího výsledku s hodnotou MAPE 23.70. Bohužel tento model není použitelný pro řízení, ani optimalizaci, jelikož taková chyba je poměrně vysoká.

Před samotným modelováním z reálných dat jsem provedl jejich analýzu. Mnoho dat bylo nepoužitelných, protože se v nich vyskytovaly nejednoznačně identifikovatelné chyby (například porucha senzoru). Výsledkem byl teda velmi omezený soubor dat, na kterém mohla identifikace a validace proběhnout. Modely na základě dat z reálné budovy dosahovaly paradoxně nižších MAPE (tzn. lépe) než ze modely vytvořené ze simulovaných dat. Tyto hodnoty ale nelze považovat za věrohodné, neboť byly vypočítány na velmi malém vzorku dat.

Mnou vytvořené black box modely nejsou vhodné pro řízení, ani optimalizaci vytápění budovy. Mohou však sloužit jako základ pro budoucí výzkum této oblasti. Tyto modely mohou být dále optimalizovány například volbou jiných vstupních dat (například zvolit jiné vstupy, menší vzorkování atp.) nebo by

stálo za úvahu zkoumat různé řády pro každý vstup zvlášť. Možností pro ladění modelů je mnoho, avšak tato oblast přesahuje rámec této práce a pro validaci by bylo zapotřebí mnohem více dat.

Výsledkem mé práce je tedy přehled black box modelů, základní implementace ze simulovaných a reálných dat a jejich hrubé porovnání. Tímto jsem poskytl základ pro možné pokračování, popsané v předchozím odstavci, čímž považuji hlavní cíl této bakalářské práce za splněný.

Příloha A

Literatura

[1] L. Ljung, *MATLAB: System Identification Toolbox, User's Guide*. USA, 2015.

[2] L. Perez-Lombard, J. Ortiz, and C. Pout, "A review on buildings energy consumption information," *Energy and buildings*, vol. 40, no. 3, pp. 394–398, 2008.

[3] H.-x. Zhao and F. Magoulès, "A review on the prediction of building energy consumption," *Renewable and Sustainable Energy Reviews*, vol. 16, no. 6, pp. 3586–3592, 2012.

[4] F. Moretti, "F40 modeling," 2014.

[5] d. M. Wit, "Hambase: heat, air and moisture model for building and systems evaluation," 2006.

[6] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," *International journal of forecasting*, vol. 22, no. 4, pp. 679–688, 2006.

[7] C. I. R. P. (CIRP), "Dimensional statistics," Feb. 2015.

[8] N. Instruments, *Selecting a Model Structure in the System Identification Process*, 2010.

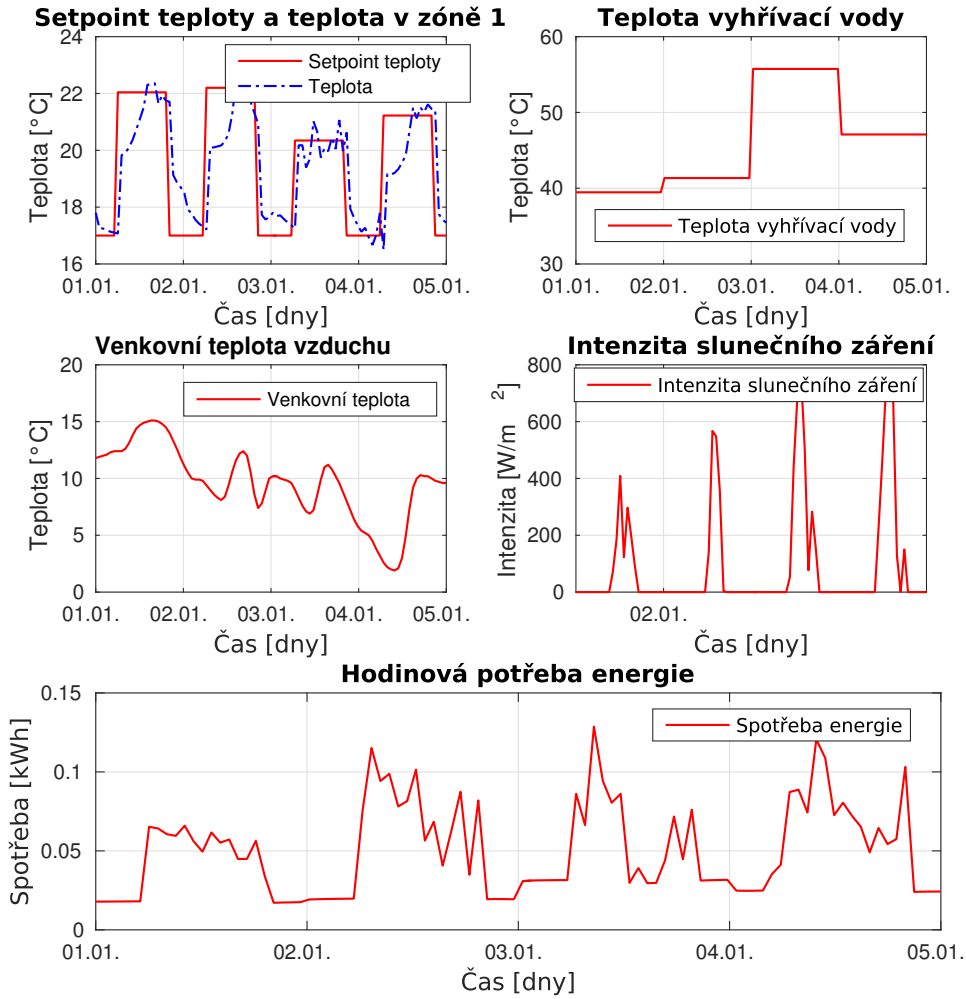
[9] O. Nelles, *Nonlinear system identification: from classical approaches to neural networks and fuzzy models*. Springer Science & Business Media, 2001.

[10] L. Ljung, ed., *System Identification (2Nd Ed.): Theory for the User*. Upper Saddle River, NJ, USA: Prentice Hall PTR, 1999.

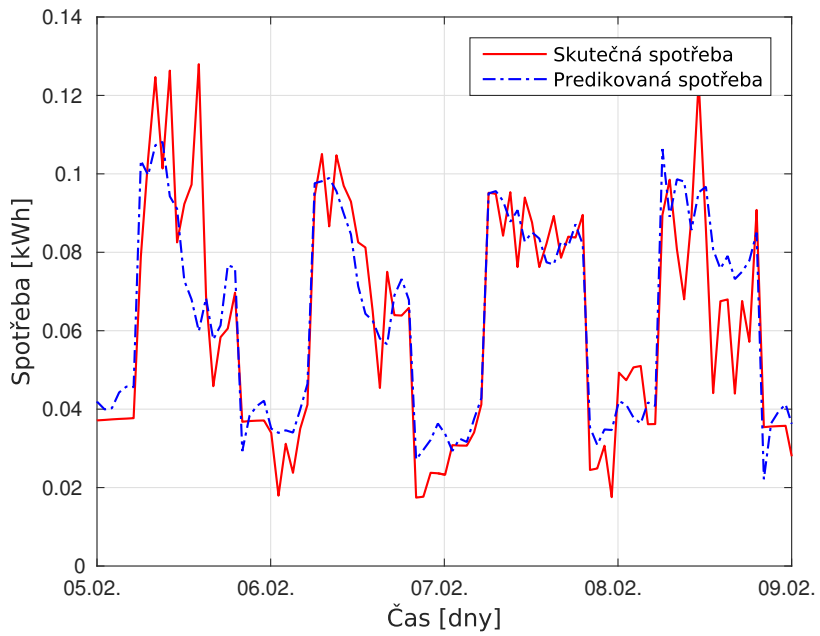
[11] P. Van Overschee and B. De Moor, "N4sid: Subspace algorithms for the identification of combined deterministic-stochastic systems," *Automatica*, vol. 30, no. 1, pp. 75–93, 1994.

Příloha B

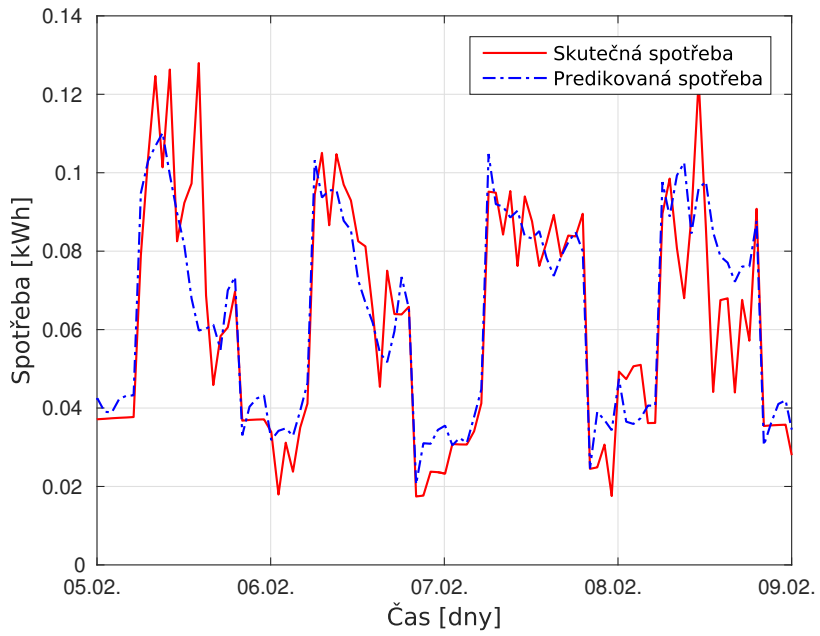
Grafy a obrázky



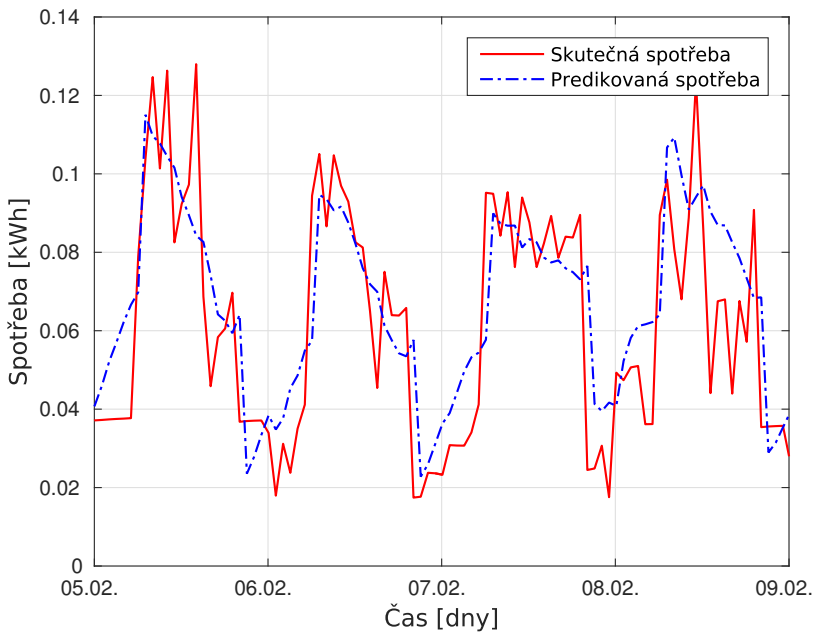
Obrázek B.1: Příklad simulovaných dat pro identifikaci
Hodinová spotřeba znázorňuje výstup a ostatní signály znamenají vstup.



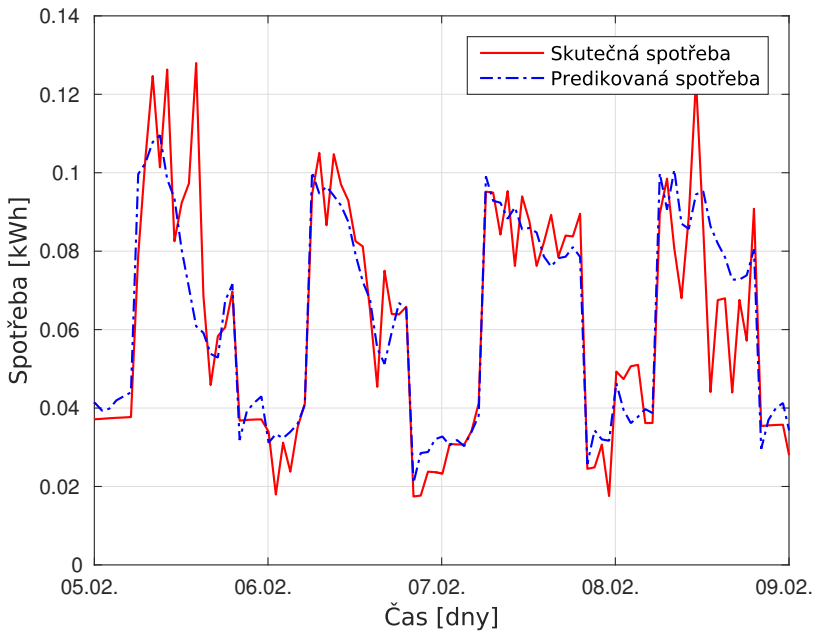
Obrázek B.2: ARX model
Příklad porovnání skutečné a predikované spotřeby modelem $arx110s$



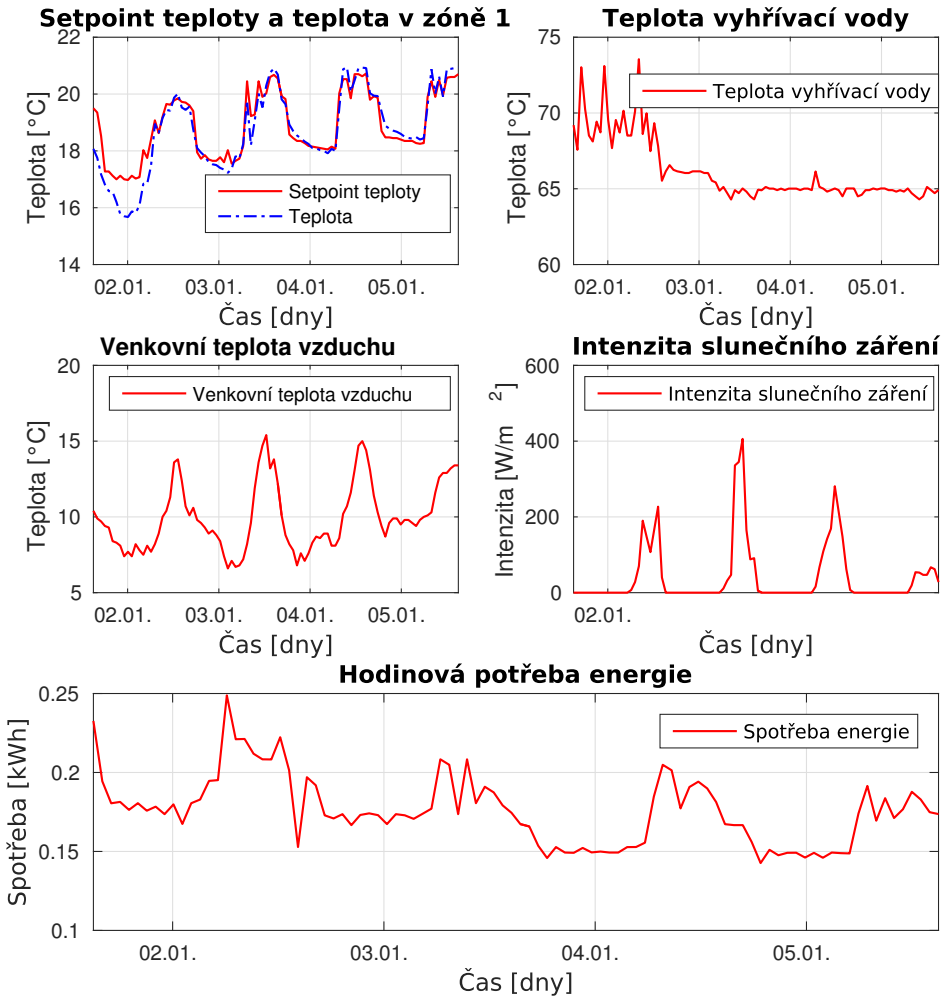
Obrázek B.3: ARMAX model
Příklad porovnání skutečné a predikované spotřeby modelem $armax2210s$



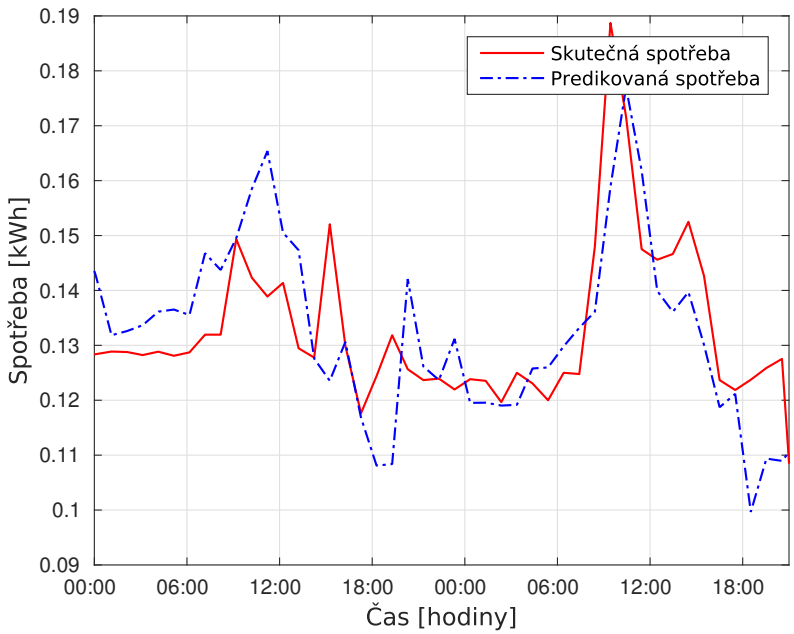
Obrázek B.4: State-Space model
Příklad porovnání skutečné a predikované spotřeby modelem $statespace1s$



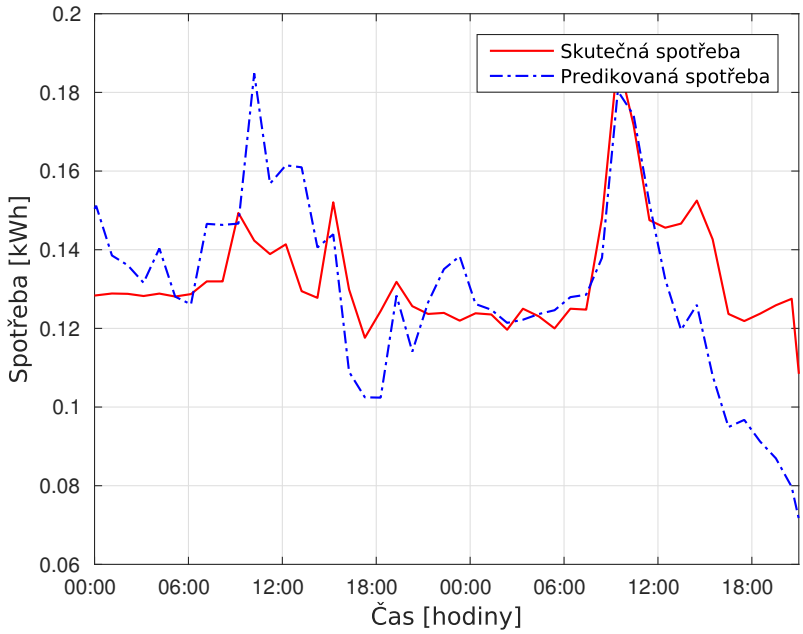
Obrázek B.5: NARX model
Příklad porovnání skutečné a predikované spotřeby modelem $narx110ss$



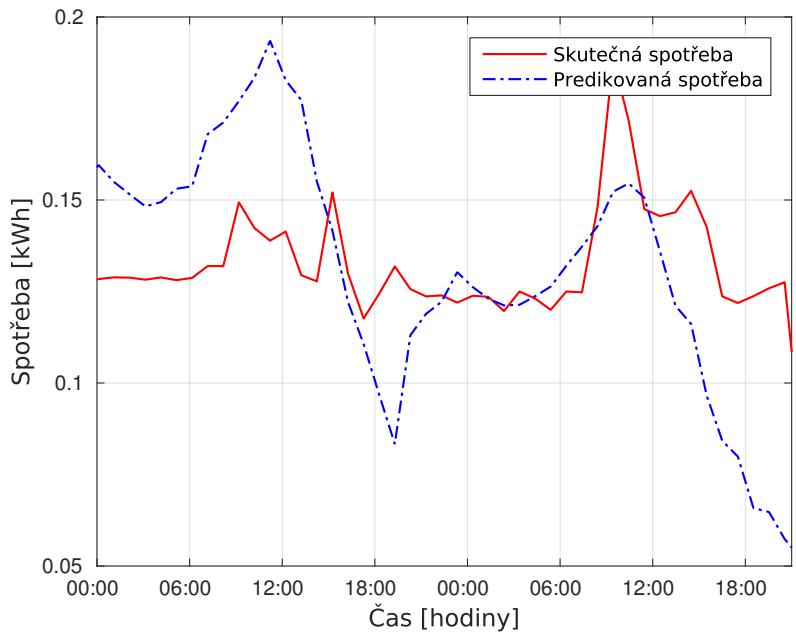
Obrázek B.6: Příklad reálných dat pro identifikaci



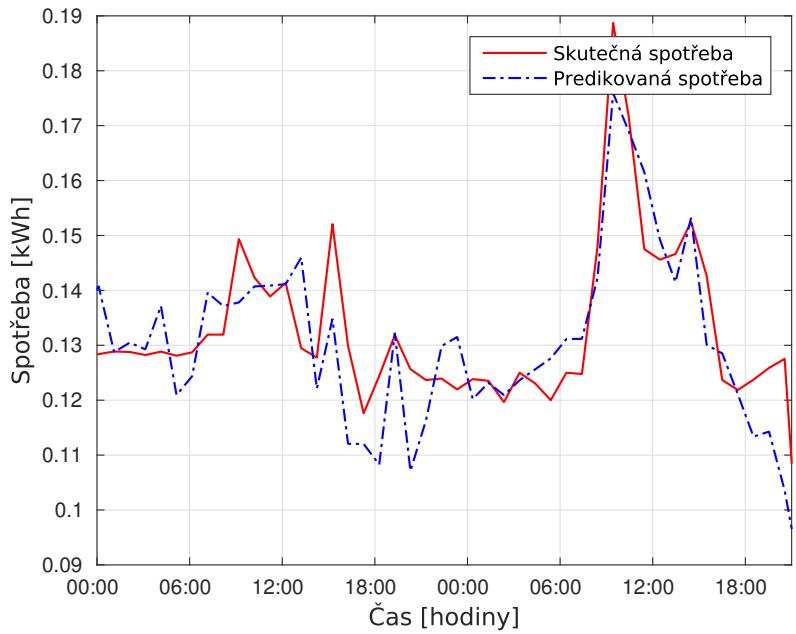
Obrázek B.7: ARX model
Příklad porovnání skutečné a predikované spotřeby modelem $arx120r$



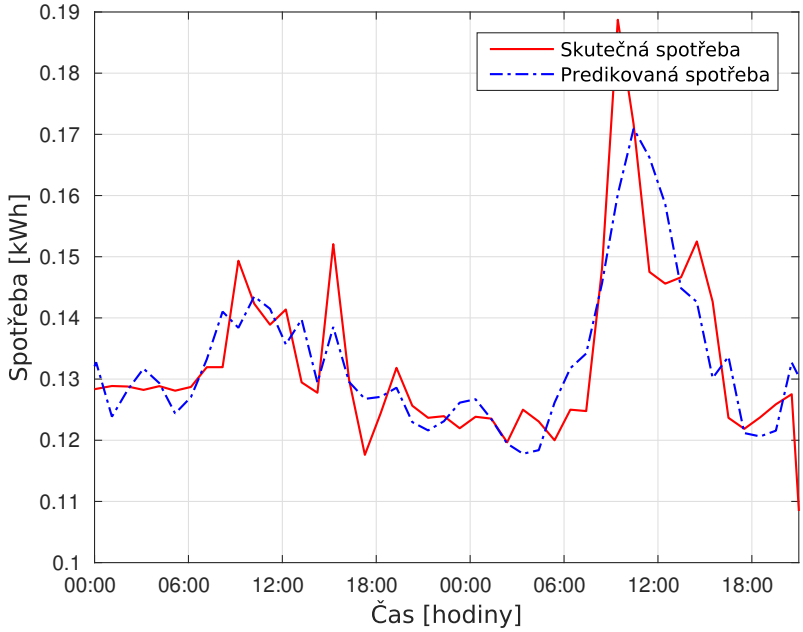
Obrázek B.8: ARMAX model
Příklad porovnání skutečné a predikované spotřeby modelem $armax1210r$



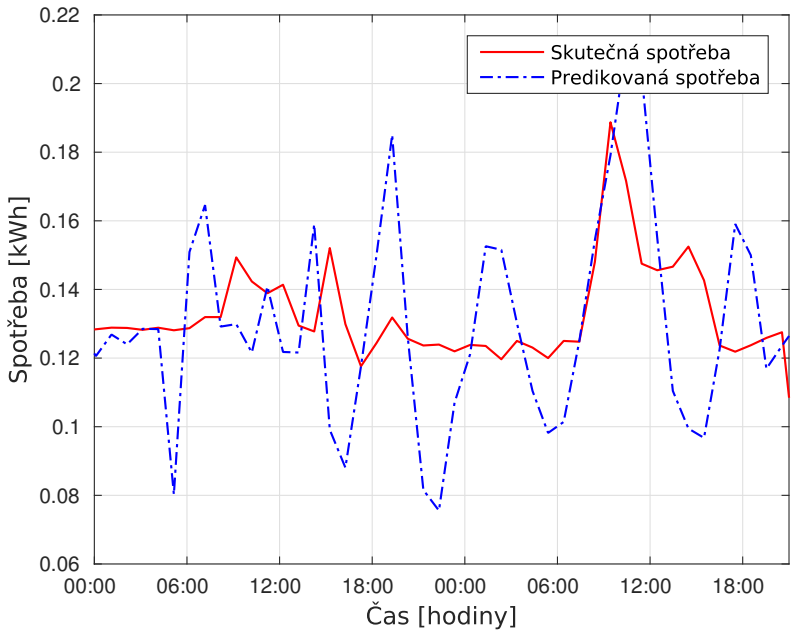
Obrázek B.9: State-Space model
Příklad porovnání skutečné a predikované spotřeby modelem $statespace1r$



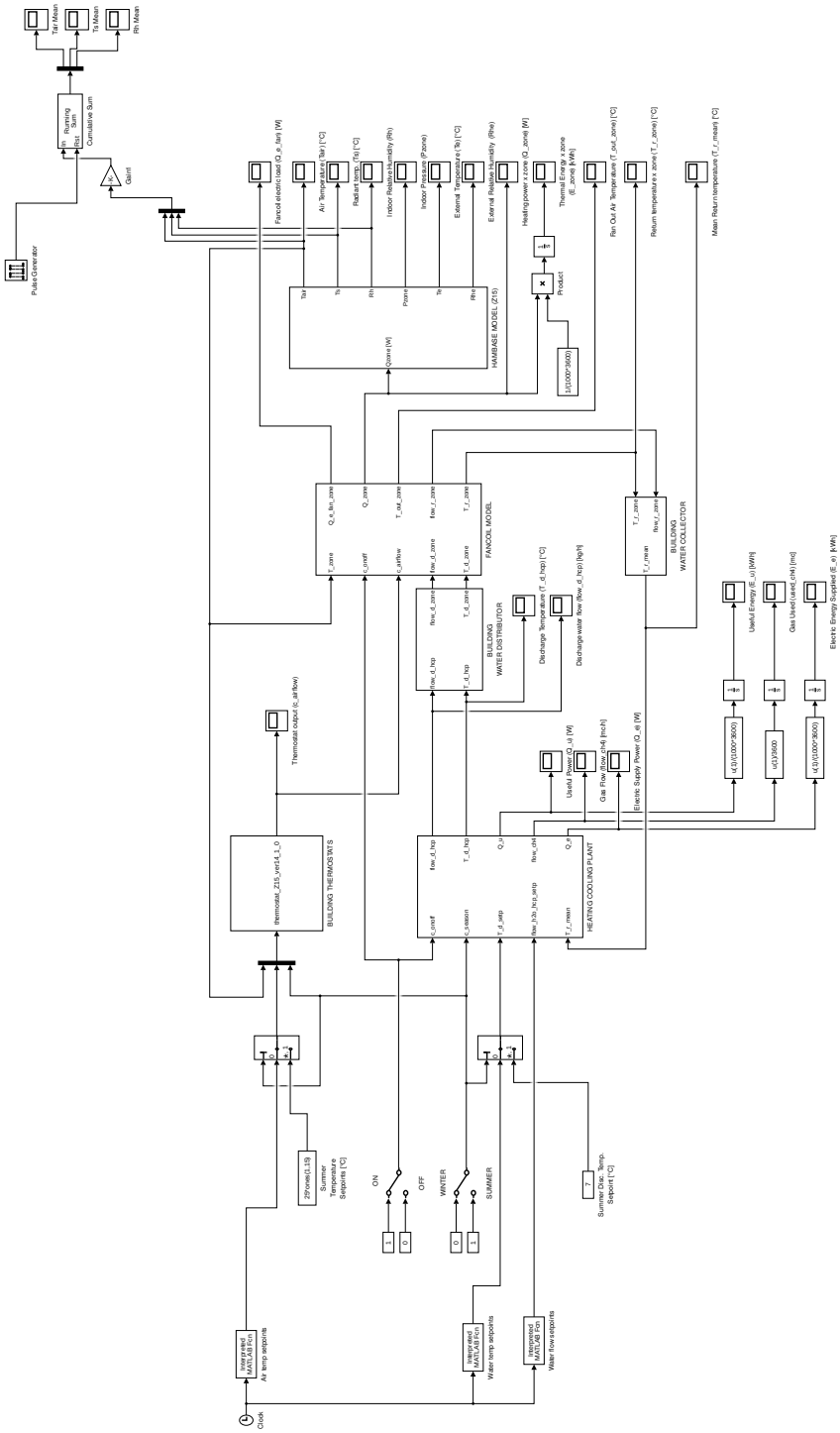
Obrázek B.10: NARX model
Příklad porovnání skutečné a predikované spotřeby modelem $narx220tr$



Obrázek B.11: Output-Error model
Příklad porovnání skutečné a predikované spotřeby modelem $oe120r$



Obrázek B.12: Hammerstein-Wiener model
Příklad porovnání skutečné a predikované spotřeby modelem $hww120pr$



Obrázek B.13: Simulinkový model

Příloha C

Tabulky

Označení	Popis	Jednotka
T_{s1} - T_{s10}	Setpoint teploty vzduchu v zóně 1-10	$[°C]$
T_1 - T_{10}	Teplota vzduchu v zóně 1-10	$[°C]$
T_w	Teplota vytápěcí vody	$[°C]$
W_1	Venkovní teplota vzduchu	$[°C]$
W_2	Venkovní relativní vlhkost	$[\%]$
W_3	Atmosférický tlak	$[mbar]$
W_4	Rychlost větru	$[ms^{-1}]$
W_5	Směr větru	[° od severu]
W_6	Množství srážek	$[mm]$
W_7	Intenzita slunečního záření	$[W/m^2]$
Q	Hodinová spotřeba pro vytápění	$[kWh]$

Tabulka C.1: Struktura reálných dat

České vysoké učení technické v Praze
Fakulta elektrotechnická

katedra řídicí techniky

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

Student: **Tran Anh Duc**

Studijní program: Kybernetika a robotika
Obor: Systémy a řízení

Název tématu: **Model vytápění budovy**

Pokyny pro vypracování:

K dispozici dostanete Simulinkový model vytápění budovy, data získaná z tohoto simulátoru a také data získaná z reálné budovy.

1. Vytvořte a vyhodnoťte black box model spotřeby tepelné energie založený na dodaných nebo vámi simulovaných datech.
2. Použijte simulovaná data ze Simulinkového modelu vytápění budovy a také z reálné budovy.
3. Vyberte vhodné metody modelování.
4. Pokuste se použít a evaluovat podobnou metodologii i na reálná data.

Seznam odborné literatury:

[1] Petr, Noskievic, Modelování a identifikace systémů.
[2] S. Haykin (1993): Neural Networks, A Comprehensive Foundation.

Vedoucí: Ing. Martin Macaš, Ph.D.

Platnost zadání: do konce letního semestru 2015/2016

L.S.

prof. Ing. Michael Šebek, DrSc.
vedoucí katedry

prof. Ing. Pavel Ripka, CSc.
děkan

V Praze dne 23. 1. 2015